STATISTICAL MECHANICS

SECOND EDITION

R. K. PATHRIA
Department of Physics,
University of Waterloo, Waterloo, Ontario, Canada

BUTTERWORTH
HEINEMANN

OXFORD AUCKLAND BOSTON JOHANNESBURG MELBOURNE NEW DELHI

Butterworth-Heinemann
Linacre House, Jordan Hill, Oxford OX2 8DP
225 Wildwood Avenue, Woburn, MA 01801-2041
A division of Reed Educational and Professional Publishing Ltd

A member of the Reed Elsevier plc group

First published 1972
Second edition 1996
Reprinted 1997, 1998, 1999, 2000, 2001

British Library Cataloguing in Publication Data
Pathria, R. K.
Statistical mechanics 2nd ed.
1. Statistical mechanics
I. Title
530.1'3

ISBN 0 7506 2469 8

Library of Congress Cataloguing in Publication Data
Pathria, R. K.
Statistical mechanics/R. K. Pathria - 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0 7506 2469 8
1. Statistical mechanics I. Title
Q175.P35 96-1679
530.1'33-dc20 CIP

Typeset by Laser Words, Madras, India
Printed and bound in Great Britain by MPG Books Ltd, Bodmin, Cornwall

CONTENTS

PREFACE TO THE SECOND EDITION

THE first edition of this book was prepared over the years 1966–70 when the subject of phase transitions was undergoing a complete overhaul. The concepts of scaling and universality had just taken root but the renormalization group approach, which converted these concepts into a calculational tool, was still obscure. Not surprisingly, my text of that time could not do justice to these emerging developments. Over the intervening years I have felt increasingly conscious of this rather serious deficiency in the text; so when the time came to prepare a new edition, my major effort went towards correcting that deficiency.

Despite the aforementioned shortcoming, the first edition of this book has continued to be popular over the last twenty years or so. I, therefore, decided not to tinker with it unnecessarily. Nevertheless, to make room for the new material, I had to remove some sections from the present text which, I felt, were not being used by the readers as much as the rest of the book was. This may turn out to be a disappointment to some individuals but I trust they will understand the logic behind it and, if need be, will go back to a copy of the first edition for reference. I, on my part, hope that a good majority of the users will not be inconvenienced by these deletions. As for the material retained, I have confined myself to making only editorial changes. The subject of phase transitions and critical phenomena, which has been my main focus of revision, has been treated in three new chapters that provide a respectable coverage of the subject and essentially bring the book up to date. These chapters, along with a collection of over sixty homework problems, will, I believe, enhance the usefulness of the book for both students and instructors.

The completion of this task has left me indebted to many. First of all, as mentioned in the Preface to the first edition, I owe a considerable debt to those who have written on this subject before and from whose writings I have benefitted greatly. It is difficult to thank them all individually; the bibliography at the end of the book is an obvious tribute to them. As for definitive help, I am most grateful to Dr Surjit Singh who advised me expertly and assisted me generously in the selection of the material that comprises Chapters 11–13 of the new text; without his help, the final product might not have been as coherent as it now appears to be. On the technical side, I am very thankful to Mrs Debbie Guenther who typed the manuscript with exceptional skill and proof-read it with extreme care; her task was clearly an arduous one but she performed it with good cheer — for which I admire her greatly.

Finally, I wish to express my heart-felt appreciation for my wife who let me devote myself fully to this task over a rather long period of time and waited for its completion ungrudgingly.

R.K.P.
Waterloo, Ontario, Canada

PREFACE TO THE FIRST EDITION

THIS book has arisen out of the notes of lectures that I gave to the graduate students at the McMaster University (1964–5), the University of Alberta (1965–7), the University of Waterloo (1969–71) and the University of Windsor (1970–1). While the subject matter, in its finer details, has changed considerably during the preparation of the manuscript, the style of presentation remains the same as followed in these lectures.

Statistical mechanics is an indispensable tool for studying physical properties of matter "in bulk" on the basis of the dynamical behavior of its "microscopic" constituents. Founded on the well-laid principles of *mathematical statistics* on one hand and *hamiltonian mechanics* on the other, the formalism of statistical mechanics has proved to be of immense value to the physics of the last 100 years. In view of the universality of its appeal, a basic knowledge of this subject is considered essential for every student of physics, irrespective of the area(s) in which he/she may be planning to specialize. To provide this knowledge, in a manner that brings out the essence of the subject with due rigor but without undue pain, is the main purpose of this work.

The fact that *the dynamics of a physical system is represented by a set of quantum states* and the assertion that *the thermodynamics of the system is determined by the multiplicity of these states* constitute the basis of our treatment. The fundamental connection between the microscopic and the macroscopic descriptions of a system is uncovered by investigating the conditions for equilibrium between two physical systems in thermodynamic contact. This is best accomplished by working in the spirit of the quantum theory right from the beginning; the entropy and other thermodynamic variables of the system then follow in a most natural manner. After the formalism is developed, one may (if the situation permits) go over to the limit of the classical statistics. This message may not be new, but here I have tried to follow it as far as is reasonably possible in a textbook. In doing so, an attempt has been made to keep the level of presentation fairly uniform so that the reader does not encounter fluctuations of too wild a character.

This text is confined to the study of the *equilibrium states* of physical systems and is intended to be used for a *graduate course* in statistical mechanics. Within these bounds, the coverage is fairly wide and provides enough material for tailoring a good two-semester course. The final choice always rests with the individual instructor; I, for one, regard Chapters 1–9 (*minus* a few sections from these chapters *plus* a few sections from Chapter 13) as the "essential part" of such a course. The contents of Chapters 10–12 are relatively advanced (not necessarily difficult); the choice of material out of these chapters will depend entirely on the taste of the instructor. To facilitate the understanding of the subject, the text has been illustrated with a large number of graphs; to assess the understanding, a large number of problems have been included. I hope these features are found useful.

I feel that one of the most essential aspects of teaching is to arouse the curiosity of the students in their subject, and one of the most effective ways of doing this is to discuss with them (in a reasonable measure, of course) the circumstances that led to the emergence of the subject. One would, therefore, like to stop occasionally to reflect upon the manner in which the various developments really came about; at the same time, one may not like the flow of the text to be hampered by the discontinuities arising from an intermittent addition of historical material. Accordingly, I decided to include in this account an Historical Introduction to the subject which stands separate from the main text. I trust the readers, especially the instructors, will find it of interest.

For those who wish to continue their study of statistical mechanics beyond the confines of this book, a fairly extensive bibliography is included. It contains a variety of references — old as well as new, experimental as well as theoretical, technical as well as pedagogical. I hope that this will make the book useful for a wider readership.

The completion of this task has left me indebted to many. Like most authors, I owe considerable debt to those who have written on the subject before. The bibliography at the end of the book is the most obvious tribute to them; nevertheless, I would like to mention, in particular, the works of the Ehrenfests, Fowler, Guggenheim, Schrödinger, Rushbrooke, ter Haar, Hill, Landau and Lifshitz, Huang, and Kubo, which have been my constant reference for several years and have influenced my understanding of the subject in a variety of ways. As for the preparation of the text, I am indebted to Robert Teshima who drew most of the graphs and checked most of the problems, to Ravindar Bansal, Vishwa Mittar and Surjit Singh who went through the entire manuscript and made several suggestions that helped me unkink the exposition at a number of points, to Mary Annetts who typed the manuscript with exceptional patience, diligence and care, and to Fred Hetzel, Jim Briante and Larry Kry who provided technical help during the preparation of the final version.

As this work progressed I felt increasingly gratified towards Professors F. C. Auluck and D. S. Kothari of the University of Delhi with whom I started my career and who initiated me into the study of this subject, and towards Professor R. C. Majumdar who took keen interest in my work on this and every other project that I have undertaken from time to time. I am grateful to Dr D. ter Haar of the University of Oxford who, as the general editor of this series, gave valuable advice on various aspects of the preparation of the manuscript and made several useful suggestions towards the improvement of the text. I am thankful to Professors J.W. Leech, J. Grindlay and A.D. Singh Nagi of the University of Waterloo for their interest and hospitality that went a long way in making this task a pleasant one.

The final tribute must go to my wife whose cooperation and understanding, at all stages of this project and against all odds, have been simply overwhelming.

R.K.P.
Waterloo, Ontario, Canada

HISTORICAL INTRODUCTION

STATISTICAL mechanics is a formalism which aims at explaining the physical properties of matter *in bulk* on the basis of the dynamical behavior of its *microscopic* constituents. The scope of the formalism is almost as unlimited as the very range of the natural phenomena, for in principle it is applicable to matter in any state whatsoever. It has, in fact, been applied, with considerable success, to the study of matter in the solid state, the liquid state or the gaseous state, matter composed of several phases and/or several components, matter under extreme conditions of density and temperature, matter in equilibrium with radiation (as, for example, in astrophysics), matter in the form of a biological specimen, etc. Furthermore, the formalism of statistical mechanics enables us to investigate the *non-equilibrium* states of matter as well as the *equilibrium* states; indeed, these investigations help us to understand the manner in which a physical system that happens to be "out of equilibrium" at a given time t approaches a "state of equilibrium" as time passes.

In contrast with the present status of its development, the success of its applications and the breadth of its scope, the beginnings of statistical mechanics were rather modest. Barring certain primitive references, such as those of Gassendi, Hooke, etc., the real work started with the contemplations of Bernoulli (1738), Herapath (1821) and Joule (1851) who, in their own individual ways, attempted to lay a foundation for the so-called *kinetic theory of gases*–a discipline that finally turned out to be the forerunner of statistical mechanics. The pioneering work of these investigators established the fact that the pressure of a gas arose from the motion of its molecules and could be computed by considering the dynamical influence of molecular bombardment on the walls of the container. Thus, Bernoulli and Herapath could show that, if temperature remained constant, the pressure P of an ordinary gas was inversely proportional to the volume V of the container (Boyle's law), and that it was essentially independent of the shape of the container. This, of course, involved the explicit assumption that, *at a given temperature* T, the (mean) speed of the molecules is independent of both pressure and volume. Bernoulli even attempted to determine the (first-order) correction to this law, arising from the *finite* size of the molecules, and showed that the volume V appearing in the statement of the law should be replaced by $(V - b)$, where b is the "actual" volume of the molecules.[1] Joule was the first to show that the pressure P is directly proportional to the square of the molecular speed c, which he had assumed to be the same for all molecules. Krönig (1856) went a step further. Introducing the "quasi-statistical" assumption that, *at any time* t, one-sixth of the molecules could be assumed to be flying in each of the six "independent" directions, namely $+x$, $-x$, $+y$, $-y$, $+z$

and $-z$, he derived the equation

$$P = \tfrac{1}{3} n m c^2, \tag{1}$$

where n is the number density of the molecules and m the molecular mass. Krönig, too, assumed the molecular speed c to be the same for all molecules; from (1), he inferred that the kinetic energy of the molecules should be directly proportional to the absolute temperature of the gas.

Krönig justified his method in these words: "The path of each molecule must be so irregular that it will defy all attempts at calculation. However, according to the laws of probability, one could assume a completely regular motion in place of a completely irregular one!" It must, however, be noted that it is only because of the special form of the summations appearing in the calculation of the pressure that Krönig's model leads to the same result as the one following from more refined models. In other problems, such as the ones involving diffusion, viscosity or heat conduction, this is no longer the case.

It was at this stage that Clausius entered the field. First of all, in 1857, he derived the ideal-gas law under assumptions far less stringent than Krönig's. He discarded both leading assumptions of Krönig and showed that eqn. (1) was still true; of course, c^2 now became the *mean square speed* of the molecules. In a later paper (1859), Clausius introduced the concept of the *mean free path* and thus became the first to analyze transport phenomena. It was in these studies that he introduced the famous "Stosszahlansatz"–the hypothesis on the number of collisions (among the molecules)–which, later on, played a prominent role in the monumental work of Boltzmann.[2] With Clausius, the introduction of the microscopic and statistical points of view into the physical theory was definitive, rather than speculative. Accordingly, Maxwell, in a popular article entitled "Molecules", written for the *Encyclopedia Britannica*, referred to him as the "principal founder of the kinetic theory of gases", while Gibbs, in his Clausius obituary notice, called him the "father of statistical mechanics".[3]

The work of Clausius attracted Maxwell to the field. He made his first appearance with the memoir "Illustrations in the dynamical theory of gases" (1860), in which he went much further than his predecessors by deriving his famous law of "distribution of molecular speeds". Maxwell's derivation was based on elementary principles of probability and was clearly inspired by the Gaussian law of "distribution of random errors". A derivation based on the requirement that "the equilibrium distribution of molecular speeds, once acquired, should remain invariant under molecular collisions" appeared in 1867. This led Maxwell to establish what is known as *Maxwell's transport equation* which, if skillfully used, leads to the same results as one gets from the more fundamental equation due to Boltzmann.[4]

Maxwell's contributions to the subject diminished considerably after his appointment, in 1871, as the Cavendish Professor at Cambridge. By that time Boltzmann had already made his first strides. In the period 1868–71 he generalized Maxwell's distribution law to polyatomic gases, also taking into account the presence of external forces, if any; this gave rise to the famous *Boltzmann factor* $\exp(-\beta\varepsilon)$, where ε denotes the *total* energy of a molecule. These investigations also led to the *equipartition theorem*. Boltzmann further showed that, just like the original distribution of Maxwell, the generalized distribution (which we now call the *Maxwell–Boltzmann distribution*) is stationary with respect to molecular collisions. In 1872 came the celebrated *H-theorem* which provided a molecular basis for the natural tendency of physical systems to approach, and stay in, a state of equilibrium. This established a connection between the microscopic approach (which characterizes statistical mechanics) and the phenomenological approach (which characterized thermodynamics) much more transparently than ever before; it also provided a direct method for computing the entropy of a given physical system from purely microscopic considerations. As a corollary to the H-theorem, Boltzmann showed that the Maxwell–Boltzmann distribution is the *only* distribution that stays invariant under molecular collisions and that any other distribution, under the influence of molecular collisions, ultimately goes over to a Maxwell–Boltzmann distribution. In 1876 Boltzmann derived his famous transport equation which, in the hands of Chapman and Enskog (1916–17), has proved to be an extremely powerful tool for investigating macroscopic properties of systems in non-equilibrium states.

Things, however, proved quite harsh for Boltzmann. His H-theorem, and the consequent *irreversible* behavior of physical systems, came under heavy attack, mainly from Loschmidt (1876–77) and Zermelo (1896). While Loschmidt wondered how the consequences of this theorem could be reconciled with the reversible character of the basic equations of motion of the molecules, Zermelo wondered how these consequences could be made to fit with the *quasi-periodic* behavior of closed systems (which arose in view of the so-called Poincaré cycles). Boltzmann defended himself against these attacks with all his might but could not convince his opponents of the correctness of his work. At the same time, the energeticists, led by Mach and Ostwald, were criticizing the very (molecular) basis of the kinetic theory,[5] while Kelvin was emphasizing the "nineteenth-century clouds hovering over the dynamical theory of light and heat".[6]

All this left Boltzmann in a state of despair and induced in him a persecution complex.[7] He wrote in the introduction to the second volume of his treatise *Vorlesungen über Gastheorie* (1898):[8]

> I am convinced that the attacks (on the kinetic theory) rest on misunderstandings and that the role of the kinetic theory is not yet played out. In my opinion it would be a blow to science if contemporary opposition were to cause kinetic theory to sink into the oblivion which was the fate suffered by the wave theory of light through the authority of Newton. I am aware of the weakness of one individual against the prevailing currents of opinion. In order to insure that not too much will have to be rediscovered when people return to the study of kinetic theory I will present the most difficult and misunderstood parts of the subject in as clear a manner as I can.

We shall not dwell any further on the kinetic theory; we would rather move onto the development of the more sophisticated approach known as the *ensemble theory*, which may in fact be regarded as the statistical mechanics proper.[9] In this approach, the dynamical state of a given system, as characterized by the generalized coordinates q_i and the generalized momenta p_i, is represented by a *phase point* $G(q_i, p_i)$ in a *phase space* of appropriate dimensionality. The evolution of the dynamical state in time is depicted by the *trajectory* of the G-point in the phase space [illegible] "geometry" of the trajectory being governed by the equations of motion of the system and by the nature of the physical constraints imposed on it. To develop an appropriate formalism, one considers the given system along with an infinitely large number of "mental copies" ther[illegible] i.e. an *ensemble* of similar systems under identical physical constraints (though, at any time t, the various systems in the ensemble would differ widely in respect of their dynamical states). In the phase space, then, one has a swarm of infinitely many G-points (which, at any time t, are widely dispersed and, with time, move along their respective trajectories). The fiction of a host of infinitely many, identical but independent, systems allows one to replace certain dubious assumptions of the kinetic theory of gases by readily acceptable statements of statistical mechanics. The explicit formulation of these statements was first given by Maxwell (1879) who on this occasion used the word "statistico-mechanical" to describe the study of ensembles (of gaseous systems) — though, eight years earlier, Boltzmann (1871) had worked with essentially the same kind of ensembles.

The most important quantity in the ensemble theory is the *density function*, $\rho(q, p; t)$, of the G-points in the phase space; a stationary distribution ($\partial\rho/\partial t = 0$) characterizes a *stationary ensemble*, which in turn represents a system *in equilibrium*. Maxwell and Boltzmann confined their study to ensembles for which the function ρ depended solely on the energy E of the system. This included the special case of *ergodic* systems which were so defined that "the undisturbed motion of such a system, if pursued for an unlimited time, would ultimately traverse (the neighborhood of) each and every phase point compatible with the *fixed* value E of the energy". Consequently, the *ensemble average*, $\langle f \rangle$, of a physical quantity f, taken at *any* time t, would be the same as the *long-time average*, $\overline{f}$, pertaining to *any* member of the ensemble. Now, $\overline{f}$ is the value we expect to obtain for the quantity in question when we make an appropriate measurement on the system; the result of this measurement should, therefore, agree with the theoretical estimate $\langle f \rangle$. We thus acquire a recipe to bring about a direct contact between theory and experiment. At the same time, we lay down a rational basis for a microscopic theory of matter as an alternative to the empirical approach of thermodynamics!

A significant advance in this direction was made by Gibbs who, with his *Elementary Principles of Statistical Mechanics* (1902), turned ensemble theory into a most efficient tool for the theorist. He emphasized the use of "generalized" ensembles and developed schemes which, in principle, enabled one to compute a complete set of thermodynamic quantities of a given system from purely mechanical properties of its microscopic constituents.[10] In its methods and results, the work of Gibbs turned out to be much more general than any preceding treatment of the subject; it applied to any physical system that met the simple-minded requirements that (i) it was mechanical in structure and (ii) it obeyed Lagrange's and Hamilton's equations of motion. In this respect, Gibbs's work may be considered to have accomplished as much for thermodynamics as Maxwell's had for electrodynamics.

These developments almost coincided with the great revolution that Planck's work of 1900 brought into physics. As is well known, Planck's *quantum hypothesis* successfully resolved the essential mysteries of the black-body radiation–a subject where the three best-established disciplines of the nineteenth century, viz. mechanics, electrodynamics and thermodynamics, were all focused. At the same time, it uncovered both the strengths and the weaknesses of these disciplines. It would have been surprising if statistical mechanics, which links thermodynamics with mechanics, could have escaped the repercussions of this revolution.

The subsequent work of Einstein (1905a) on the photoelectric effect and of Compton (1923a,b) on the scattering of x-rays established, so to say, the "existence" of the *quantum of radiation*, or the *photon* as we now call it.[11] It was then natural that someone tried to derive Planck's radiation formula by treating black-body radiation as a *gas of photons* in much the same way as Maxwell had derived his law of distribution (of molecular speeds) for a gas of conventional molecules. But, then, does a gas of photons differ so radically from a gas of conventional molecules that the two laws of distribution should be so different from one another?

The answer to this question came from the manner in which Planck's formula was derived by Bose. In his historic paper of 1924, Bose treated black-body radiation as a gas of photons; however, instead of considering the allocation of the "individual" photons to the various energy states of the system, he fixed his attention on the number of states that contained "a particular number" of photons. Einstein, who seems to have translated Bose's paper into German from an English manuscript sent to him by the author, at once recognized the importance of this approach and added the following note to his translation: "Bose's derivation of Planck's formula is in my opinion an important step forward. The method employed here would also yield the quantum theory of an ideal gas, which I propose to demonstrate elsewhere."

Implicit in Bose's approach was the fact that in the case of photons what really mattered was "the set of numbers (of photons) in various energy states of the system" and not the specification as to "which photon was in which state"; in other words, photons were *mutually indistinguishable*. Einstein argued that what Bose had implied for photons should be true for material particles as well (for the property of indistinguishability arose essentially from the wave character of these entities and, according to de Broglie, material particles also possessed that character).[12] In two papers, which appeared soon after, Einstein (1924, 1925) applied this method to the study of an ideal gas and thereby developed what we now call *Bose–Einstein statistics*. In the second of these papers, the fundamental difference between the new statistics and the classical *Maxwell–Boltzmann* statistics comes out transparently in terms of the indistinguishability of the molecules.[13] In the same paper Einstein discovered the phenomenon of *Bose–Einstein condensation* which, thirteen years later, was adopted by London (1938a,b) as the basis for a microscopic understanding of the curious properties of liquid He^4 at low temperatures.

Following the enunciation of Pauli's exclusion principle (1925), Fermi (1926) showed that certain physical systems would obey a different kind of statistics, viz. the *Fermi–Dirac statistics*, in which not more than one particle could occupy the same energy state ($n_i = 0, 1$). It seems important to mention here that Bose's method of 1924 leads to the Fermi–Dirac distribution as well, provided that one limits the occupancy of an energy state to *at most* one particle.[14]

Soon after its appearance, the Fermi–Dirac statistics were applied by Fowler (1926) to discuss the equilibrium states of white dwarf stars and by Pauli (1927) to explain the weak, temperature-independent paramagnetism of alkali metals; in each case, one had to deal with a "highly degenerate" gas of electrons which obey Fermi–Dirac statistics. In the wake of this, Sommerfeld produced his monumental work of 1928 which not only put the electron theory of metals on a physically secure foundation but also gave it a fresh start in the right direction. Thus, Sommerfeld could explain practically all the major properties of metals that arose from conduction electrons and, in each case, obtained results which showed much better agreement with experiment than the ones following from the classical theories of Riecke (1898), Drude (1900) and Lorentz (1904–05). Around the same time, Thomas (1927) and Fermi (1928) investigated the electron distribution in heavier atoms and obtained theoretical estimates for the relevant binding energies; these investigations led to the development of the so-called *Thomas–Fermi model* of the atom, which was later extended so that it could be applied to molecules, solids and nuclei as well.[15]

Thus, the whole structure of statistical mechanics was overhauled by the introduction of the concept of indistinguishability of (identical) particles.[16] The statistical aspect of the problem, which was already there in view of the large number of particles present, was now augmented by another statistical aspect that arose from the probabilistic nature of the wave mechanical description. One had, therefore, to carry out a *two-fold* averaging of the dynamical variables over the states of the given system in order to obtain the relevant expectation values. That sort of a situation was bound to necessitate a reformulation of the ensemble theory itself, which was carried out step by step. First, Landau (1927) and von Neumann (1927) introduced the so-called *density matrix*, which was the quantum-mechanical analogue of the *density function* of the classical phase space; this was discussed, both from statistical and quantum-mechanical points of view, by Dirac (1929–31). Guided by the classical ensemble theory, these authors considered both *microcanonical* and *canonical* ensembles; the introduction of *grand canonical* ensembles in quantum statistics was made by Pauli (1927).[17]

The important question as to which particles would obey Bose–Einstein statistics and which Fermi–Dirac remained theoretically unsettled until Belinfante (1939) and Pauli (1940) discovered the vital connection between spin and statistics.[18] It turns out that those particles whose spin is an integral multiple of $\hbar$ obey Bose–Einstein statistics while those whose spin is a half-odd integral multiple of $\hbar$ obey Fermi–Dirac statistics. To date, no third category of particles has been discovered.

Apart from the foregoing milestones, several notable contributions towards the development of statistical mechanics have been made from time to time; however, most of these contributions are concerned with the development or perfection of mathematical techniques which make application of the basic formalism to actual physical problems more fruitful. A review of these developments is out of place here; they will be discussed at their appropriate place in the text.

Notes

[1] As is well known, this "correction" was correctly evaluated, much later, by van der Waals (1873) who showed that, for large V, b is *four times* the "actual" volume of the molecules; see Problem 1.4.

[2] For an excellent review of this and related topics, see P. and T. Ehrenfest (1912).

[3] For further details, refer to Montroll (1963) where an account is also given of the pioneering work of Waterston (1846, 1892).

[4] This has been demonstrated in Guggenheim (1960) where t[illegible]efficients of viscosity, thermal conductivity and diffusion of a gas of hard spheres have been c[illegible]ted on the basis of Maxwell's transport equation.

[5] These critics were silenced by Einstein whose work on the Brownian motion (1905b) established atomic theory *once and for all.*

[6] The first of these clouds was concerned with the mysteries of the "aether", and was dispelled by the theory of relativity. The second was concerned with the inadequacy of the "equipartition theorem", and was dispelled by the quantum theory.

[7] Some people attribute Boltzmann's suicide on September 5, 1906 to this cause.

[8] Quotation from Montroll (1963).

[9] For a review of the historical development of kinetic theory leading to statistical mechanics, see Brush (1957, 1958, 1961, 1965–66).

[10] In much the same way as Gibbs, but quite independently of him, Einstein (1902, 1903) also developed the theory of ensembles.

[11] Strictly speaking, it might be somewhat misleading to cite Einstein's work on the photoelectric effect as a proof of the existence of photons. In fact, many of the effects (including the photoelectric effect), for which it seems necessary to invoke photons, can be explained away on the basis of a wave theory of radiation. The only phenomena for which photons seem indispensable are the ones involving *fluctuations*, such as the Hanbury Brown–Twiss effect or the Lamb shift. For the relevance of fluctuations to the problem of radiation, see ter Haar (1967, 1968).

[12] Of course, in the case of material particles, the total number N (of the particles) will also have to be conserved; this had not to be done in the case of photons. For details, see Sec. 6.1.

[13] It is here that one encounters the *correct* method of counting "the number of distinct ways in which g_i energy states can accommodate n_i particles", depending upon whether the particles are (i) distinguishable or (ii) indistinguishable. The occupancy of the individual states was, in each case, *unrestricted*, i.e. $n_i = 0, 1, 2, \ldots$.

[14] Dirac, who was the first to investigate the connection between statistics and wave mechanics, showed, in 1926, that the wave functions describing a system of identical particles obeying Bose–Einstein (or Fermi–Dirac) statistics must be symmetric (or antisymmetric) with respect to an interchange of two particles.

[15] For an excellent review of this model, see N. H. March (1957).

[16] Of course, in many a situation where the wave nature of the particles is not so important classical statistics continue to apply.

[17] A detailed treatment of this development has been given by Kramers (1938).

[18] See also Lüders and Zumino (1958).

CHAPTER 1

THE STATISTICAL BASIS OF THERMODYNAMICS

IN THE annals of thermal physics, the fifties of the last century mark a very definite epoch. By that time the science of thermodynamics, which grew essentially out of an experimental study of the macroscopic behavior of physical systems, had become, through the work of Carnot, Joule, Clausius and Kelvin, a secure and stable discipline of physics. The theoretical conclusions following from the first two laws of thermodynamics were found to be in very good agreement with the corresponding experimental results.[1] At the same time, the kinetic theory of gases, which aimed at explaining the macroscopic behavior of gaseous systems in terms of the motion of the molecules and had so far thrived more on speculation than calculation, began to emerge as a real, mathematical theory. Its initial successes were glaring; however, a real contact with thermodynamics could not be made until about 1872 when Boltzmann developed his H-theorem and thereby established a direct connection between entropy on one hand and molecular dynamics on the other. Almost simultaneously, the conventional (kinetic) theory began giving way to its more sophisticated successor–the ensemble theory. The power of the techniques that finally emerged reduced thermodynamics to the status of an "essential" consequence of the get-together of the *statistics* and the *mechanics* of the molecules constituting a given physical system. It was then natural to give the resulting formalism the name *Statistical Mechanics.*

As a preparation towards the development of the formal theory, we start with a few general considerations regarding the statistical nature of a macroscopic system. These considerations will provide ground for a statistical interpretation of thermodynamics. It may be mentioned here that, unless a statement is made to the contrary, the system under study is supposed to be in one of its equilibrium states.

1.1. The macroscopic and the microscopic states

We consider a physical system composed of N identical particles confined to a space of volume V. In a typical case, N would be an extremely large number–generally, of order 10^{23}. In view of this, it is customary to carry out analysis in the so-called *thermodynamic limit*, viz. $N \to \infty$, $V \to \infty$ (such that the ratio N/V, which represents the *particle density* n, stays fixed at a preassigned value). In this limit, the *extensive* properties of the system become directly proportional to the size of the system (i.e. proportional to N or to V), while the *intensive* properties become independent thereof; the particle density, of course, remains an important parameter for all physical properties of the system.

Next we consider the total energy E of the system. If the particles comprising the system could be regarded as non-interacting, the total energy E would be equal to the sum of the energies ε_i of the individual particles:

$$E = \sum_i n_i \varepsilon_i, \tag{1}$$

where n_i denotes the number of particles each with energy ε_i. Clearly,

$$N = \sum_i n_i. \tag{2}$$

According to quantum mechanics, the single-particle energies ε_i are discrete and their values depend crucially on the volume V to which the particles are confined. Accordingly, the possible values of the total energy E are also discrete. However, for large V, the spacing of the different energy values is so small in comparison with the total energy of the system that the parameter E might well be regarded as a *continuous* variable. This would be true even if the particles were mutually interacting; of course, in that case the total energy E cannot be written in the form (1).

The specification of the actual values of the parameters N, V and E then defines a *macrostate* of the given system.

At the molecular level, however, a large number of possibilities still exist because at that level there will *in general* be a large number of different ways in which the macrostate (N, V, E) of the given system can be realized. In the case of a non-interacting system, since the total energy E consists of a simple sum of the N single-particle energies ε_i, there will obviously be a large number of different ways in which the individual ε_i can be chosen so as to make the total energy equal to E. In other words, there will be a large number of different ways in which the total energy E of the system can be distributed among the N particles constituting it. Each of these (different) ways specifies a *microstate*, or *complexion*, of the given system. In general, the various microstates, or complexions, of a given system can be identified with the independent solutions $\psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$ of the Schrödinger equation of the system, corresponding to the eigenvalue E of the relevant Hamiltonian. In any case, to a given macrostate of the system there does in general correspond a large number of microstates and it seems natural to assume, when there are no other constraints, that at any time t the system is *equally likely* to be in any one of these microstates. This assumption forms the backbone of our formalism and is generally referred to as the postulate of "equal *a priori* probabilities" for all microstates consistent with a given macrostate.

The actual number of all possible microstates will, of course, be a function of N, V and E and may be denoted by the symbol $\Omega(N, V, E)$; the dependence on V comes in because the possible values ε_i of the single-particle energy ε are themselves a function of this parameter.[2] Remarkably enough, it is from the magnitude of the number Ω, and from its dependence on the parameters N, V and E, that complete thermodynamics of the given system can be derived!

We shall not stop here to discuss the ways in which the number $\Omega(N, V, E)$ can be computed; we shall do that only after we have developed our considerations sufficiently so that we can carry out further derivations from it. First we have to discover the manner in which this number is related to any of the leading thermodynamic quantities. To do this, we consider the problem of "thermal contact" between two given physical systems, in the hope that this consideration will bring out the true nature of the number Ω.

1.2. Contact between statistics and thermodynamics: physical significance of the number $\Omega(N, V, E)$

We consider two physical systems, A_1 and A_2, which are separately in equilibrium; see Fig. 1.1. Let the macrostate of A_1 be represented by the parameters N_1, V_1 and E_1, so that it has $\Omega_1(N_1, V_1, E_1)$ possible microstates, and the macrostate of A_2 be represented by the parameters N_2, V_2 and E_2, so that it has $\Omega_2(N_2, V_2, E_2)$ possible microstates. The mathematical form of the function Ω_1 may not be the same as that of the function Ω_2, because that ultimately depends upon the nature of the system. We do, of course, believe that all thermodynamic properties of the systems A_1 and A_2 can be derived from the functions $\Omega_1(N_1, V_1, E_1)$ and $\Omega_2(N_2, V_2, E_2)$, respectively.

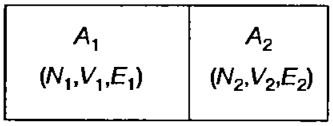


FIG. 1.1. Two physical systems, being brought into thermal contact.

We now bring the two systems into thermal contact with each other, thus allowing the possibility of exchange of energy between the two; this can be done by sliding in a conducting wall and removing the impervious one. For simplicity, the two systems are still separated by a rigid, impenetrable wall, so that the respective volumes V_1 and V_2 and the respective particle numbers N_1 and N_2 remain fixed. The energies E_1 and E_2, however, become variable and the only condition that restricts their variation is

$$E^{(0)} = E_1 + E_2 = \text{const.} \tag{1}$$

Here, $E^{(0)}$ denotes the energy of the composite system $A^{(0)}(\equiv A_1 + A_2)$; the energy of interaction between A_1 and A_2, if any, is being neglected. Now, at any time t, the sub-system A_1 is equally likely to be in any one of the $\Omega_1(E_1)$ microstates while the sub-system A_2 is equally likely to be in any one of the $\Omega_2(E_2)$ microstates; therefore, the composite system $A^{(0)}$ is equally likely to be in any one of the $\Omega^{(0)}(E_1, E_2)$ microstates, where

$$\Omega^{(0)}(E_1, E_2) = \Omega_1(E_1)\Omega_2(E_2) = \Omega_1(E_1)\Omega_2(E^{(0)} - E_1) = \Omega^{(0)}(E^{(0)}, E_1), \text{ say.}$$
(2)[3]

Clearly, the number $\Omega^{(0)}$ itself varies with E_1. The question now arises: at what value of the variable E_1 will the composite system be in equilibrium? In other words, how far will the energy exchange go in order to bring the sub-systems A_1 and A_2 into mutual equilibrium?

We assert that this will happen at that value of E_1 which *maximizes* the number $\Omega^{(0)}(E^{(0)}, E_1)$. The philosophy behind this assertion is that a physical system, left to itself, proceeds naturally in a direction that enables it to assume an ever-increasing number of microstates until it finally settles down in a macrostate that affords the *largest possible* number of microstates. Statistically speaking, we regard a macrostate with a larger number of microstates as a more probable state, and the one with the largest number of microstates as the most probable one. Detailed studies show that, for a typical system, the number of microstates pertaining to any macrostate that departs even slightly from the most probable one is "orders of magnitude" smaller than the number pertaining to the latter. Thus, the most probable state of a system is *the* state in which the system spends an "overwhelmingly" large fraction of its time. It is then natural to identify this state with the *equilibrium* state of the system.

Denoting the equilibrium value of E_1 by $\overline{E}_1$ (and that of E_2 by $\overline{E}_2$), we obtain, on maximizing $\Omega^{(0)}$,

$$\left(\frac{\partial \Omega_1(E_1)}{\partial E_1}\right)_{E_1=\overline{E}_1} \Omega_2(\overline{E}_2) + \Omega_1(\overline{E}_1) \left(\frac{\partial \Omega_2(E_2)}{\partial E_2}\right)_{E_2=\overline{E}_2} \cdot \frac{\partial E_2}{\partial E_1} = 0.$$

Since $\partial E_2/\partial E_1 = -1$, see eqn. (1), the foregoing condition can be written as

$$\left(\frac{\partial \ln \Omega_1(E_1)}{\partial E_1}\right)_{E_1=\overline{E}_1} = \left(\frac{\partial \ln \Omega_2(E_2)}{\partial E_2}\right)_{E_2=\overline{E}_2}.$$

Thus, our condition for equilibrium reduces to the equality of the parameters β_1 and β_2 of the sub-systems A_1 and A_2, respectively, where β is defined by

$$\beta \equiv \left(\frac{\partial \ln \Omega(N, V, E)}{\partial E}\right)_{N,V,E=\overline{E}}. \tag{3}$$

We thus find that when two physical systems are brought into thermal contact, which allows an exchange of energy between them, this exchange continues *until* the equilibrium values $\overline{E}_1$ and $\overline{E}_2$ of the variables E_1 and E_2 are reached. Once these values are reached, there is no more net exchange of energy between the two systems; the systems are then said to have attained a state of mutual equilibrium. According to our analysis, this happens only when the respective values of the parameter β, namely β_1 and β_2, become equal.[4] It is then natural to expect that the parameter β is somehow related to the *thermodynamic temperature* T of a given system. To determine this relationship, we recall the thermodynamic formula

$$\left(\frac{\partial S}{\partial E}\right)_{N,V} = \frac{1}{T}, \tag{4}$$

where S is the entropy of the system in question. Comparing eqns (3) and (4), we conclude that an equally intimate relationship exists between the thermodynamic quantity S and the statistical quantity Ω; we may, in fact, write for any physical system

$$\frac{\Delta S}{\Delta(\ln \Omega)} = \frac{1}{\beta T} = \text{const.} \tag{5}$$

This correspondence was first established by Boltzmann who also believed that, since the relationship between the thermodynamic approach and the statistical approach seems to be of a fundamental character, the constant appearing in (5) must be a *universal constant.* It was Planck who first wrote the explicit formula

$$S = k \ln \Omega, \tag{6}$$

without any additive constant S_0. As it stands, formula (6) determines the *absolute* value of the entropy of a given physical system in terms of the total number of microstates accessible to it in conformity with the given macrostate. The zero of entropy then corresponds to the special state for which only one microstate is accessible ($\Omega = 1$)–the so-called "unique configuration"; the statistical approach thus provides a theoretical basis for the third law of thermodynamics as well. Formula (6) is of fundamental importance in physics; it provides a bridge between the microscopic and the macroscopic.

Now, in the study of the second law of thermodynamics we are told that the law of increase of entropy is related to the fact that the energy content of the universe, in its natural course, is becoming less and less available for conversion into work; accordingly, the entropy of a given system may be regarded as a measure of the so-called disorder or chaos prevailing in the system. Formula (6) tells us how disorder arises microscopically. Clearly, disorder is a manifestation of the largeness of the number of microstates the system can have. The larger the choice of microstates, the lesser the degree of predictability or the level of order in the system. Complete order prevails when and only when the system has no other choice but to be in a unique state ($\Omega = 1$); this, in turn, corresponds to a state of vanishing entropy.

By (5) and (6), we also have

$$\beta = \frac{1}{kT}. \tag{7}$$

The universal constant k is generally referred to as the *Boltzmann constant.* In Sec. 1.4 we shall discover how k is related to the gas constant R and the Avogadro number N_A; see eqn. (1.4.3).[5]

1.3. Further contact between statistics and thermodynamics

In continuation of the preceding considerations, we now examine a more elaborate exchange between the sub-systems A_1 and A_2. If we assume that the wall separating the two sub-systems is movable as well as conducting, then the respective volumes V_1 and V_2 (of sub-systems A_1 and A_2) also become variable; indeed, the total volume $V^{(0)} (= V_1 + V_2)$ remains constant, so that effectively we have only one more independent variable. Of course, the wall is still assumed to be impenetrable to particles, so the numbers N_1 and N_2 remain fixed. Arguing as before, the state of equilibrium for the composite system $A^{(0)}$ will obtain when the number $\Omega^{(0)}(V^{(0)}, E^{(0)}; V_1, E_1)$ attains its largest value, i.e. when not only

$$\left(\frac{\partial \ln \Omega_1}{\partial E_1}\right)_{N_1, V_1;\ E_1=\bar{E}_1} = \left(\frac{\partial \ln \Omega_2}{\partial E_2}\right)_{N_2, V_2;\ E_2=\bar{E}_2}, \tag{1a}$$

but also

$$\left(\frac{\partial \ln \Omega_1}{\partial V_1}\right)_{N_1,E_1;\ V_1=\bar{V}_1} = \left(\frac{\partial \ln \Omega_2}{\partial V_2}\right)_{N_2,E_2;\ V_2=\bar{V}_2}. \tag{1b}$$

Our conditions for equilibrium now take the form of an equality between the pair of parameters (β_1, η_1) of the sub-system A_1 and the parameters (β_2, η_2) of the sub-system A_2 where, by definition,

$$\eta \equiv \left(\frac{\partial \ln \Omega(N, V, E)}{\partial V}\right)_{N,E,V=\bar{V}}. \tag{2}$$

Similarly, if A_1 and A_2 came into contact through a wall which allowed an exchange of particles as well, the conditions for equilibrium would be further augmented by the equality of the parameter ζ_1 of sub-system A_1 and the parameter ζ_2 of sub-system A_2 where, by definition,

$$\zeta \equiv \left(\frac{\partial \ln \Omega(N, V, E)}{\partial N}\right)_{V,E,N=\bar{N}}. \tag{3}$$

To determine the physical meaning of the parameters η and ζ, we make use of eqn. (1.2.6) and the basic formula of thermodynamics, viz.

$$dE = TdS - PdV + \mu\, dN, \tag{4}$$

where P is the thermodynamic pressure and μ the chemical potential of the given system. It follows that

$$\eta = \frac{P}{kT} \quad \text{and} \quad \zeta = -\frac{\mu}{kT}. \tag{5}$$

From a physical point of view, these results are completely satisfactory because, thermodynamically too, the conditions of equilibrium between two systems A_1 and A_2, if the wall separating them is both conducting and movable (thus making their respective energies and volumes variable), are indeed the same as the ones contained in eqns (1), namely

$$T_1 = T_2 \quad \text{and} \quad P_1 = P_2. \tag{6}$$

On the other hand, if the two systems can exchange particles as well as energy but have their volumes fixed, the conditions of equilibrium, obtained thermodynamically, are indeed

$$T_1 = T_2 \quad \text{and} \quad \mu_1 = \mu_2. \tag{7}$$

And finally, if the exchange is such that all three (macroscopic) parameters become variable, then the conditions of equilibrium become

$$T_1 = T_2, \quad P_1 = P_2 \quad \text{and} \quad \mu_1 = \mu_2. \tag{8}$$[6]

It is gratifying that these conclusions are identical with the ones following from statistical considerations.

Combining the results of the foregoing discussi[illegible] we arrive at the following recipe for deriving thermodynamics from a statistic[illegible]ginning: determine, for the macrostate (N, V, E) of the given system, the number of all possible microstates accessible to the system; call this number $\Omega(N, V, E)$. Then, the entropy of the system in that state follows from the fundamental formula

$$S(N, V, E) = k \ln \Omega(N, V, E), \tag{9}$$

while the leading *intensive* parameters, viz. temperature, pressure and chemical potential, are given by

$$\left(\frac{\partial S}{\partial E}\right)_{N,V} = \frac{1}{T}; \quad \left(\frac{\partial S}{\partial V}\right)_{N,E} = \frac{P}{T}; \quad \left(\frac{\partial S}{\partial N}\right)_{V,E} = -\frac{\mu}{T}. \tag{10}$$

Alternatively, we can write[7]

$$P = \left(\frac{\partial S}{\partial V}\right)_{N,E} \Big/ \left(\frac{\partial S}{\partial E}\right)_{N,V} = -\left(\frac{\partial E}{\partial V}\right)_{N,S} \tag{11}$$

and

$$\mu = -\left(\frac{\partial S}{\partial N}\right)_{V,E} \Big/ \left(\frac{\partial S}{\partial E}\right)_{N,V} = \left(\frac{\partial E}{\partial N}\right)_{V,S}, \tag{12}$$

while

$$T = \left(\frac{\partial E}{\partial S}\right)_{N,V}. \tag{13}$$

Formulae (11)–(13) follow equally well from eqn. (4). The evaluation of P, μ and T from these formulae indeed requires that the energy E be expressed as a function of the quantities N, V and S; this should, in principle, be possible once S is known as a function of N, V and E; see (9).

The rest of the thermodynamics follows straightforwardly. For instance, the Helmholtz free energy A, the Gibbs free energy G and the enthalpy H are given by

$$A = E - TS, \tag{14}$$

$$\begin{aligned} G &= A + PV = E - TS + PV \\ &= \mu N \end{aligned} \tag*{(15)[8]}$$

and

$$H = E + PV = G + TS. \tag{16}$$

The specific heat at constant volume, C_V, and the one at constant pressure, C_P, would be given by

$$C_V \equiv T\left(\frac{\partial S}{\partial T}\right)_{N,V} = \left(\frac{\partial E}{\partial T}\right)_{N,V} \tag{17}$$

and

$$C_P \equiv T\left(\frac{\partial S}{\partial T}\right)_{N,P} = \left(\frac{\partial (E + PV)}{\partial T}\right)_{N,P} = \left(\frac{\partial H}{\partial T}\right)_{N,P}. \tag{18}$$

1.4. The classical ideal gas

To illustrate the approach developed in the preceding sections, we shall now derive the various thermodynamic properties of a classical ideal gas composed of monatomic molecules. The main reason why we choose this highly specialized system for consideration is that it affords an explicit, though asymptotic, evaluation of the number $\Omega(N, V, E)$. This example becomes all the more instructive when we find that its study enables us, in a most straightforward manner, to identify the Boltzmann constant k in terms of other physical constants; see eqn. (3). Moreover, the behavior of this system serves as a useful reference with which the behavior of other physical systems, especially real gases (with or without quantum effects), can be compared. And, indeed, in the limit of high temperatures and low densities the ideal-gas behavior becomes typical of most real systems.

Before undertaking a detailed study of this case it appears worth while to make a remark which applies to all *classical* systems composed of *non-interacting* particles, irrespective of the internal structure of the particles. This remark is related to the explicit dependence of the number $\Omega(N, V, E)$ on V and hence to the *equation of state* of these systems. Now, if there do not exist any spatial correlations among the particles, i.e. if the probability of any one of them being found in a particular region of the available space is completely independent of the location of the other particles,[9] then the total number of ways in which the N particles can be spatially distributed in the system will be simply equal to the product of the numbers of ways in which the individual particles can be accommodated in the same space independently of one another. With N and E fixed, each of these numbers will be directly proportional to V, the volume of the container; accordingly, the total number of ways will be directly proportional to the Nth power of V:

$$\Omega(N, E, V) \propto V^N. \tag{1}$$

Combined with eqns (1.3.9) and (1.3.10), this gives

$$\frac{P}{T} = k\left(\frac{\partial \ln \Omega(N, E, V)}{\partial V}\right)_{N,E} = k\frac{N}{V}. \tag{2}$$

If the system contains n moles of the gas, then $N = nN_A$, where N_A is the *Avogadro number*. Equation (2) then becomes

$$PV = nRT \quad (R = kN_A), \tag{3}$$

which is the famous *ideal-gas law*, R being the *gas constant* per mole. Thus, for any classical system composed of non-interacting particles the ideal-gas law holds.

For deriving other thermodynamic properties of this system, we require a detailed knowledge of the way Ω depends on the parameters N, V and E. The problem essentially reduces to determining the total number of ways in which eqns (1.1.1) and (1.1.2) can be mutually satisfied. In other words, we have to determine the total number of (independent) ways of satisfying the equation

$$\sum_{r=1}^{3N} \varepsilon_r = E, \tag{4}$$

where the ε_r are the energies associated with the various degrees of freedom of the N particles. The reason why this number should depend upon the parameters N and E is quite obvious. Nevertheless, this number would also depend upon the "spectrum of values" which the variables ε_r can assume; it is through this spectrum that the dependence on V comes in. Now, the energy eigenvalues for a *free, nonrelativisitic* particle confined to a cubical box of side $L(V = L^3)$, under the condition that the wave function $\psi(\boldsymbol{r})$ vanishes everywhere on the boundary, are given by

$$\varepsilon(n_x, n_y, n_z) = \frac{h^2}{8mL^2}(n_x^2 + n_y^2 + n_z^2); \quad n_x, n_y, n_z = 1, 2, 3, .. \tag{5}$$

where h is Planck's constant and m the mass of the particle. The number of distinct eigenfunctions (or microstates) for a particle of energy ε would, therefore, be equal to the number of independent, positive-integral solutions of the equation

$$(n_x^2 + n_y^2 + n_z^2) = \frac{8mV^{2/3}\varepsilon}{h^2} = \varepsilon^*, \quad \text{say.} \tag{6}$$

We may denote this number by $\Omega(1, \varepsilon, V)$. Extending the argument it follows that the desired number $\Omega(N, E, V)$ would be equal to the number of independent, positive-integral solutions of the equation

$$\sum_{r=1}^{3N} n_r^2 = \frac{8mV^{2/3}E}{h^2} = E^*, \quad \text{say.} \tag{7}$$

An important result follows straightforwardly from eqn. (7), even before the number $\Omega(N, E, V)$ is explicitly evaluated. From the nature of the expression appearing on the right-hand side of this equation we conclude that the volume V and the energy E of the system enter into the expression for Ω in the form of the combination $(V^{2/3}E)$. Consequently,

$$S(N, V, E) \equiv S(N, V^{2/3}E). \tag{8}$$

Hence, for the constancy of S and N, which defines a *reversible adiabatic* process,

$$V^{2/3}E = \text{const.} \tag{9}$$

Equation (1.3.11) then gives

$$P = -\left(\frac{\partial E}{\partial V}\right)_{N,S} = \frac{2}{3}\frac{E}{V}, \tag{10}$$

that is, the pressure of a system of nonrelativistic, non-interacting particles is precisely equal to two-thirds of its energy density.[10] It should be noted here that, since an explicit computation of the number Ω has not yet been done, results (9) and (10) hold for *quantum* as well as *classical* statistics; equally general is the result obtained by combining these, viz.

$$PV^{5/3} = \text{const.}, \tag{11}$$

which tells us how P varies with V during a *reversible adiabatic* process.

We shall now attempt to evaluate the number Ω. In this evaluation we shall explicitly assume the particles to be *distinguishable*, so that if a particle in state i gets interchanged with a particle in state j the resulting microstate is counted as distinct. Consequently, the number $\Omega(N, V, E)$, or better $\Omega_N(E^*)$ {see eqn. (7)}, is equal to the number of positive-integral lattice points lying on the surface of a $3N$-dimensional sphere of radius $\sqrt{E^*}$.[11] Clearly, this number will be an extremely irregular function of E^*, in that for two given values of E^* which may be very close to one another the values of this number could be very different. In contrast, the number $\Sigma_N(E^*)$, which denotes the number of positive-integral lattice points lying *on or within* the surface of a $3N$-dimensional sphere of radius $\sqrt{E^*}$, will be much less irregular. In terms of our physical problem, this would correspond to the number, $\Sigma(N, V, E)$, of microstates of the given system consistent with *all* macrostates characterized by the specified values of the parameters N and V but having energy *less than or equal to* E, i.e.

$$\Sigma(N, V, E) = \sum_{E' \leq E} \Omega(N, V, E') \tag{12}$$

or

$$\Sigma_N(E^*) = \sum_{E^{*'} \leq E^*} \Omega_N(E^{*'}). \tag{13}$$

Of course, the number Σ will also be somewhat irregular; however, we expect that its asymptotic behavior, as $E^* \to \infty$, would be a lot smoother than that of Ω. We shall see in the sequel that the thermodynamics of the system follows equally well from the number Σ as from Ω.

To appreciate the point made here, let us digress a little to examine the behavior of the numbers $\Omega_1(\varepsilon^*)$ and $\Sigma_1(\varepsilon^*)$, which correspond to the case of a single particle confined to the given volume V. The *exact* values of these numbers, for $\varepsilon^* \leq 10000$, can be extracted from a table compiled by Gupta (1947). The wild irregularities of the number $\Omega_1(\varepsilon^*)$ can hardly be missed. The number $\Sigma_1(\varepsilon^*)$, on the other hand, exhibits a much smoother asymptotic behavior. From the geometry of the problem, we note that, *asymptotically*, $\Sigma_1(\varepsilon^*)$ should be equal to the volume of an octant of a three-dimensional sphere of radius $\sqrt{\varepsilon^*}$, i.e.

$$\lim_{\varepsilon^* \to \infty} \frac{\Sigma_1(\varepsilon^*)}{(\pi/6)\varepsilon^{*3/2}} = 1. \tag{14}$$

A more detailed analysis shows (see Pathria, 1966) that, to the next approximation,

$$\Sigma_1(\varepsilon^*) \approx \frac{\pi}{6}\varepsilon^{*3/2} - \frac{3\pi}{8}\varepsilon^*; \tag{15}$$

the correction term arises from the fact that the volume of an octant somewhat overestimates the number of desired lattice points, for it includes, partly though, some points with one or more coordinates equal to zero. Figure 1.2 shows a histogram of the actual values of $\Sigma_1(\varepsilon^*)$ for ε^* lying between 200 and 300; the theoretical estimate (15) is also shown. In the figure, we have also included a histogram of the actual values of the corresponding number of microstates, $\Sigma_1'(\varepsilon^*)$, when the quantum numbers n_x, n_y and n_z can assume the value zero as well. In the latter case, the volume of an octant somewhat underestimates the number of desired lattice points; we now have

$$\Sigma_1'(\varepsilon^*) \approx \frac{\pi}{6}\varepsilon^{*3/2} + \frac{3\pi}{8}\varepsilon^*. \tag{16}$$

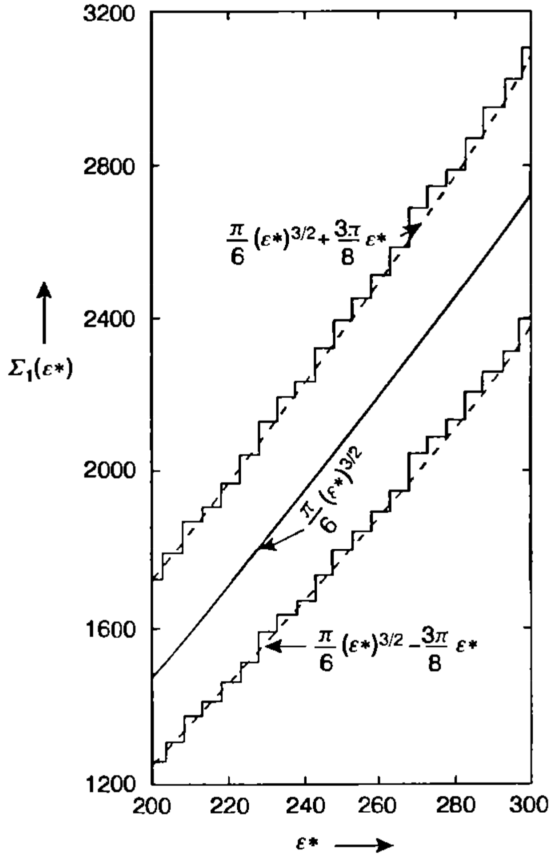


FIG. 1.2. Histograms showing the actual number of microstates available to a particle in a cubical enclosure; the lower histogram corresponds to the so-called Dirichlet boundary conditions, while the upper one corresponds to the Neumann boundary conditions (see Appendix A). The corresponding theoretical estimates, (15) and (16), are shown by dashed lines; the customary estimate, (14), is shown by a solid line.

Asymptotically, however, the number $\Sigma_1'(\varepsilon^*)$ also satisfies eqn. (14).

Returning to the N-particle problem, the number $\Sigma_N(E^*)$ should be *asymptotically* equal to the "volume" of the "positive compartment" of a $3N$-dimensional sphere of radius $\sqrt{E^*}$. Referring to eqn. (7) of Appendix C, we obtain

$$\Sigma_N(E^*) \approx \left(\frac{1}{2}\right)^{3N} \left\{ \frac{\pi^{3N/2}}{(3N/2)!} E^{*3N/2} \right\}$$

which, on substitution for E^*, gives

$$\Sigma(N, V, E) \approx \left(\frac{V}{h^3}\right)^N \frac{(2\pi m E)^{3N/2}}{(3N/2)!}. \tag{17}$$

Taking logarithms and applying Stirling's formula (B.29),

$$\ln(n!) \approx n \ln n - n \quad (n \gg 1), \tag{18}$$

we get

$$\ln \Sigma(N, V, E) \approx N \ln \left[\frac{V}{h^3} \left(\frac{4\pi m E}{3N} \right)^{3/2} \right] + \frac{3}{2} N. \tag{19}$$

For deriving the thermodynamic properties of the given system we must somehow fix the precise value of, or limits for, the energy of the system. In view of the extremely irregular nature of the function $\Omega(N, V, E)$, the specification of a precise value for the energy of the system cannot be justified on physical grounds, for that would never yield well-behaved expressions for the thermodynamic functions of the system. From a practical point of view, too, an absolutely isolated system is too much of an idealization. In the real world, almost every system has some contact with its surroundings, however little it may be; as a result, its energy cannot be defined sharply.[12] Of course, the effective width of the range over which the energy may vary would, in general, be small in comparison with the mean value of the energy. Let us specify this range by the limits $(E - \frac{1}{2}\Delta)$ and $(E + \frac{1}{2}\Delta)$ where, by assumption, $\Delta \ll E$; typically, $\Delta/E = O(1/\sqrt{N})$. The corresponding number of microstates, $\Gamma(N, V, E; \Delta)$, is then given by

$$\Gamma(N, V, E; \Delta) \simeq \frac{\partial \Sigma(N, V, E)}{\partial E}\Delta \approx \frac{3N}{2}\frac{\Delta}{E}\Sigma(N, V, E) \tag{17a}$$

whence

$$\ln \Gamma(N, V, E; \Delta) \approx N \ln\left[\frac{V}{h^3}\left(\frac{4\pi m E}{3N}\right)^{3/2}\right] + \frac{3}{2}N + \left\{\ln\left(\frac{3N}{2}\right) + \ln\left(\frac{\Delta}{E}\right)\right\}. \tag{19a}$$

Now, for $N \gg 1$, the first term in the curly bracket is negligible in comparison with any of the terms outside this bracket, for $\lim_{N\to\infty} (\ln N)/N = 0$. Furthermore, for any reasonable value of Δ/E, the same is true of the second term in this bracket.[13] Hence, for all practical purposes,

$$\ln \Gamma \approx \ln \Sigma \approx N \ln\left[\frac{V}{h^3}\left(\frac{4\pi m E}{3N}\right)^{3/2}\right] + \frac{3}{2}N. \tag{20}$$

Thus, we arrive at the baffling result that, for all practical purposes, the actual width of the range allowed for the energy of the system does not make much difference; the energy could lie between $(E - \frac{1}{2}\Delta)$ and $(E + \frac{1}{2}\Delta)$ or equally well between 0 and E. The reason underlying this situation is that the rate at which the number of microstates of the system increases with energy is so fantastic, see eqn. (17), that even if we allow *all* values of energy between zero and a particular value E it is only the "immediate neighborhood" of E that makes an overwhelmingly dominant contribution to this number! And since we are finally concerned only with the logarithm of this number, even the "width" of that neighborhood is inconsequential!

The stage is now set for deriving the thermodynamics of our system. First of all, we have

$$S(N, V, E) = k \ln \Gamma = Nk \ln\left[\frac{V}{h^3}\left(\frac{4\pi m E}{3N}\right)^{3/2}\right] + \frac{3}{2}Nk \tag{21^{14}}$$

whence

$$E(S, V, N) = \frac{3h^2 N}{4\pi m V^{2/3}} \exp\left(\frac{2S}{3Nk} - 1\right). \tag{22}$$

The tempera ‸ of the gas then follows with the help of formula (1.3.10) or (1.3.13), which leads to the energy–temperature relationship

$$E = N\left(\tfrac{3}{2}kT\right) = n\left(\tfrac{3}{2}RT\right), \tag{23}$$

where n is the number of moles of the gas. The specific heat at constant volume now follows with the help of formula (1.3.17):

$$C_V = \left(\frac{\partial E}{\partial T}\right)_{N,V} = \frac{3}{2}Nk = \frac{3}{2}nR. \tag{24}$$

For the equation of state, we obtain

$$P = -\left(\frac{\partial E}{\partial V}\right)_{N,S} = \frac{2}{3}\frac{E}{V}, \tag{25}$$

which agrees with our earlier result (10). Combined with (23), the foregoing result becomes

$$P = \frac{NkT}{V} \quad \text{or} \quad PV = nRT, \tag{26}$$

which is the same as (3). The specific heat at constant pressure is given by, see (1.3.18),

$$C_P = \left(\frac{\partial(E+PV)}{\partial T}\right)_{N,P} = \frac{5}{2}nR, \tag{27}$$

so that, for the ratio of the two specific heats, we have

$$\gamma = C_P/C_V = \tfrac{5}{3}. \tag{28}$$

Now, suppose that the gas undergoes an *isothermal* change of state (T = const. and N = const.); then, according to (23), the total energy of the gas would remain constant while, according to (26), its pressure would vary inversely with volume (Boyle's law). The change in the entropy of the gas, between the initial state i and the final state f, would be, see eqn. (21),

$$S_f - S_i = Nk\ln(V_f/V_i). \tag{29}$$

On the other hand, if the gas undergoes a *reversible adiabatic* change of state (S = const. and N = const.), then, according to (22) and (23), both E and T would vary as $V^{-2/3}$; moreover, according to (25) or (26), P would vary as $V^{-5/3}$. These results agree with the conventional thermodynamic ones, namely

$$PV^{\gamma} = \text{const.} \quad \text{and} \quad TV^{\gamma-1} = \text{const.}, \tag{30}$$

with $\gamma = \frac{5}{3}$. It may be noted that, thermodynamically, the change in E during an adiabatic process arises solely from the external work done by the gas on the surroundings or vice versa:

$$(dE)_{\text{adiab}} = -PdV = -\frac{2E}{3V}dV; \tag{31}$$

see eqns (1.3.4) and (25). The dependence of E on V follows readily from this relationship.

The considerations of this section have clearly demonstrated the manner in which the thermodynamics of a macroscopic system can be derived from the multiplicity of its microstates (as represented by the number Ω or Γ or Σ). The whole problem then hinges on an asymptotic enumeration of these numbers which unfortunately is tractable only in a few idealized cases, such as the one considered in this section; see also Problems 1.7 and 1.8. Even in an idealized case like this, there remains an inadequacy which could not be detected in the derivations made so far; this relates to the *explicit* dependence of S on N. The discussion of the next section is intended not only to bring out this inadequacy but also to provide the necessary remedy for it.

1.5. The entropy of mixing and the Gibbs paradox

One thing we readily observe from expression (1.4.21) is that, contrary to what is logically desired, the entropy of an ideal gas, as given by this expression, is *not* an extensive property of the system! That is, if we increase the size of the system by a factor α, keeping the intensive variables unchanged,[15] then the entropy of the system, which should also increase by the same factor α, does not do so; the presence of the $\ln V$ term in the expression affects the result adversely. This in a way means that the entropy of the system is different from the sum of the entropies of its parts, which is absurd. A more common way of looking at this problem is to consider the so-called *Gibbs paradox.*

Gibbs visualized the mixing of two ideal gases 1 and 2, both being initially at the same temperature T; see Fig. 1.3. Clearly, the temperature of the mixture would also be the same. Now, before the mixing took place, the respective entropies of the two gases were, see eqns (1.4.21) and (1.4.23),

$$S_i = N_i k \ln V_i + \frac{3}{2} N_i k \left\{1 + \ln\left(\frac{2\pi m_i kT}{h^2}\right)\right\}; \quad i = 1, 2. \tag{1}$$

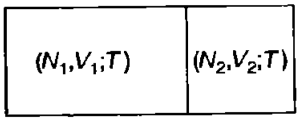


FIG. 1.3. The mixing together of two ideal gases 1 and 2.

After the mixing has taken place, the total entropy would be

$$S_T = \sum_{i=1}^{2} \left[N_i k \ln V + \frac{3}{2} N_i k \left\{1 + \ln\left(\frac{2\pi m_i kT}{h^2}\right)\right\}\right], \tag{2}$$

where $V = V_1 + V_2$. Thus, the net increase in the value of S, which may be called the *entropy of mixing*, is given by

$$(\Delta S) = S_T - \sum_{i=1}^{2} S_i = k\left[N_1 \ln \frac{V_1 + V_2}{V_1} + N_2 \ln \frac{V_1 + V_2}{V_2}\right]; \tag{3}$$

the quantity ΔS is indeed positive, as it must be for an *irreversible* process like mixing. Now, in the special case when the initial particle densities of the two gases (and, hence, the particle density of the mixture) are also the same, eqn. (3) becomes

$$(\Delta S)^* = k\left[N_1 \ln \frac{N_1+N_2}{N_1} + N_2 \ln \frac{N_1+N_2}{N_2}\right], \tag{4}$$

which is again positive.

So far, it seems all right. However, a paradoxical situation arises if we consider the mixing of two samples of the same gas. Once again, the entropies of the individual samples will be given by (1); of course, now $m_1 = m_2 = m$, say. And the entropy after mixing will be given by

$$S_T = Nk \ln V + \frac{3}{2}Nk\left\{1 + \ln\left(\frac{2\pi mkT}{h^2}\right)\right\}, \tag{2a}$$

where $N = N_1 + N_2$; note that this expression is numerically the same as (2), with $m_i = m$. Therefore, the entropy of mixing in this case will also be given by expression (3) and, if $N_1/V_1 = N_2/V_2 = (N_1+N_2)/(V_1+V_2)$, by expression (4). The last conclusion, however, is unacceptable because the mixing of two samples of the same gas, with a common initial temperature T and a common initial particle density n, is clearly a *reversible* process, for we can simply reinsert the partitioning wall into the system and obtain a situation which is in no way different from the one we had before mixing. Of course, we tacitly imply that in dealing with (a system of) identical particles we cannot track them down individually; all we can reckon with is their numbers. When two dissimilar gases, even with a common initial temperature and a common initial particle density, mixed togéther the process was irreversible, for by reinserting the partitioning wall one would obtain two samples of the mixture and not the two gases that were originally present; to that case, expression (4) would indeed apply. However, in the present case, the corresponding result should be

$$(\Delta S)^*_{1\equiv 2} = 0. \tag{4a}$$ [16]

The foregoing result would also be consistent with the requirement that the entropy of a given system be equal to the sum of the entropies of its parts. Of course, we had already noticed that this is not ensured by expression (1.4.21). Thus, once again we are led to believe that there is something basically wrong with that expression.

To see how the above paradoxical situation can be avoided, we recall that, for the entropy of mixing of two samples of the same gas, with a common T and a common n, we were led to result (4), which can also be written as

$$(\Delta S)^* = S_T - (S_1 + S_2) \approx k[\ln\{(N_1+N_2)!\} - \ln(N_1!) - \ln(N_2!)], \tag{4}$$

instead of the logical result (4a). A closer look at this expression shows that we would indeed obtain the correct result if our original expression for S were diminished by an [illegible]*oc* term, $k\ln(N!)$, for that would diminish S_1 by $k\ln(N_1!)$, S_2 by $k\ln(N_2!)$ and S_T by $k\ln\{(N_1+N_2)!\}$, with the result that $(\Delta S)^*$ would turn out to be zero instead of the expression appearing in (4). Clearly, this would amount to an *ad hoc* reduction of the statistical number and Σ by a factor $N!$. This is precisely the remedy proposed by Gibbs to avoid the paradox in question.

If we agree to the foregoing suggestion, the modified expression for the entropy of a classical ideal gas would be

$$S(N, V, E) = Nk \ln \left[\frac{V}{Nh^3} \left(\frac{4\pi mE}{3N} \right)^{3/2} \right] + \frac{5}{2} Nk \tag{1.4.21a}$$

$$= Nk \ln \left(\frac{V}{N} \right) + \frac{3}{2} Nk \left\{ \frac{5}{3} + \ln \left(\frac{2\pi mkT}{h^2} \right) \right\}, \tag{1a}$$

which indeed is truly extensive! If we now mix two samples of the same gas at a common initial temperature T, the entropy of mixing turns out to be

$$(\Delta S)_{1\equiv 2} = k \left[(N_1 + N_2) \ln \left(\frac{V_1 + V_2}{N_1 + N_2} \right) - N_1 \ln \left(\frac{V_1}{N_1} \right) - N_2 \ln \left(\frac{V_2}{N_2} \right) \right] \tag{3a}$$

and, if the initial particle densities of the samples were also equal, the result would be

$$(\Delta S)^*_{1\equiv 2} = 0. \tag{4a}$$

It may be noted that for the mixing of two dissimilar gases, the original expressions (3) and (4) would continue to hold even when (1.4.21) is replaced by (1.4.21a).[17] The paradox of Gibbs is thereby resolved.

Equation (1a) is generally referred to as the *Sackur–Tetrode* equation. We reiterate the fact that, by this equation, the entropy of the system does indeed become a truly extensive quantity. Thus, the very root of the trouble has been eliminated by the recipe of Gibbs. We shall discuss the physical implications of this recipe in Sec. 1.6; here, let us jot down some of its immediate consequences.

First of all, we note that the expression for the energy E of the gas, written as a function of S, V and N, is also modified. We now have

$$E(S, V, N) = \frac{3h^2 N^{5/3}}{4\pi m V^{2/3}} \exp \left(\frac{2S}{3Nk} - \frac{5}{3} \right), \tag{1.4.22a}$$

which, unlike its predecessor (1.4.22), makes energy too a truly extensive quantity. Of course, the thermodynamic results (1.4.23) through (1.4.31), derived in the previous section, remain unchanged. However, there are some which were intentionally left out, for they would come out correct only from the modified expression for $S(N, V, E)$ or $E(S, V, N)$. The most important of these is the chemical potential of the gas, for which we obtain

$$\mu \equiv \left(\frac{\partial E}{\partial N} \right)_{V,S} = E \left[\frac{5}{3N} - \frac{2S}{3N^2 k} \right]. \tag{5}$$

In view of eqns (1.4.23) and (1.4.25), this becomes

$$\mu = \frac{1}{N} [E + PV - TS] \equiv \frac{G}{N}, \tag{6}$$

where G is the Gibbs free energy of the system. In terms of the variables N, V and T, expression (5) takes the form

$$\mu(N, V, T) = kT \ln\left\{\frac{N}{V}\left(\frac{h^2}{2\pi mkT}\right)^{3/2}\right\}. \tag{7}$$

Another quantity of importance is the Helmholtz free energy:

$$A = E - TS = G - PV = NkT\left[\ln\left\{\frac{N}{V}\left(\frac{h^2}{2\pi mkT}\right)^{3/2}\right\} - 1\right]. \tag{8}$$

It will be noted that while A is an extensive property of the system, μ is intensive.

1.6. The "correct" enumeration of the microstates

In the preceding section we saw that an *ad hoc* diminution in the entropy of an N-particle system by an amount $k\ln(N!)$, which implies an *ad hoc* reduction in the number of microstates accessible to the system by a factor $(N!)$, was able to correct the unphysical features of some of our former expressions. It is now natural to ask: why, *in principle*, should the number of microstates, computed in Sec. 1.4, be reduced in this manner? The physical reason for doing so is that the particles constituting the given system are not only identical but also *indistinguishable*; accordingly, it is unphysical to label them as No. 1, No. 2, No. 3, etc., and to speak of their being *individually* in the various single-particle states ε_i. All we can sensibly speak of is their distribution over the states ε_i by *numbers*, e.g. n_1 particles being in the state ε_1, n_2 in the state ε_2, and so on. Thus, the correct way of specifying a microstate of the system is through the distribution numbers $\{n_i\}$, and not through the statement as to "which particle is in which state". To elaborate the point, we may say that if we consider two microstates which differ from one another merely in an interchange of two particles in different energy states, then according to our original mode of counting we would regard these microstates as distinct; in view of the indistinguishability of the particles, however, these microstates are not distinct (for, physically, there exists no way of distinguishing between them).[18]

Now, the total number of permutations that can be effected among N particles, distributed according to the set $\{n_i\}$, is

$$\frac{N!}{n_1! n_2! \ldots}, \tag{1}$$

where the n_i must be consistent with the basic constraints (1.1.1) and (1.1.2).[19] If our particles were distinguishable, then all these permutations would lead to "distinct" microstates. However, in view of the indistinguishability of the particles, these permutations must be regarded as leading to one and the same thing; consequently, for *any* distribution set $\{n_i\}$, we have one, and only one, distinct microstate. As a result, the total number of distinct microstates accessible to the system, consistent with a given macrostate (N, V, E), would be severely cut down. However, since factor (1) itself depends upon the numbers n_i constituting a particular distribution set and for a given macrostate there will be many such sets, there is no straightforward way to "correct down" the number of microstates computed on the basis of the classical concept of "distinguishability" of the particles.

The recipe of Gibbs clearly amounts to disregarding the details of the numbers n_i and slashing the whole sequence of microstates by a *common* factor $N!$; this is correct for situations in which all N particles happen to be in different energy states but is certainly wrong for other situations. We must keep in mind that by adopting this recipe we are still using a spurious *weight factor*,

$$w\{n_i\} = \frac{1}{n_1!n_2!\ldots}, \tag{2}$$

for the distribution set $\{n_i\}$ whereas in principle we should use a factor of *unity*, irrespective of the values of the numbers n_i.[20] Nonetheless, the recipe of Gibbs does correct the situation in a gross manner, though in matters of detail it is still inadequate. In fact, it is only by taking $w\{n_i\}$ to be equal to unity (or zero) that we obtain true *quantum statistics*!

We thus see that the recipe of Gibbs corrects the enumeration of the microstates, as necessitated by the indistinguishability of the particles, only in a gross manner. Numerically, this would approach closer and closer to reality as the probability of the n_i being greater than 1 becomes less and less. This in turn happens when the given system is at a sufficiently high temperature (so that many more energy states become accessible) and has a sufficiently low density (so that there are not as many particles to accommodate). It follows that the "corrected" classical statistics represents truth more closely if the expectation values of the *occupation numbers* n_i are much less than unity:

$$\langle n_i \rangle \ll 1, \tag{3}$$

i.e. if the numbers n_i are generally 0, occasionally 1, and rarely greater than 1. Condition (3) in a way defines the *classical limit*. We must, however, remember that it is because of the application of the correction factor $1/N!$, which replaces (1) by (2), that our results agree with reality *at least* in the classical limit.

In Sec. 5.5 we shall demonstrate, in an independent manner, that the factor by which the number of microstates, as computed for the "labeled" molecules, be reduced so that the formalism of classical statistical mechanics becomes a true limit of the formalism of quantum statistical mechanics is indeed $N!$.

Problems

1.1. (a) Show that, for two *large* systems in thermal contact, the number $\Omega^{(0)}(E^{(0)}, E_1)$ of Sec. 1.2 can be expressed as a Gaussian in the variable E_1. Determine the root-mean-square deviation of E_1 from the mean value $\overline{E}_1$ in terms of other quantities pertaining to the problem.

(b) Make an explicit evaluation of the root-mean-square deviation of E_1 in the special case when the systems A_1 and A_2 are ideal classical gases.

1.2. Assuming that the entropy S and the statistical number Ω of a physical system are related through an arbitrary functional form

$$S = f(\Omega),$$

show that the additive character of S and the multiplicative character of Ω *necessarily* require that the function $f(\Omega)$ be of the form (1.2.6).

1.3. Two systems A and B, of identical composition, are bro[illegible] together and allowed to exchange both energy and particles, keeping volumes V_A and V_B cons[illegible]. Show that the minimum value of the quantity (dE_A/dN_A) is given by

$$\frac{\mu_A T_B - \mu_B T_A}{T_B - T_A},$$

where the μ and the T are the respective chemical potentials and temperatures.

1.4. In a classical gas of hard spheres (of diameter σ), the spatial distribution of the particles is no longer uncorrelated. Roughly speaking, the presence of n particles in the system leaves only a volume $(V - nv_0)$ available for the $(n+1)$th particle; clearly, v_0 would be proportional to σ^3. Assuming that $Nv_0 \ll V$, determine the dependence of $\Omega(N, V, E)$ on V {cf. eqn. (1.4.1)} and show that, as a result of this, V in the gas law (1.4.3) gets replaced by $(V - b)$, where b is four times the actual space occupied by the particles.

1.5. Read Appendix A and establish formulae (1.4.15) and (1.4.16). Estimate the importance of the linear term in these formulae, relative to the main term $(\pi/6)\varepsilon^{*3/2}$, for an oxygen molecule confined to a cube of side 10 cm; take $\varepsilon = 0.05$ eV.

1.6. A cylindrical vessel 1 m long and 0.1 m in diameter is filled with a monatomic gas at $P = 1$ atm and $T = 300$ K. The gas is heated by an electrical discharge, along the axis of the vessel, which releases an energy of 10^4 joules. What will the temperature of the gas be immediately after the discharge?

1.7. Study the statistical mechanics of an extreme relativisitic gas characterized by the single-particle energy states

$$\varepsilon(n_x, n_y, n_z) = \frac{hc}{2L}(n_x^2 + n_y^2 + n_z^2)^{1/2},$$

instead of (1.4.5), along the lines followed in Sec. 1.4. Show that the ratio C_P/C_V in this case is 4/3, instead of 5/3.

1.8. Consider a system of quasi-particles whose energy eigenvalues are given by

$$\varepsilon(n) = nh\nu; \quad n = 0, 1, 2, \ldots$$

Obtain an asymptotic expression for the number Ω of this system, for a given number N of the quasi-particles and a given total energy E. Determine the temperature T of the system as a function of E/N and $h\nu$, and examine the situation for which $E/(Nh\nu) \gg 1$.

1.9. Making use of the fact that the entropy $S(N, V, E)$ of a thermodynamic system is an extensive quantity, show that

$$N\left(\frac{\partial S}{\partial N}\right)_{V,E} + V\left(\frac{\partial S}{\partial V}\right)_{N,E} + E\left(\frac{\partial S}{\partial E}\right)_{N,V} = S.$$

Note that this result implies: $(-N\mu + PV + E)/T = S$, i.e. $N\mu = E + PV - TS$.

1.10. A mole of argon and a mole of helium are contained in vessels of equal volume. If argon is at 300 K, what should the temperature of helium be so that the two have the same entropy?

1.11. Four moles of nitrogen and one mole of oxygen at $P = 1$ atm and $T = 300$ K are mixed together to form air at the same pressure and temperature. Calculate the entropy of mixing per mole of the air formed.

1.12. Show that the various expressions for the entropy of mixing, derived in Sec. 1.5, satisfy the following relations:

(a) For all N_1, V_1, N_2 and V_2,

$$(\Delta S)_{1=2} = \{(\Delta S) - (\Delta S)^*\} \geq 0,$$

the equality holding when and only when $N_1/V_1 = N_2/V_2$.

(b) For a given value of $(N_1 + N_2)$,

$$(\Delta S)^* \leq (N_1 + N_2)k \ln 2,$$

the equality holding when and only when $N_1 = N_2$.

1.13. If the two gases considered in the mixing process of Sec. 1.5 were initially at different temperatures, say T_1 and T_2, what would the entropy of mixing be in that case? Would the contribution arising from this cause depend on whether the two gases were different or identical?

1.14. Show that for an ideal gas composed of monatomic molecules the entropy change, between any two temperatures, when the pressure is kept constant is 5/3 times the corresponding entropy change when the volume is kept constant. Verify this result *numerically* by calculating the actual values of $(\Delta S)_P$ and $(\Delta S)_V$ per mole of an ideal gas whose temperature is raised from 300 K to 400 K.

1.15. We hav n that the P–V relationship during a reversible adiabatic process in an ideal gas is governed by xponent γ, such that

$$PV^{\gamma} = \text{const.}$$

Consider a mixture of two ideal gases, with mole fraction and f_2 and respective exponents γ_1 and γ_2. Show that the effective exponent γ for the mixture is given by

$$\frac{1}{\gamma - 1} = \frac{f_1}{\gamma_1 - 1} + \frac{f_2}{\gamma_2 - 1}.$$

1.16. Establish thermodynamically the formulae

$$V\left(\frac{\partial P}{\partial T}\right)_\mu = S \quad \text{and} \quad V\left(\frac{\partial P}{\partial \mu}\right)_T = N.$$

Express the pressure P of an ideal classical gas in terms of the variables μ and T, and verify the above formulae

Notes

[1] The third law, which is also known as *Nernst's heat theorem*, did not arrive until about 1906. For a general discussion of this law, see Simon (1930) and Wilks (1961); these references also provide an extensive bibliography on this subject.

[2] It may be noted that the manner in which the ε_i depend on V is itself determined by the nature of the system. For instance, it is not the same for relativistic systems as it is for nonrelativistic ones; compare, for instance, the cases dealt with in Sec. 1.4 and in Problem 1.7. We should also note that, *in principle*, the dependence of Ω on V arises from the fact that it is the *physical dimensions* of the container that appear in the boundary conditions imposed on the wave functions of the system.

[3] It is obvious that the macrostate of the composite system $A^{(0)}$ has to be defined by two energies, viz. E_1 and E_2 (or else $E^{(0)}$ and E_1).

[4] This result may be compared with the so-called "zeroth law of thermodynamics", which stipulates the existence of a *common* parameter T for two or more physical systems in mutual equilibrium.

[5] We follow the notation whereby eqn. (1.4.3) means eqn. (3) of Sec. 1.4. However, while referring to an equation in the same section, we will omit the mention of the section number.

[6] It may be noted that the same would be true for any two parts of a single thermodynamic system; consequently, in equilibrium, the parameters T, P and μ would be constant *throughout* the system.

[7] In writing these formulae, we have made use of the well-known relationship in partial differential calculus, namely that "if three variables x, y and z are mutually related, then

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1".$$

[8] The relation $E - TS + PV = \mu N$ follows directly from (4). For this, all we have to do is to regard the given system as having grown to its present size in a gradual manner, such that the intensive parameters, T, P and μ stayed constant throughout the process while the extensive parameters N, V and E (and hence S) grew *proportionately* with one another.

[9] This will be true if (i) the mutual interactions among particles are negligible, and (ii) the wave packets of individual particles do not significantly overlap (or, in other words, the quantum effects are also negligible).

[10] Combining (10) with (2), we obtain for the classical ideal gas: $E = \frac{3}{2}NkT$. Accordingly, eqn. (9) reduces to the well-known thermodynamic relationship: $V^{\gamma-1}T = \text{const.}$, which holds during a *reversible adiabatic* process, with $\gamma = \frac{5}{3}$.

[11] If the particles are regarded as *indistinguishable*, the evaluation of the number Ω by counting lattice points becomes quite intricate. The problem is then solved by having recourse to the theory of "partitions of numbers"; see Auluck and Kothari (1946, 1947).

[12] Actually, the very act of making measurements on a system brings about, inevitably, a contact between the system and the surroundings.

[13] It should be clear that while Δ/E is much less than 1, it must not tend to 0, for that would make $\Gamma \to 0$ and $\ln\Gamma \to -\infty$. A situation of that kind would be too artificial and would have nothing to do with reality. Actually, in most physical systems, $\Delta/E = O(N^{-1/2})$ whereby $\ln(\Delta/E)$ becomes of order $\ln N$, which again is negligible in comparison with the terms outside the curly bracket.

[14] Henceforth, we shall replace the sign $\approx$, which characterizes the *asymptotic* character of a relationship, by the sign of equality because for most physical systems the asymptotic results are as good as exact.

[15] This mea[illegible]n increase of the parameters N, V and E to αN, αV and αE, so that the energy per particle and the volume per particle remain unchanged.

[16] In view of this, we fear that expression (3) may also be inapplicable to this case.

[17] Because, in this case, the entropy S_T of the mixture would be diminished by $k \ln (N_1!N_2!)$, rather than by $k \ln \{(N_1 + N_2)!\}$.

[18] Of course, if an interchange took place among particles in the same energy state, then even our original mode of counting did not regard the two microstates as distinct.

[19] The presence of the factors $(n_i!)$ in the denominator is related to the comment made in the preceding note.

[20] Or a factor of *zero* if the distribution set $\{n_i\}$ is disallowed on certain physical grounds, such as the Pauli exclusion principle.

CHAPTER 2

ELEMENTS OF ENSEMBLE THEORY

IN THE preceding chapter we noted that, for a given *macrostate* (N, V, E), a statistical system, at any time t, is equally likely to be in any one of an extremely large number of distinct *microstates*. As time passes, the system continually switches from one microstate to another, with the result that, over a reasonable span of time, all one observes is a behavior "averaged" over the variety of microstates through which the system passes. It may, therefore, make sense if we consider, at a *single* instant of time, a rather large number of systems–all being some sort of "mental copies" of the given system–which are characterized by the same macrostate as the original system but are, naturally enough, in all sorts of possible microstates. Then, under ordinary circumstances, we may expect that the average behavior of any system in this collection, which we call an *ensemble*, would be identical with the time-averaged behavior of the given system. It is on the basis of this expectation that we proceed to develop the so-called *ensemble theory*.

For classical systems, the most appropriate workshop for developing the desired formalism is the *phase space*. Accordingly, we begin our study of the various ensembles with an analysis of the basic features of this space.

2.1. Phase space of a classical system

The microstate of a given classical system, at any time t, may be defined by specifying the *instantaneous* positions and momenta of all the particles constituting the system. Thus, if N is the number of particles in the system, the definition of a microstate requires the specification of $3N$ position coordinates $q_1, q_2, \ldots, q_{3N}$ and $3N$ momentum coordinates $p_1, p_2, \ldots, p_{3N}$. Geometrically, the set of coordinates (q_i, p_i), where $i = 1, 2, \ldots, 3N$, may be regarded as a point in a space of $6N$ dimensions. We refer to this space as the *phase space*, and the phase point (q_i, p_i) as a *representative point*, of the given system.

Of course, the coordinates q_i and p_i are functions of the time t; the precise manner in which they vary with t is determined by the canonical equations of motion,

$$\left.\begin{aligned}\dot{q}_i &= \frac{\partial H(q_i, p_i)}{\partial p_i}\\ \dot{p}_i &= -\frac{\partial H(q_i, p_i)}{\partial q_i}\end{aligned}\right\} \quad i = 1, 2, \ldots, 3N, \tag{1}$$

where $H(q_i, p_i)$ is the *Hamiltonian* of the system. Now, as time passes, the set of coordinates (q_i, p_i), which also defines the microstate of the system, undergoes a continual change. Correspondingly, our representative point in the phase space carves out a *trajectory* whose direction, at any time t, is determined by the *velocity vector* $\boldsymbol{v} \equiv (\dot{q}_i, \dot{p}_i)$, which in turn is given by the equations of motion (1). It is not difficult to see that the trajectory of the representative point must remain within a limited region of the phase space; this is so because a finite volume V directly limits the values of the coordinates q_i, while a finite energy E limits the values of both the q_i and the p_i [through the Hamiltonian $H(q_i, p_i)$]. In particular, if the total energy of the system is known to have a *precise* value, say E, the corresponding trajectory will be restricted to the "hypersurface"

$$H(q_i, p_i) = E \tag{2}$$

of the phase space; on the other hand, if the total energy may lie anywhere in the range $(E - \frac{1}{2}\Delta, E + \frac{1}{2}\Delta)$, the corresponding trajectory will be restricted to the "hypershell" defined by these limits.

Now, if we consider an ensemble of systems (i.e. the given system, along with a large number of mental copies of it) then, at any time t, the various members of the ensemble will be in all sorts of possible microstates; indeed, each one of these microstates must be consistent with the given macrostate which is supposed to be common to all members of the ensemble. In the phase space, the corresponding picture will consist of a swarm of representative points, one for each member of the ensemble, all lying within the "allowed" region of this space. As time passes, every member of the ensemble undergoes a continual change of microstates; correspondingly, the representative points constituting the swarm continually move along their respective trajectories. The overall picture of this movement possesses some important features which are best illustrated in terms of what we call a *density function* $\rho(q, p; t)$. This function is such that, at any time t, the number of representative points in the "volume element" $(d^{3N}q\, d^{3N}p)$ around the point (q, p) of the phase space is given by the product $\rho(q, p; t)\, d^{3N}q\, d^{3N}p$.[1] Clearly, the density function $\rho(q, p; t)$ symbolizes the manner in which the members of the ensemble are distributed over all possible microstates at different instants of time. Accordingly, the *ensemble average* $\langle f \rangle$ of a given physical quantity $f(q, p)$, which may be different for systems in different microstates, would be given by

$$\langle f \rangle = \frac{\int f(q, p)\rho(q, p; t)\, d^{3N}q\, d^{3N}p}{\int \rho(q, p; t)\, d^{3N}q\, d^{3N}p}. \tag{3}$$

The integrations in (3) extend over the whole of the phase space; however, it is only the populated regions of the phase space $(\rho \neq 0)$ that really contribute. We note that, in general, the ensemble average $\langle f \rangle$ may itself be a function of time.

An ensemble is said to be *stationary* if ρ does not depend explicitly on time, i.e. at all times

$$\frac{\partial \rho}{\partial t} = 0. \tag{4}$$

Clearly, for such an ensemble the average value $\langle f \rangle$ of *any* physical quantity $f(q, p)$ will be independent of time. Naturally, a stationary ensemble qualifies to represent a system in *equilibrium*. To determine the circumstances under which eqn. (4) may hold, we have to make a rather detailed study of the movement of the representative points in the phase space.

2.2. Liouville's theorem and its consequences

Consider an arbitrary "volume" ω in the relevant region of the phase space and let the "surface" enclosing this volume be denoted by σ; see Fig. 2.1. Then, the rate at which the number of representative points in this volume increases with time is written as

$$\frac{\partial}{\partial t} \int_{\omega} \rho \, d\omega, \tag{1}$$

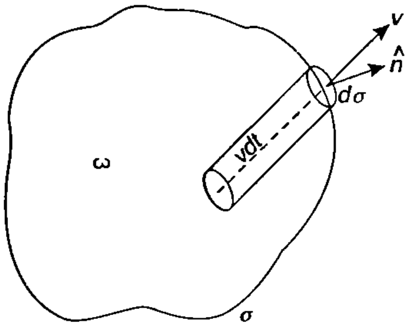


FIG. 2.1. The "hydrodynamics" of the representative points in the phase space.

where $d\omega \equiv (d^{3N}q \, d^{3N}p)$. On the other hand, the *net* rate at which the representative points "flow" out of ω (across the bounding surface σ) is given by

$$\int_{\sigma} \rho(\boldsymbol{v} \cdot \hat{\boldsymbol{n}}) \, d\sigma; \tag{2}$$

here, $\boldsymbol{v}$ is the velocity vector of the representative points in the region of the surface element $d\sigma$ while $\hat{\boldsymbol{n}}$ is the (outward) unit vector normal to this element. By the divergence theorem, (2) can be written as

$$\int_{\omega} \operatorname{div}(\rho \boldsymbol{v}) \, d\omega; \tag{3}$$

of course, the operation of divergence here means the following:

$$\operatorname{div}(\rho \boldsymbol{v}) \equiv \sum_{i=1}^{3N} \left\{ \frac{\partial}{\partial q_i}(\rho \dot{q}_i) + \frac{\partial}{\partial p_i}(\rho \dot{p}_i) \right\}. \tag{4}$$

In view of fact that there are no "sources" or "sinks" in the phase space and hence the total number of representative points remains conserved,[2] we have, by (1) and (3),

$$\frac{\partial}{\partial t}\int_{\omega} \rho \, d\omega = -\int_{\omega} \operatorname{div}(\rho \boldsymbol{v})\, d\omega, \tag{5}$$

that is,

$$\int_{\omega} \left\{ \frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \boldsymbol{v}) \right\} d\omega = 0. \tag{6}$$

Now, the necessary and sufficient condition that integral (6) vanish for all arbitrary volumes ω is that the integrand itself vanish *everywhere* in the relevant region of the phase space. Thus, we must have

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \boldsymbol{v}) = 0, \tag{7}$$

which is the *equation of continuity* for the swarm of the representative points.

Combining (4) and (7), we obtain

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left(\frac{\partial \rho}{\partial q_i}\dot{q}_i + \frac{\partial \rho}{\partial p_i}\dot{p}_i \right) + \rho \sum_{i=1}^{3N} \left(\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right) = 0. \tag{8}$$

The last group of terms vanishes identically because, by the equations of motion, we have, for all i,

$$\frac{\partial \dot{q}_i}{\partial q_i} = \frac{\partial^2 H(q_i, p_i)}{\partial q_i \partial p_i} \equiv \frac{\partial^2 H(q_i, p_i)}{\partial p_i \partial q_i} = -\frac{\partial \dot{p}_i}{\partial p_i}. \tag{9}$$

Further, since $\rho \equiv \rho(q_i, p_i; t)$, the remaining terms in (8) may be combined to form the "total" time derivative of ρ, with the result that

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + [\rho, H] = 0. \qquad (10)^3$$

Equation (10) embodies the so-called *Liouville's theorem* (1838). According to this theorem, the "local" density of the representative points, *as viewed by an observer moving with a representative point*, stays constant in time. Thus, the swarm of the representative points moves in the phase space in essentially the same manner as an incompressible fluid moves in the physical space!

A distinction must be made, however, between eqn. (10) on one hand and eqn. (2.1.4) on the other. While the former derives from the basic mechanics of the particles and is therefore *quite generally* true, the latter is only a requirement for equilibrium which, in a given case, may or may not be satisfied. The condition that ensures simultaneous validity of the two equations is clearly

$$[\rho, H] = \sum_{i=1}^{3N} \left(\frac{\partial \rho}{\partial q_i}\dot{q}_i + \frac{\partial \rho}{\partial p_i}\dot{p}_i \right) = 0. \tag{11}$$

Now, one possible way of satisfying (11) is to assume that ρ, which is already assumed to have no explicit dependence on time, is *independent* of the coordinates (q, p) as well, i.e.

$$\rho(q, p) = \text{const.} \tag{12}$$

over the relevant region of the phase space (and, of course, is zero everywhere else). Physically, this choice corresponds to an ensemble of systems which at *all* times are *uniformly* distributed over all possible microstates. The ensemble average (2.1.3) then reduces to

$$\langle f \rangle = \frac{1}{\omega} \int_{\omega} f(q, p)\, d\omega; \tag{13}$$

here, ω denotes the total "volume" of the relevant region of the phase space. Clearly, in this case, *any* member of the ensemble is equally likely to be in *any* one of the various possible microstates, inasmuch as *any* representative point in the swarm is equally likely to be in the neighborhood of *any* phase point in the allowed region of the phase space. This statement is usually referred to as the postulate of "equal *a priori* probabilities" for the various possible microstates (or for the various volume elements in the allowed region of the phase space); the resulting ensemble is referred to as the *microcanonical ensemble.*

A more general way of satisfying (11) is to assume that the dependence of ρ on (q, p) comes only through an explicit dependence on the Hamiltonian $H(q, p)$, i.e.

$$\rho(q, p) = \rho[H(q, p)]; \tag{14}$$

condition (11) is then identically satisfied. Equation (14) provides a class of density functions for which the corresponding ensemble is stationary. In Chap. 3 we shall see that the most natural choice in this class of ensembles is the one for which

$$\rho(q, p) \propto \exp[-H(q, p)/kT]. \tag{15}$$

The ensemble so defined is referred to as the *canonical ensemble.*

2.3. The microcanonical ensemble

In this ensemble the macrostate of a system is defined by the number of molecules N, the volume V and the energy E. However, in view of the considerations expressed in Sec. 1.4, we may prefer to specify a range of energy values, say from $(E - \frac{1}{2}\Delta)$ to $(E + \frac{1}{2}\Delta)$, rather than a sharply defined value E. With the macrostate specified, a choice still remains for the systems of the ensemble to be in *any one* of a large number of possible microstates. In the phase space, correspondingly, the representative points of the ensemble have a choice to lie *anywhere* within a "hypershell" defined by the condition

$$\left(E - \tfrac{1}{2}\Delta\right) \leq H(q, p) \leq \left(E + \tfrac{1}{2}\Delta\right) \tag{1}$$

The volume of the phase space enclosed within this shell is given by

$$\omega = \int' d\omega \equiv \int'(d^{3N}q\, d^{3N}p), \tag{2}$$

where the primed integration extends only over that part of the phase space which conforms to condition (1). It is clear that ω will [illegible] function of the parameters N, V, E and Δ.

Now, the microcanonical ensemble is a collection of systems for which the density function ρ is, at all times, given by

$$\left.\begin{array}{ll}\rho(q, p) = \text{const.} & \text{if } (E - \tfrac{1}{2}\Delta) \le H(q, p) \le (E + \tfrac{1}{2}\Delta) \\ \qquad 0 & \text{otherwise}\end{array}\right\}. \tag{3}$$

Accordingly, the expectation value of the number of representative points lying in a volume element $d\omega$ of the relevant hypershell is simply proportional to $d\omega$. In other words, the *a priori* probability of finding a representative point in a given volume element $d\omega$ is the same as that of finding a representative point in an equivalent volume element $d\omega$ located *anywhere* in the hypershell. In our original parlance, this means an equal *a priori* probability for a given member of the ensemble to be in *any one* of the various possible microstates. In view of these considerations, the ensemble average $\langle f \rangle$, as given by eqn. (2.2.13), acquires a simple physical meaning. To see this, we proceed as follows.

Since the ensemble under study is a stationary one, the ensemble average of any physical quantity f will be independent of time; accordingly, taking a time average thereof will not produce any new result. Thus

$$\begin{aligned}\langle f \rangle &\equiv \text{ the ensemble average of } f \\ &= \text{the time average of (the ensemble average of } f).\end{aligned}$$

Now, the processes of time averaging and ensemble averaging are completely independent, so the order in which they are performed may be reversed without causing any change in the value of $\langle f \rangle$. Thus

$$\langle f \rangle = \text{ the ensemble average of (the time average of } f).$$

Now, the time average of any physical quantity, taken over a sufficiently long interval of time, must be the same for *every* member of the ensemble, for after all we are dealing with only *mental copies* of a given system.[4] Therefore, taking an ensemble average thereof should be inconsequential, and we may write

$$\langle f \rangle = \text{ the long-time average of } f,$$

where the latter may be taken over *any* member of the ensemble. Furthermore, the long-time average of a physical quantity is all one obtains by making a measurement of that quantity on the given system; therefore, it may be identified with the value one expects to obtain through experiment. Thus, we finally have

$$\langle f \rangle = f_{\text{exp}}. \tag{4}$$

This brings us to the most important result: *the ensemble average of any physical quantity f is identical with the value one expects to obtain on making an appropriate measurement on the given system.*

The next thing we look for is the establishment of a connection between the mechanics of the microcanonical ensemble and the thermodynamics of the member systems. To do this, we observe that there exists a direct correspondence between the various microstates of the given system and the various locations in the phase space. The volume ω (of the allowed region of the phase space) is, therefore, a direct measure of the multiplicity Γ of the microstates accessible to the system.

To establish a numerical correspondence between ω and Γ, we need to discover a *fundamental volume* ω_0 which could be regarded as "equivalent to one microstate". Once this is done, we may say that, asymptotically,

$$\Gamma = \omega/\omega_0. \tag{5}$$

The thermodynamics of the system would then follow in the same way as in Secs 1.2–1.4, viz. through the relationship

$$S = k \ln \Gamma = k \ln (\omega/\omega_0), \quad \text{etc.} \tag{6}$$

The basic problem then consists in determining ω_0. From dimensional considerations, see (2), ω_0 must be in the nature of an "angular momentum raised to the power $3N$". To determine it exactly, we consider certain simplified systems, both from the point of view of the phase space and from the point of view of the distribution of quantum states.

2.4. Examples

We consider, first of all, the problem of a classical ideal gas composed of monatomic particles; see Sec. 1.4. In the microcanonical ensemble, the volume ω of the phase space accessible to the representative points of the (member) systems is given by

$$\omega = \int' \ldots \int' (d^{3N}q \, d^{3N}p), \tag{1}$$

where the integrations are restricted by the conditions that (i) the particles of the system are confined in physical space to volume V, and (ii) the total energy of the system lies between the limits $(E - \frac{1}{2}\Delta)$ and $(E + \frac{1}{2}\Delta)$. Since the Hamiltonian in this case is a function of the p_i alone, integrations over the q_i can be carried out straightforwardly; these give a factor of V^N. The remaining integral is

$$\underset{\left(E-\frac{1}{2}\Delta\right) \le \sum_{i=1}^{3N} (p_i^2/2m) \le \left(E+\frac{1}{2}\Delta\right)}{\int \ldots \int} d^{3N}p = \underset{2m\left(E-\frac{1}{2}\Delta\right) \le \sum_{i=1}^{3N} y_i^2 \le 2m\left(E+\frac{1}{2}\Delta\right)}{\int \ldots \int} d^{3N}y,$$

which is equal to the volume of a $3N$-dimensional hypershell, bounded by hyperspheres of radii

$$\sqrt{[2m(E + \tfrac{1}{2}\Delta)]} \quad \text{and} \quad \sqrt{[2m(E - \tfrac{1}{2}\Delta)]}.$$

For $\Delta \ll E$, this is given by the thickness of the shell, which is almost equal to $\Delta(m/2E)^{1/2}$, multiplied by the surface area of a $3N$-dimensional hypersphere of radius $\sqrt{(2mE)}$. By eqn. (7) of Appendix C, we obtain for this integral

$$\Delta \left(\frac{m}{2E}\right)^{1/2} \left\{ \frac{2\pi^{3N/2}}{[(3N/2)-1]!} (2mE)^{(3N-1)/2} \right\},$$

whence

$$\omega \simeq \frac{\Delta}{E} V^N \frac{(2\pi m E)^{3N/2}}{[(3N/2)-1]!}. \tag{2}$$

Comparing (2) with (1.4.17, 17a), we obtain the desired correspondence, viz.

$$(\omega/\Gamma)_{\text{asymp}} \equiv \omega_0 = h^{3N};$$

see also Problem 2.9. Quite generally, if the system under study has N' degrees of freedom, the desired conversion factor is

$$\omega_0 = h^{N'}. \tag{3}$$

In the case of a single particle, $N' = 3$; accordingly, the number of microstates available would asymptotically be equal to the volume of the allowed region of the phase space divided by h^3. Let $\Sigma(P)$ denote the number of microstates available to a free particle confined to volume V of the physical space, its momentum p being less than or equal to a specified value P. Then

$$\Sigma(P) \approx \frac{1}{h^3} \int \underset{p \le P}{\cdots} \int (d^3q\, d^3p) = \frac{V}{h^3}\frac{4\pi}{3}P^3, \tag{4}$$

whence we obtain for the number of microstates with momentum lying between p and $p + dp$

$$g(p)\,dp = \frac{d\Sigma(p)}{dp}\,dp \approx \frac{V}{h^3}4\pi p^2\,dp. \tag{5}$$

Expressed in terms of the particle energy, these expressions assume the form

$$\Sigma(E) \approx \frac{V}{h^3}\frac{4\pi}{3}(2mE)^{3/2} \tag{6}$$

and

$$a(\varepsilon)\,d\varepsilon = \frac{d\Sigma(\varepsilon)}{d\varepsilon}\,d\varepsilon \approx \frac{V}{h^3}2\pi(2m)^{3/2}\varepsilon^{1/2}\,d\varepsilon. \tag{7}$$

The next case we shall consider here is that of a one-dimensional *simple harmonic oscillator*. The classical expression for the Hamiltonian of this system is

$$H(q, p) = \frac{1}{2}kq^2 + \frac{1}{2m}p^2, \tag{8}$$

where k is the spring constant and m the mass of the oscillating particle. The space coordinate q and the momentum coordinate p of the system are given by

$$q = A\cos(\omega t + \phi), \quad p = m\dot{q} = -m\omega A\sin(\omega t + \phi), \tag{9}$$

A being the amplitude and ω the (angular) frequency of vibration:

$$\omega = \sqrt{(k/m)}. \tag{10}$$

The energy of the oscillator is a constant of the motion, and is given by

$$E = \tfrac{1}{2}m\omega^2A^2. \tag{11}$$

The phase-space trajectory of the representative point (q, p) of this system is determined by eliminating t between expressions (9) for $q(t)$ and $p(t)$; we obtain

$$\frac{q^2}{(2E/m\omega^2)} + \frac{p^2}{(2mE)} = 1, \tag{12}$$

which is an ellipse, with axes proportional to $\sqrt{E}$ and hence area proportional to E; to be precise, the area of this ellipse is $2\pi E/\omega$. Now, if we restrict the oscillator energy to the interval $(E-\frac{1}{2}\Delta, E+\frac{1}{2}\Delta)$, its representative point in the phase space will be confined to the region bounded by elliptical trajectories corresponding to the energy values $(E+\frac{1}{2}\Delta)$ and $(E-\frac{1}{2}\Delta)$. The "volume" (in this case, the area) of this region will be

$$\underset{\left(E-\frac{1}{2}\Delta\right)\le H(q,p)\le\left(E+\frac{1}{2}\Delta\right)}{\int\cdots\int}(dq\,dp)=\frac{2\pi\left(E+\frac{1}{2}\Delta\right)}{\omega}-\frac{2\pi\left(E-\frac{1}{2}\Delta\right)}{\omega}=\frac{2\pi\Delta}{\omega}. \tag{13}$$

According to quantum mechanics, the energy eigenvalues of the harmonic oscillator are given by

$$E_n=\left(n+\tfrac{1}{2}\right)\hbar\omega;\qquad n=0,1,2,\ldots \tag{14}$$

In terms of phase space, one could say that the representative point of the system must move along one of the "chosen" trajectories, as shown in Fig. 2.2; the area of the phase space between two consecutive trajectories, for which $\Delta=\hbar\omega$, is simply $2\pi\hbar$.[5] For arbitrary values of E and Δ, such that $E \gg \Delta \gg \hbar\omega$, the number of eigenstates within the allowed energy interval is very nearly equal to $\Delta/\hbar\omega$. Hence, the area of the phase space equivalent to one eigenstate is, asymptotically, given by

$$\omega_0=(2\pi\Delta/\omega)/(\Delta/\hbar\omega)=2\pi\hbar=h. \tag{15}$$

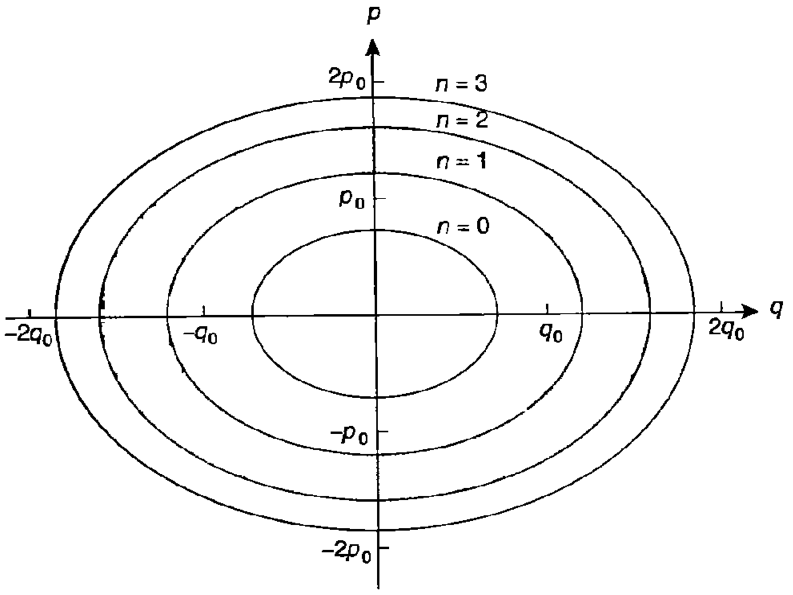


FIG. 2.2. Eigenstates of a linear harmonic oscillator, in relation to its phase space.

If, on the other hand, we consider a system of N harmonic oscillators along the same lines as above, we arrive at the result: $\omega_0=h^N$ (see Problem 2.7). Thus, our findings in these cases are consistent with our earlier result (3).

2.5. Quantum states and the phase space

At this stage we would like to say a few words on the central role played here by the Planck constant h. The best way to appreciate this role is to recall the implications of the Heisenberg uncertainty principle, according to which we cannot specify *simultaneously* both the position and the momentum of a particle exactly. An element of uncertainty is inherently present and can be expressed as follows: assuming that all conceivable uncertainties of measurement are eliminated, even then, by the very nature of things, the product of the uncertainties Δq and Δp in the *simultaneous* measurement of the canonically conjugate coordinates q and p would be of order $\hbar$:

$$(\Delta q \Delta p)_{\min} \sim \hbar. \tag{1}$$

Thus, it is impossible to define the position of a representative point in the phase space of the given system more accurately than is allowed by condition (1). In other words, around any point (q, p) in the (two-dimensional) phase space, there exists an area of order $\hbar$ within which the position of the representative point cannot be pin-pointed. In a phase space of $2f$ dimensions, the corresponding "volume of uncertainty" around any point would be of order $\hbar^f$. Therefore, it seems reasonable to regard the phase space as made up of elementary cells, of volume $\sim \hbar^f$, and to consider the various positions within such a cell as non-distinct. These cells could then be put into one-to-one correspondence with the quantum-mechanical states of the system.

It is, however, obvious that considerations of uncertainty alone cannot give us the *exact* value of the conversion factor ω_0. This could only be done by an *actual* counting of microstates on one hand and a computation of volume of the relevant region of the phase space on the other, as was done in the examples of the previous section. Clearly, a procedure along these lines could not be possible until after the work of Schrödinger and others. Historically, however, the first to establish the result (2.4.3) was Tetrode (1912) who, in his well-known work on the chemical constant and the entropy of a monatomic gas, assumed that

$$\omega_0 = (zh)^f \tag{2}$$

where z was supposed to be an unknown numerical factor. Comparing theoretical results with the experimental data on mercury, Tetrode found that z was very nearly equal to unity; from this he concluded that "it seems rather plausible that z is *exactly* equal to unity, as has already been taken by O. Sackur (1911)".[6]

In the extreme relativistic limit, the same result was established by Bose (1924). In his famous treatment of the photon gas, Bose made use of Einstein's relationship between the momentum of a photon and the frequency of the associated vibration, namely

$$p = \frac{h\nu}{c}, \tag{3}$$

and observed that, for a photon confined to a three-dimensional cavity of volume V, the relevant "volume" of the phase space,

$$\int (d^3q\, d^3p) = V 4\pi p^2\, dp = V(4\pi h^3 \nu^2/c^3)\, d\nu, \tag{4}$$

would correspond exactly to the Rayleigh expres

$$V(4\pi \nu^2/c^3)\,d\nu, \tag{5}$$

for the number of normal modes of a radiation oscillator, ***provided that*** we divide phase space into elementary cells of volume h^3 and put these cells into one-to-one correspondence with the vibrational modes of Rayleigh. It may, however, be added that a two-fold multiplicity of these states ($g = 2$) arises from the spin orientations of the photon (or from the states of polarization of the vibrational modes); this requires a multiplication of both expressions (4) and (5) by a factor of 2, leaving the conversion factor h^3 unchanged.

Problems

2.1. Show that the volume element

$$d\omega = \prod_{i=1}^{3N}(dq_i\,dp_i)$$

of the phase space remains *invariant* under a canonical transformation of the (generalized) coordinates (q, p) to any other set of (generalized) coordinates (Q, P).

[*Hint*: Before considering the most general transformation of this kind, which is referred to as a *contact* transformation, it may be helpful to consider a *point* transformation–one in which the new coordinates Q_i and the old coordinates q_i transform only among themselves.]

2.2. (a) Verify *explicitly* the invariance of the volume element $d\omega$ of the phase space of a single particle under transformation from the Cartesian coordinates (x, y, z, p_x, p_y, p_z) to the spherical polar coordinates $(r, \theta, \phi, p_r, p_\theta, p_\phi)$.

(b) The foregoing result seems to contradict the intuitive notion of "equal weights for equal solid angles", because the factor $\sin\theta$ is invisible in the expression for $d\omega$. Show that if we average out any physical quantity, whose dependence on p_θ and p_ϕ comes only through the kinetic energy of the particle, then as a result of integration over these variables we do indeed recover the factor $\sin\theta$ to appear with the sub-element $(d\theta\, d\phi)$.

2.3. Starting with the line of zero energy and working in the (two-dimensional) phase space of a classical rotator, draw lines of constant energy which divide phase space into cells of "volume" h. Calculate the energies of these states and compare them with the energy eigenvalues of the corresponding quantum-mechanical rotator.

2.4. By evaluating the "volume" of the relevant region of its phase space, show that the number of microstates available to a rigid rotator with angular momentum $\leq M$ is $(M/\hbar)^2$. Hence determine the number of microstates that may be associated with the quantized angular momentum $M_j = \sqrt{\{j(j+1)\}}\hbar$, where $j = 0, 1, 2, \ldots$ or $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$. Interpret the result physically.

[*Hint*: It simplifies to consider motion in the variables θ and φ, with $M^2 = p_\theta^2 + (p_\phi/\sin\theta)^2$.]

2.5. Consider a particle of energy E moving in a one-dimensional potential well $V(q)$, such that

$$m\hbar\left|\frac{dV}{dq}\right| \ll \{m(E-V)\}^{3/2}.$$

Show that the allowed values of the momentum p of the particle are such that

$$\oint p\,dq = \left(n + \tfrac{1}{2}\right)h,$$

where n is an integer.

2.6. The generalized coordinates of a simple pendulum are the angular displacement θ and the angular momentum $ml^2\dot{\theta}$. Study, both mathematically and graphically, the nature of the corresponding trajectories in the phase space of the system, and show that the area A enclosed by a trajectory is equal to the product of the total energy E and the time period τ of the pendulum.

2.7. Derive (i) an *asymptotic* expression for the number of ways in which a given energy E can be distributed among a set of N one-dimensional harmonic oscillators, the energy eigenvalues of the oscillators being $(n + \frac{1}{2})\hbar\omega$; $n = 0, 1, 2, \ldots$, and (ii) the corresponding expression for the "volume" of the relevant region of the phase space of this system. Establish the correspondence between the two results, showing that the conversion factor ω_0 is precisely h^N.

2.8. Following the method of Appendix C, replacing eqn. (C.4) by the integral

$$\int_0^\infty e^{-r} r^2 \, dr = 2,$$

show that

$$V_{3N} = \underset{0 \le \sum_{i=1}^{N} r_i \le R}{\int \cdots \int} \prod_{i=1}^{N} (4\pi r_i^2 \, dr_i) = (8\pi R^3)^N / (3N)!.$$

Using this result, compute the "volume" of the relevant region of the phase space of an extreme relativistic gas ($\varepsilon = pc$) of N particles moving in three dimensions. Hence, derive expressions for the various thermodynamic properties of this system and compare your results with those of Problem 1.7.

2.9. (a) Solve the integral

$$\underset{0 \le \sum_{i=1}^{3N} |x_i| \le R}{\int \cdots \int} (dx_1 \ldots dx_{3N})$$

and use it to determine the "volume" of the relevant region of the phase space of an extreme relativistic gas ($\varepsilon = pc$) of $3N$ particles moving in one dimension. Determine, as well, the number of ways of distributing a given energy E among this system of particles and show that, asymptotically, $\omega_0 = h^{3N}$.

(b) Compare the thermodynamics of this system with that of the system considered in Problem 2.8.

Notes

[1] Note that (q, p) is a further abbreviation of $(q_i, p_i) \equiv (q_1, \ldots, q_{3N}, p_1, \ldots, p_{3N})$.

[2] This means that in the ensemble under consideration neither are any new members being admitted nor are any old ones being expelled.

[3] We recall that the *Poisson bracket* $[\rho, H]$ stands for the sum

$$\sum_{i=1}^{3N} \left(\frac{\partial \rho}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial \rho}{\partial p_i} \frac{\partial H}{\partial q_i} \right),$$

which is identical with the group of terms in the middle of (8).

[4] To provide a *rigorous* justification for this assertion is not trivial. One can readily see that if, for any particular member of the ensemble, the quantity f is averaged only over a *short* span of time, the result is bound to depend upon the relevant "subset of microstates" through which the system passes during that time. In the phase space, this will mean an averaging over only a "part of the allowed region". However, if we employ instead a sufficiently long interval of time, the system may be expected to pass through *almost* all possible microstates "without fear or favor"; consequently, the result of the averaging process would depend only upon the macrostate of the system, and not upon a subset of microstates. Correspondingly, the averaging in the phase space would go over *practically* all parts of the allowed region, again "without fear or favor". In other words, the representative point of our system will have traversed each and every part of the allowed region *almost* uniformly. This statement embodies the so-called *ergodic theorem* or *ergodic hypothesis*, which was first introduced by Boltzmann (1871). According to this hypothesis, the trajectory of a representative point passes, in the course of time, through *each and every* point of the relevant region of the phase space. A little reflection, however, shows that the statement as such requires a qualification; we better replace it by the so-called *quasi-ergodic hypothesis*, according to which the trajectory of a representative point

traverses, in the course of time, *any neighborhood of any point* of the relevant region. For further details, see ter Haar (1954, 1955), Farquhar (1964).

Now, when we consider an ensemble of systems, the foregoing statement should hold for every member of the ensemble; thus, *irrespective of the initial (and final) states* of the various systems, the long-time average of any physical quantity f should be the same for every member system.

[5] Strictly speaking, the very concept of phase space is invalid in quantum mechanics because there it is wrong, in principle, to assign to a particle the coordinates q and p *simultaneously*. Nevertheless, the ideas discussed here are tenable in the correspondence limit.

[6] For a more satisfactory proof, see Sec. 5.5, especially eqn. (5.5.22).

CHAPTER 3

THE CANONICAL ENSEMBLE

IN THE preceding chapter we established the basis of ensemble theory and made a somewhat detailed study of the microcanonical ensemble. In that ensemble the macrostate of the systems was defined through a fixed number of particles N, a fixed volume V and a fixed energy E [or, preferably, a fixed energy range $(E - \frac{1}{2}\Delta, E + \frac{1}{2}\Delta)$]. The basic problem then consisted in determining the number $\Omega(N, V, E)$, or $\Gamma(N, V, E; \Delta)$, of *distinct* microstates accessible to the system. From the asymptotic expressions of these numbers, complete thermodynamics of the system could be derived in a straightforward manner. However, for most physical systems, the mathematical problem of determining these numbers is quite formidable. For this reason alone, a search for an alternative approach within the framework of the ensemble theory seems necessary.

Physically, too, the concept of a fixed energy (or even an energy range) for a system belonging to the real world does not appear satisfactory. For one thing, the total energy E of a system is hardly ever measured; for another, it is hardly possible to keep its value under strict physical control. A far better alternative appears to be to speak of a fixed temperature T of the system–a parameter which is not only directly observable (by placing a "thermometer" in contact with the system) but also controllable (by keeping the system in contact with an appropriate "heat reservoir"). For most purposes, the precise nature of the reservoir is not very relevant; all one needs is that it should have an infinitely large heat capacity, so that, irrespective of energy exchange between the system and the reservoir, an overall constant temperature can be maintained. Now, if the reservoir consists of an infinitely large number of mental copies of the given system we have once again an ensemble of systems–this time, however, it is an ensemble in which the macrostate of the systems is defined through the parameters N, V and T. Such an ensemble is referred to as a *canonical* ensemble.

In the canonical ensemble, the energy E of a system is variable; in principle, it can take values anywhere between zero and infinity. The question then arises: what is the probability that, at any time t, a system in the ensemble is found to be in one of the states characterized by the energy value E_r?[1] We denote this probability by the symbol P_r. Clearly, there are two ways in which the dependence of P_r on E_r can be determined. One consists in regarding the system as in equilibrium with a heat reservoir at a *common* temperature T and studying the statistics of the energy exchange between the two. The other consists in regarding the system as a member of a canonical ensemble (N, V, T), in which an energy $\mathscr{E}$ is being shared by $\mathcal{N}$ identical systems constituting the ensemble, and studying the statistics of this sharing process. We expect that in the thermodynamic limit the final result in either case would be the same. Once P_r is determined, the rest follows without difficulty.

3.1. Equilibrium between a system and a heat reservoir

We consider the given system A, immersed in a very large heat reservoir A'; see Fig. 3.1. On attaining a state of mutual equilibrium, the system and the reservoir would have a *common* temperature, T say. Their energies, however, would be variable and, in principle, could have, at any time t, values lying anywhere between 0 and $E^{(0)}$, where $E^{(0)}$ denotes the energy of the composite system $A^{(0)}(\equiv A + A')$. If, at any particular instant of time, the system A happens to be in a state characterized by the energy value E_r, then the reservoir would have an energy E'_r such that

$$E_r + E'_r = E^{(0)} = \text{const.} \tag{1}$$

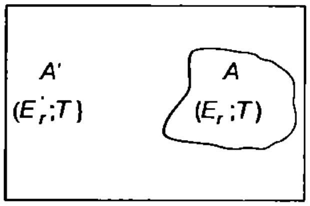


FIG. 3.1. A given system A immersed in a heat reservoir A'; in equilibrium, the two have a common temperature T.

Of course, since the reservoir is supposed to be much larger than the given system, any *practical* value of E_r would be a very small fraction of $E^{(0)}$; therefore, for all practical purposes,

$$\frac{E_r}{E^{(0)}} = \left(1 - \frac{E'_r}{E^{(0)}}\right) \ll 1. \tag{2}$$

With, the state of the system A having been specified, the reservoir A' can still be in *any one* of a large number of states compatible with the energy value E'_r. Let the number of these states be denoted by $\Omega'(E'_r)$. The prime on the symbol Ω emphasizes the fact that its functional form will depend upon the nature of the reservoir; of course, the details of this dependence are not going to be of any particular relevance to our final results. Now, the larger the number of states available to the reservoir, the larger the probability of the reservoir assuming that particular energy value E'_r (and, hence, of the system A assuming the corresponding energy value E_r). Moreover, since the various possible states (with a given energy value) are *equally likely* to occur, the relevant probability would be directly proportional to this number; thus,

$$P_r \propto \Omega'(E'_r) \equiv \Omega'(E^{(0)} - E_r). \tag{3}$$

In view of (2), we may carry out an expansion of (3) around the value $E'_r = E^{(0)}$, i.e. around $E_r = 0$. However, for reasons of convergence, it is essential to effect the expansion of its logarithm instead:

$$\begin{aligned}\ln \Omega'(E'_r) &= \ln \Omega'(E^{(0)}) + \left(\frac{\partial \ln \Omega'}{\partial E'}\right)_{E'=E^{(0)}} (E'_r - E^{(0)}) + \cdots \\ &\simeq \text{const} - \beta' E_r, \end{aligned} \tag{4}$$

where use has been made of formula (1.2.3), whereby

$$\left(\frac{\partial \ln \Omega}{\partial E}\right)_{N,V} \equiv \beta; \tag{5}$$

note that, in equilibrium, $\beta' = \beta = 1/kT$. From (3) and (4), we obtain the desired result:

$$P_r \propto \exp(-\beta E_r). \tag{6}$$

Normalizing (6), we get

$$P_r = \frac{\exp(-\beta E_r)}{\sum_r \exp(-\beta E_r)}, \tag{7}$$

where the summation in the denominator goes over *all* states accessible to the system A. We note that our final formula (7) bears no relation whatsoever to the physical nature of the reservoir A'.

We now examine the same problem from the ensemble point of view.

3.2. A system in the canonical ensemble

We consider an ensemble of $\mathcal{N}$ identical systems (which may be labelled as $1, 2, \ldots, \mathcal{N}$), sharing a total energy $\mathcal{E}$; let $E_r (r = 0, 1, 2, \ldots)$ denote the energy eigenvalues of the systems. If n_r denotes the number of systems which, at any time t, have the energy value E_r, then the set of numbers $\{n_r\}$ must satisfy the obvious conditions

$$\left.\begin{aligned} \sum_r n_r &= \mathcal{N} \\ \sum_r n_r E_r &= \mathcal{E} = \mathcal{N} U, \end{aligned}\right\} \tag{1}$$

where $U(= \mathcal{E}/\mathcal{N})$ denotes the average energy per system in the ensemble. Any set $\{n_r\}$ which satisfies the restrictive conditions (1) represents a possible mode of distribution of the total energy $\mathcal{E}$ among the $\mathcal{N}$ members of the ensemble. Furthermore, any such mode can be realized in a number of ways, for we may effect a reshuffle among those members of the ensemble for which the energy values are different and thereby obtain a state of the ensemble which is distinct from the original one. Denoting the number of different ways of doing so by the symbol $W\{n_r\}$, we have

$$W\{n_r\} = \frac{\mathcal{N}!}{n_0! n_1! n_2! \ldots}. \tag{2}$$

In view of the fact that all possible states of the ensemble, which are compatible with conditions (1), are *equally likely* to occur, the frequency with which the distribution set $\{n_r\}$ may appear will be directly proportional to the number $W\{n_r\}$. Accordingly, the "most probable" mode of distribution will be the one for which the number W is a maximum. We denote the corresponding distribution set by $\{n_r^*\}$; clearly, the set $\{n_r^*\}$ must also satisfy conditions (1). As will be seen in the sequel, the probability of appearance of other modes of distribution, however little they may be differing from the most probable mode, is extremely low! Therefore, for all practical purposes, the *most probable distribution set* $\{n_r^*\}$ is the only one we have to contend with.

However, unless this is mathematically demonstrated, one must take into account *all* possible modes of distribution, as characterized by the various distribution sets $\{n_r\}$, along with their respective weight factors $W\{n_r\}$. Accordingly, the *expectation values*, or *mean values*, $\langle n_r \rangle$ of the numbers n_r would be given by

$$\langle n_r \rangle = \frac{\sum_{\{n_r\}}' n_r W\{n_r\}}{\sum_{\{n_r\}}' W\{n_r\}}, \tag{3}$$

where the primed summations go over all distribution sets that conform to conditions (1). In principle, the mean value $\langle n_r \rangle$, as a fraction of the total number $\mathcal{N}$, should be a natural analogue of the probability P_r evaluated in the preceding section. In practice, however, the fraction $n_r^*/\mathcal{N}$ is also the same.

We now proceed to derive expressions for the numbers n_r^* and $\langle n_r \rangle$, and to show that, in the limit $\mathcal{N} \to \infty$, they are identical.

(i) *The method of most probable values*. Our aim here is to determine that distribution set which, while satisfying conditions (1), maximizes the weight factor (2). For simplicity, we work with $\ln W$ instead:

$$\ln W = \ln(\mathcal{N}!) - \sum_r \ln(n_r!). \tag{4}$$

Since in the end we propose to resort to the limit $\mathcal{N} \to \infty$, the values of n_r (which are going to be of any practical significance) would also, in that limit, tend to infinity. It is, therefore, justified to apply the Stirling formula, $\ln(n!) \approx n \ln n - n$, to (4) and write

$$\ln W = \mathcal{N} \ln \mathcal{N} - \sum_r n_r \ln n_r. \tag{5}$$

If we shift from the set $\{n_r\}$ to a slightly different set $\{n_r + \delta n_r\}$, then expression (5) would change by an amount

$$\delta(\ln W) = -\sum_r (\ln n_r + 1)\, \delta n_r. \tag{6}$$

Now, if the set $\{n_r\}$ is maximal, the variation $\delta(\ln W)$ should vanish. At the same time, in view of the restrictive conditions (1), the variations δn_r themselves must satisfy the conditions

$$\left.\begin{aligned} \sum_r \delta n_r &= 0 \\ \sum_r E_r\, \delta n_r &= 0. \end{aligned}\right\} \tag{7}$$

The desired set $\{n_r^*\}$ is then determined by the method of *Lagrange multipliers*,[2] by which the condition determining this set becomes

$$\sum_r \{-(\ln n_r^* + 1) - \alpha - \beta E_r\}\,\delta n_r = 0, \tag{8}$$

where α and β are the Lagrangian undetermined multipliers that take care of the restrictive conditions (7). In (8), the variations δn_r become completely arbitrary; accordingly, the only way to satisfy this condition is that all its coefficients must vanish identically, i.e. for *all r*,

$$\ln n_r^* = -(\alpha + 1) - \beta E_r,$$

whence

$$n_r^* = C \exp(-\beta E_r), \tag{9}$$

where C is again an undetermined parameter. To determine C and β, we subject (9) to conditions (1), with the result that

$$\frac{n_r^*}{\mathcal{N}} = \frac{\exp(-\beta E_r)}{\sum_r \exp(-\beta E_r)}, \tag{10}$$

the parameter β being a solution of the equation

$$\frac{\mathcal{E}}{\mathcal{N}} = U = \frac{\sum_r E_r \exp(-\beta E_r)}{\sum_r \exp(-\beta E_r)}. \tag{11}$$

Combining statistical considerations with thermodynamic ones, see Sec. 3.3, we can show that the parameter β here is exactly the same as the one appearing in Sec. 3.1, i.e. $\beta = 1/kT$.

(ii) *The method of mean values.* Here we attempt to evaluate expression (3) for $\langle n_r \rangle$, taking into account the weight factors (2) and the restrictive conditions (1). To do this, we replace (2) by

$$\tilde{W}\{n_r\} = \frac{\mathcal{N}!\,\omega_0^{n_0}\omega_1^{n_1}\omega_2^{n_2}\ldots}{n_0!\,n_1!\,n_2!\ldots}, \tag{12}$$

with the understanding that in the end all the ω_r will be set equal to unity, and introduce a function

$$\Gamma(\mathcal{N}, U) = \sum_{\{n_r\}}{}' \tilde{W}\{n_r\}. \tag{13}$$

where the primed summation, as before, goes over all distribution sets that conform to conditions (1). Expression (3) can then be written as

$$\langle n_r \rangle = \omega_r \frac{\partial}{\partial \omega_r}(\ln \Gamma)\bigg|_{\text{all } \omega_r = 1} \tag{14}$$

Thus, all we need to know here is the dependence of the quantity $\ln \Gamma$ on the parameters ω_r. Now

$$\Gamma(\mathcal{N}, U) = \mathcal{N}! \sum_{\{n_r\}}{}' \left(\frac{\omega_0^{n_0}}{n_0!} \cdot \frac{\omega_1^{n_1}}{n_1!} \cdot \frac{\omega_2^{n_2}}{n_2!} \cdots \right), \tag{15}$$

but the summation appearing here cannot be evaluated explicitly because it is restricted to those sets only which conform to the pair of conditions (1). If our distribution sets were restricted by the condition $\sum_r n_r = \mathcal{N}$ alone, then the evaluation of (15) would have been trivial; by the multinomial theorem, $\Gamma(\mathcal{N})$ would have been simply $(\omega_0 + \omega_1 + \cdots)^{\mathcal{N}}$. The added restriction $\sum_r n_r E_r = \mathcal{N} U$, however, permits the inclusion of only a "limited" number of terms in the sum–and that constitutes the real difficulty of the problem. Nevertheless, we can still hope to make some progress because, from a physical point of view, we do not require anything more than an *asymptotic* result–one that holds in the limit $\mathcal{N} \to \infty$. The method commonly used for this purpose is the one developed by Darwin and Fowler (1922a,b, 1923), which itself makes use of the *saddle-point method* of integration or the so-called *method of steepest descent.*

We construct a *generating function* $G(\mathcal{N}, z)$ for the quantity $\Gamma(\mathcal{N}, U)$:

$$G(\mathcal{N}, z) = \sum_{U=0}^{\infty} \Gamma(\mathcal{N}, U) z^{\mathcal{N} U} \tag{16}$$

which, in view of eqn. (15) and the second of the restrictive conditions (1), may be written as

$$G(\mathcal{N}, z) = \sum_{U=0}^{\infty} \left[\sum_{\{n_r\}}{}' \frac{\mathcal{N}!}{n_0! n_1! \ldots} \left(\omega_0 z^{E_0}\right)^{n_0} \left(\omega_1 z^{E_1}\right)^{n_1} \ldots \right]. \tag{17}$$

It is easy to see that the summation over *doubly* restricted sets $\{n_r\}$, followed by a summation over all possible values of U, is equivalent to a summation over *singly* restricted sets $\{n_r\}$, viz. the ones that satisfy only one condition: $\sum_r n_r = \mathcal{N}$. Expression (17) can, therefore, be evaluated with the help of the multinomial theorem, with the result

$$\begin{aligned} G(\mathcal{N}, z) &= \left(\omega_0 z^{E_0} + \omega_1 z^{E_1} + \cdots\right)^{\mathcal{N}} \\ &= [f(z)]^{\mathcal{N}}, \text{ say.} \end{aligned} \tag{18}$$

Now, if we suppose that the E_r (and hence the total energy values $\mathcal{E} = \mathcal{N} U$) are all integers, then, by (16), the quantity $\Gamma(\mathcal{N}, U)$ is simply the coefficient of $z^{\mathcal{N} U}$ in the expansion of the function $G(\mathcal{N}, z)$ as a power series in z. It can, therefore, be evaluated by the method of residues in the complex z-plane.

To make this plan work, we assume to have chosen, *right at the outset*, a unit of energy so small that, to any desired degree of accuracy, we can regard the energies E_r (and the prescribed total energy $\mathcal{N} U$) as integral multiples of this unit. In terms of this unit, any energy value we come across will be an integer. We further assume, without loss of generality, that the sequence $E_0, E_1, \ldots$ is a *nondecreasing* sequence, *with no common divisor*;[3] also, for the sake of simplicity, we assume that $E_0 = 0$.[4] The solution now is

$$\Gamma(\mathcal{N}, U) = \frac{1}{2\pi i} \oint \frac{[f(z)]^{\mathcal{N}}}{z^{\mathcal{N} U + 1}} dz, \tag{19}$$

where the integration is carried along any closed contour around the origin; of course, we should stay *within* the circle of convergence of the function $f(z)$, so that a need for analytic continuation does not arise.

First of all we examine the behavior of the integrand as we proceed from the origin along the real positive axis, remembering that all our ω are virtually equal to unity and that $0 = E_0 \leq E_1 \leq E_2 \cdots$. We find that the factor $[f(z)]^N$ starts from the value 1 at $z = 0$, increases monotonically and tends to infinity as z approaches the circle of convergence of $f(z)$, wherever that may be. The factor $z^{-(NU+1)}$, on the other hand, starts from a positive, infinite value at $z = 0$ and decreases monotonically as z increases. Moreover, the relative rate of increase of the factor $[f(z)]^N$ itself increases monotonically while the relative rate of decrease of the factor $z^{-(NU+1)}$ decreases monotonically. Under these circumstances, the integrand must exhibit a minimum (and no other extremum) at some value of z, say x_0, within the circle of convergence. And, in view of the largeness of the numbers N and NU, this minimum may indeed be very steep!

Thus, at $z = x_0$ the first derivative of the integrand must vanish, while the second derivative must be positive and, hopefully, very large. Accordingly, if we proceed through the point $z = x_0$ in a direction orthogonal to the real axis, the integrand must exhibit an equally steep maximum.[5] Thus, in the complex z-plane, as we move along the real axis our integrand shows a minimum at $z = x_0$, whereas if we move along a path parallel to the imaginary axis but passing through the point $z = x_0$, the integrand shows a maximum there. It is natural to call the point x_0 a *saddle point*; see Fig. 3.2. For the contour of integration we choose a circle, with center at $z = 0$ and radius equal to x_0, hoping that on integration along this contour only the immediate neighborhood of the sharp maximum at the point x_0 will make the most dominant contribution to the value of the integral.[6]

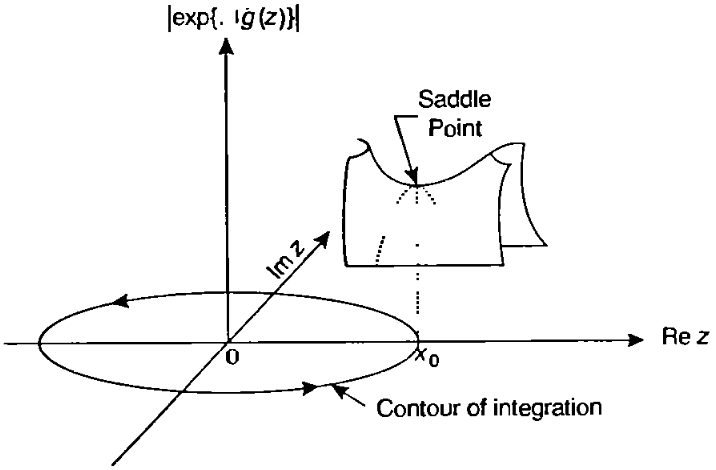


FIG. 3.2. The saddle point.

To carry out the integration we first locate the point x_0. For this we write our integrand as

$$\frac{[f(z)]^N}{z^{NU+1}} = \exp[Ng(z)]. \tag{20}$$

where

$$g(z) = \ln f(z) - \left(U + \frac{1}{\mathcal{N}}\right) \ln z, \tag{21}$$

while

$$f(z) = \sum_r \omega_r z^{E_r}. \tag{22}$$

The number x_0 is then determined by the equation

$$g'(x_0) = \frac{f'(x_0)}{f(x_0)} - \frac{\mathcal{N}U + 1}{\mathcal{N}x_0} = 0 \tag{23}$$

which, in view of the fact that $\mathcal{N}U \gg 1$, can be written as

$$U \approx x_0 \frac{f'(x_0)}{f(x_0)} = \frac{\sum_r \omega_r E_r x_0^{E_r}}{\sum_r \omega_r x_0^{E_r}}. \tag{24}$$

We further have

$$\begin{aligned} g''(x_0) &= \left(\frac{f''(x_0)}{f(x_0)} - \frac{[f'(x_0)]^2}{[f(x_0)]^2}\right) + \frac{\mathcal{N}U + 1}{\mathcal{N}x_0^2} \\ &\approx \frac{f''(x_0)}{f(x_0)} - \frac{U^2 - U}{x_0^2}. \end{aligned} \tag{25}$$

It will be noted here that, in the limit $\mathcal{N} \to \infty$ and $\mathcal{E}(\equiv \mathcal{N}U) \to \infty$, *with U staying constant*, the number x_0 and the quantity $g''(x_0)$ become independent of $\mathcal{N}$.

Expanding $g(z)$ about the point $z = x_0$, along the direction of integration, i.e. along the line $z = x_0 + iy$, we have

$$g(z) = g(x_0) - \tfrac{1}{2} g''(x_0) y^2 + \cdots;$$

accordingly, the integrand (20) might be approximated as

$$\frac{[f(x_0)]^{\mathcal{N}}}{x_0^{\mathcal{N}U+1}} \exp\left[-\frac{\mathcal{N}}{2} g''(x_0) y^2\right]. \tag{26}$$

Equation (19) then gives

$$\begin{aligned} \Gamma(\mathcal{N}, U) &\simeq \frac{1}{2\pi i} \frac{[f(x_0)]^{\mathcal{N}}}{x_0^{\mathcal{N}U+1}} \int_{-\infty}^{\infty} \exp\left[-\frac{\mathcal{N}}{2} g''(x_0) y^2\right] i\,dy \\ &= \frac{[f(x_0)]^{\mathcal{N}}}{x_0^{\mathcal{N}U+1}} \cdot \frac{1}{\{2\pi \mathcal{N} g''(x_0)\}^{1/2}}, \end{aligned} \tag{27}$$

whence

$$\frac{1}{\mathcal{N}} \ln \Gamma(\mathcal{N}, U) = \{\ln f(x_0) - U \ln x_0\} - \frac{1}{\mathcal{N}} \ln x_0 - \frac{1}{2\mathcal{N}} \ln \{2\pi \mathcal{N} g''(x_0)\}. \tag{28}$$

In the limit $\mathcal{N} \to \infty$ (with U staying constant), the last two terms in this expression tend to zero, with the result

$$\frac{1}{\mathcal{N}} \ln \Gamma(\mathcal{N}, U) = \ln f(x_0) - U \ln x_0. \tag{29}$$

Substituting for $f(x_0)$ and introducing a new variable β, defined by the relationship

$$x_0 \equiv \exp(-\beta), \tag{30}$$

we get

$$\frac{1}{\mathcal{N}} \ln \Gamma(\mathcal{N}, U) = \ln \left\{ \sum_r \omega_r \exp(-\beta E_r) \right\} + \beta U. \tag{31}$$

The expectation value of the number n_r then follows from (14) and (31):

$$\frac{\langle n_r \rangle}{\mathcal{N}} = \left[\frac{\omega_r \exp(-\beta E_r)}{\sum_r \omega_r \exp(-\beta E_r)} + \left\{ -\frac{\sum_r \omega_r E_r \exp(-\beta E_r)}{\sum_r \omega_r \exp(-\beta E_r)} + U \right\} \omega_r \frac{\partial \beta}{\partial \omega_r} \right]_{\text{all } \omega_r = 1} \tag{32}$$

The term inside the curly brackets vanishes identically because of (24) and (30). It has been included here to emphasize the fact that, for a fixed value of U, the number $\beta (\equiv -\ln x_0)$ in fact depends upon the choice of the ω_r; see (24). We will appreciate the importance of this fact when we evaluate the mean square fluctuation in the number n_r; in the calculation of the expectation value of n_r, this does not really matter. We thus obtain

$$\frac{\langle n_r \rangle}{\mathcal{N}} = \frac{\exp(-\beta E_r)}{\sum_r \exp(-\beta E_r)}, \tag{33}$$

which is identical with expression (10) for $n_r^*/\mathcal{N}$. The physical significance of the parameter β is also the same as in that expression, for it is determined by eqn. (24), with all $\omega_r = 1$, i.e. by eqn. (11) which fits naturally with eqn. (33) because U is nothing but the ensemble average of the variable E_r:

$$U = \sum_r E_r P_r = \frac{1}{\mathcal{N}} \sum_r E_r \langle n_r \rangle. \tag{34}$$

Finally, we compute fluctuations in the values of the numbers n_r. We have, first of all,

$$\langle n_r^2 \rangle \equiv \frac{\sum'_{\{n_r\}} n_r^2 W\{n_r\}}{\sum'_{\{n_r\}} W\{n_r\}} = \frac{1}{\Gamma} \left(\omega_r \frac{\partial}{\partial \omega_r} \right) \left(\omega_r \frac{\partial}{\partial \omega_r} \right) \Gamma \bigg|_{\text{all } \omega_r = 1}; \tag{35}$$

see eqns (12)–(14). It follows that

$$\langle (\Delta n_r)^2 \rangle \equiv \langle \{ n_r - \langle n_r \rangle \}^2 \rangle = \langle n_r^2 \rangle - \langle n_r \rangle^2 = \left(\omega_r \frac{\partial}{\partial \omega_r} \right) \left(\omega_r \frac{\partial}{\partial \omega_r} \right) (\ln \Gamma) \bigg|_{\text{all } \omega_r = 1}. \tag{36}$$

Substituting from (31) and making use of (32), we get

$$\frac{\langle(\Delta n_r)^2\rangle}{\mathcal{N}} = \omega_r\frac{\partial}{\partial\omega_r}\left[\frac{\omega_r\exp(-\beta E_r)}{\sum\limits_r \omega_r\exp(-\beta E_r)} + \left\{-\frac{\sum\limits_r \omega_r E_r\exp(-\beta E_r)}{\sum\limits_r \omega_r\exp(-\beta E_r)} + U\right\}\omega_r\frac{\partial\beta}{\partial\omega_r}\right]_{\text{all }\omega_r=1} . \qquad \text{(37}$$

We note that the term in the curly brackets would not make any contribution because it is identically zero, *whatever the choice of the* ω_r. However, in the differentiation of the first term, we must not forget to take into account the *implicit* dependence of β on the ω_r, which arises from the fact that unless the ω are set equal to unity the relation determining β does contain ω_r; see eqns (24) and (30), whereby

$$U = \left.\frac{\sum\limits_r \omega_r E_r\exp(-\beta E_r)}{\sum\limits_r \omega_r\exp(-\beta E_r)}\right|_{\text{all }\omega_r=1} . \qquad \text{(38}$$

A straightforward calculation gives

$$\left.\left(\frac{\partial\beta}{\partial\omega_r}\right)_U\right|_{\text{all }\omega_r=1} = \frac{E_r - U}{\langle E_r^2\rangle - U^2}\frac{\langle n_r\rangle}{\mathcal{N}} . \qquad \text{(39}$$

We can now evaluate (37), with the result

$$\begin{aligned}\frac{\langle(\Delta n_r)^2\rangle}{\mathcal{N}} &= \frac{\langle n_r\rangle}{\mathcal{N}} - \left(\frac{\langle n_r\rangle}{\mathcal{N}}\right)^2 + \frac{\langle n_r\rangle}{\mathcal{N}}(U - E_r)\left.\left(\frac{\partial\beta}{\partial\omega_r}\right)_U\right|_{\text{all }\omega_r=1}\\ &= \frac{\langle n_r\rangle}{\mathcal{N}}\left[1 - \frac{\langle n_r\rangle}{\mathcal{N}} - \frac{\langle n_r\rangle}{\mathcal{N}}\frac{(E_r - U)^2}{\langle(E_r - U)^2\rangle}\right] .\end{aligned} \qquad \text{(40}$$

For the relative fluctuation in n_r, we have

$$\left\langle\left(\frac{\Delta n_r}{\langle n_r\rangle}\right)^2\right\rangle = \frac{1}{\langle n_r\rangle} - \frac{1}{\mathcal{N}}\left\{1 + \frac{(E_r - U)^2}{\langle(E_r - U)^2\rangle}\right\} . \qquad \text{(41}$$

As $\mathcal{N} \to \infty$, $\langle n_r\rangle$ also $\to \infty$, with the result that the relative fluctuations in n tend to zero; accordingly, the (canonical) distribution becomes infinitely sharp an with it the mean value, the most probable value–in fact, any values of n_r th appear with nonvanishing probability–become essentially the same. And that is t reason why two wildly different methods of obtaining the canonical distributic followed in this section have led to identical results.

3.3. Physical significance of the various statistical quantities in the canonical ensemble

We start with the *canonical distribution*

$$P_r \equiv \frac{\langle n_r\rangle}{\mathcal{N}} = \frac{\exp(-\beta E_r)}{\sum\limits_r \exp(-\beta E_r)}, \tag{1}$$

where β is determined by the equation

$$U = \frac{\sum\limits_r E_r \exp(-\beta E_r)}{\sum\limits_r \exp(-\beta E_r)} = -\frac{\partial}{\partial\beta}\ln\left\{\sum_r \exp(-\beta E_r)\right\} \tag{2}$$

We now look for a general recipe to extract information about the various macroscopic properties of the given system on the basis of the foregoing statistical results. For this we recall certain thermodynamic relationships involving the Helmholtz free energy $A(= U - TS)$, namely

$$dA = dU - T\,dS - S\,dT = -S\,dT - P\,dV + \mu\,dN, \tag{3}$$

$$S = -\left(\frac{\partial A}{\partial T}\right)_{N,V} \quad P = -\left(\frac{\partial A}{\partial V}\right)_{N,T} \quad \mu = \left(\frac{\partial A}{\partial N}\right)_{V,T}, \tag{4}$$

and

$$U = A + TS = A - T\left(\frac{\partial A}{\partial T}\right)_{N,V} = -T^2\left[\frac{\partial}{\partial T}\left(\frac{A}{T}\right)\right]_{N,V} = \left[\frac{\partial(A/T)}{\partial(1/T)}\right]_{N,V} \tag{5}$$

where the various symbols have their usual meanings. Comparing (5) with (2), we infer that there exists a close correspondence between the quantities that enter through the statistical treatment and the ones that come from thermodynamics, viz.

$$\beta = \frac{1}{kT}, \quad \ln\left\{\sum_r \exp(-\beta E_r)\right\} = -\frac{A}{kT}, \tag{6}$$

where k is a universal constant yet to be determined; soon we shall see that k is indeed the *Boltzmann constant.*

Equations (6) constitute the most fundamental result of the canonical ensemble theory. Customarily, we write it in the form

$$A(N, V, T) = -kT\ln Q_N(V, T), \tag{7}$$

where

$$Q_N(V, T) = \sum_r \exp(-E_r/kT). \tag{8}$$

The quantity $Q_N(V, T)$ is referred to as the *partition function* of the system; sometimes, it is also called the "sum-over-states" (German: *Zustandssumme*). The dependence of Q on T is quite obvious. The dependence on N and V comes through the energy eigenvalues E_r; in fact, any other parameters that might govern the values E_r should also appear in the argument of Q. Moreover, for the quantity $A(N, V, T)$ to be an extensive property of the system, $\ln Q$ must also be an extensive quantity.

Once the Helmholtz free energy is known, the rest of the thermodynamic quantities follow straightforwardly. While the entropy, the pressure and the chemical potential are obtained from formulae (4), the specific heat at constant volume follows from

$$C_V = \left(\frac{\partial U}{\partial T}\right)_{N,V} = -T\left(\frac{\partial^2 A}{\partial T^2}\right)_{N,V} \tag{9}$$

and the Gibbs free energy from

$$G = A + PV = A - V\left(\frac{\partial A}{\partial V}\right)_{N,T} = N\left(\frac{\partial A}{\partial N}\right)_{V,T} = N\mu; \tag{10}$$

see Problem 3.5.

At this stage it appears worthwhile to make a few remarks on the foregoing results. First of all, we note from eqns (4) and (6) that the pressure P is given by

$$P = -\frac{\sum_r \frac{\partial E_r}{\partial V}\exp(-\beta E_r)}{\sum_r \exp(-\beta E_r)}, \tag{11}$$

so that

$$P\,dV = -\sum_r P_r\,dE_r = -dU. \tag{12}$$

The quantity on the right-hand side of this equation is clearly the change in the average energy of a system (in the ensemble) during a process that alters the energy levels E_r, leaving the probabilities P_r unchanged. The left-hand side then tells us that the volume change dV provides an example of such a process, and the pressure P is the "force" accompanying that process. The quantity P, which was introduced here through the thermodynamic relationship (3), thus acquires a mechanical meaning as well.

Next, about the entropy. Since $P_r = Q^{-1}\exp(-\beta E_r)$, it follows that

$$\langle \ln P_r\rangle = -\ln Q - \beta\langle E_r\rangle = \beta(A - U) = -S/k,$$

with the result that

$$S = -k\langle \ln P_r\rangle = -k\sum_r P_r \ln P_r. \tag{13}$$

This is an extremely interesting relationship, for it shows that the entropy of a physical system is *solely* and *completely* determined by the probability values P_r (of the system being in different dynamical states accessible to it)!

From the very look of it, eqn. (13) appears to be of fundamental importance; indeed, it admits of a number of interesting conclusions. One of these relates to a system in its ground state ($T = 0$ K). If the ground state is unique, then the system is sure to be found in this particular state and in no other; consequently, P_r is equal to 1 for this state and 0 for all others. Equation (13) then tells us that the entropy of the system is precisely zero, which is essentially the content of the *Nernst heat theorem* or the *third law of thermodynamics.*[7] We also infer that vanishing entropy and perfect statistical order (which implies complete predictability about the system) go together. As the number of accessible states increases, more and more of the P become nonzero; the entropy of the system thereby increases. As the number of states becomes exceedingly large, most of the P-values become exceedingly small (and their logarithms assume large, negative values); the net result is that the entropy becomes exceedingly large. Thus, the largeness of entropy and the high degree of statistical disorder (or unpredictability) in the system also go together.

It is because of this fundamental connection between entropy on one hand and lack of information on the other that formula (13) became the starting point of the pioneering work of Shannon (1948, 1949) in the theory of communication.

It may be pointed out that formula (13) applies in the microcanonical ensemble as well. There, for each member system of the ensemble, we have a group of Ω states, all *equally likely* to occur. The value of P_r is, then, $1/\Omega$ for each of these states and 0 for all others. Consequently,

$$S = -k \sum_{r=1}^{\Omega} \left\{ \frac{1}{\Omega} \ln \left(\frac{1}{\Omega} \right) \right\} = k \ln \Omega, \tag{14}$$

which is precisely the central result in the microcanonical ensemble theory; see eqn. (1.2.6) or (2.3.6).

3.4. Alternative expressions for the partition function

In most physical cases the energy levels accessible to a system are *degenerate*, i.e. one has a group of states, g_i in number, all belonging to the same energy value E_i. In such cases it is more useful to write the partition function (3.3.8) as

$$Q_N(V, T) = \sum_i g_i \exp(-\beta E_i); \tag{1}$$

the corresponding expression for P_i, the probability that the system be in a state with energy E_i, would be

$$P_i = \frac{g_i \exp(-\beta E_i)}{\sum_i g_i \exp(-\beta E_i)}. \tag{2}$$

Clearly, the g_i states with a common energy E_i are all equally likely to occur. As a result, the probability of a system having energy E_i becomes proportional to the multiplicity g_i of this level; g_i thus plays the role of a "weight factor" for the level E_i. The actual probability is then determined by the weight factor g_i as well as by the Boltzmann factor $\exp(-\beta E_i)$ of the level, as we have in (2). The basic relations established in the preceding section remain unaffected.

Now, in view of the largeness of the number of particles constituting a given system and the largeness of the volume to which these particles are confined, the consecutive energy values E_i of the system are, in general, very close to one another. Accordingly, there lie, within a reasonable interval of energy $(E, E+dE)$, a very large number of energy levels. One may then regard E as a *continuous* variable and write $P(E)\,dE$ for the probability that the given system, as a member of the canonical ensemble, may have its energy in the range $(E, E+dE)$. Clearly, this will be given by the product of the relevant single-state probability and the number of energy states lying in the specified range. Denoting the latter by $g(E)\,dE$, where $g(E)$ denotes the *density of states* around the energy value E, we have

$$P(E)\,dE \propto \exp(-\beta E)g(E)\,dE \tag{3}$$

which, on normalization, becomes

$$P(E)\,dE = \frac{\exp(-\beta E)g(E)\,dE}{\int_0^\infty \exp(-\beta E)g(E)\,dE} \tag{4}$$

The denominator is clearly another expression for the partition function of the system:

$$Q_N(V,T) = \int_0^\infty e^{-\beta E}g(E)\,dE. \tag{5}$$

The expression for $\langle f\rangle$, the expectation value of a physical quantity f, may now be written as

$$\langle f\rangle \equiv \sum_i f_i P_i = \frac{\sum_i f(E_i)g_i e^{-\beta E_i}}{\sum_i g_i e^{-\beta E_i}} \to \frac{\int_0^\infty f(E)e^{-\beta E}g(E)\,dE}{\int_0^\infty e^{-\beta E}g(E)\,dE}. \tag{6}$$

Before proceeding further, we take a closer look at eqn. (5). With $\beta > 0$, the partition function $Q(\beta)$ is just the Laplace transform of the density of states $g(E)$. We may, therefore, write $g(E)$ as the inverse Laplace transform of $Q(\beta)$:

$$g(E) = \frac{1}{2\pi i}\int_{\beta'-i\infty}^{\beta'+i\infty} e^{\beta E}Q(\beta)\,d\beta \quad (\beta' > 0) \tag{7}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{(\beta'+i\beta'')E}Q(\beta'+i\beta'')\,d\beta''. \tag{8}$$

where β is now treated as a complex variable, $\beta' + i\beta''$, while the path of integration runs parallel to, and to the right of, the imaginary axis, i.e. along the straight line Re $\beta = \beta' > 0$. Of course, the path may be continuously deformed so long as the integral converges.

3.5. The classical systems

The theory developed in the preceding sections is of very general applicability. It applies to systems in which quantum-mechanical effects are important as well as to those t..... can be treated classically. In the latter case, our formalism may be written in the language of the phase space; as a result, the summations over quantum states get replaced by integrations over phase space.

We recall the concepts developed in Secs 2.1 and 2.2, especially formula (2.1.3) for the ensemble average $\langle f \rangle$ of a physical quantity $f(q, p)$, namely

$$\langle f \rangle = \frac{\int f(q, p)\rho(q, p)\, d^{3N}q\, d^{3N}p}{\int \rho(q, p)\, d^{3N}q\, d^{3N}p}, \tag{1}$$

where $\rho(q, p)$ denotes the density of the representative points (of the systems) in the phase space; we have omitted here the explicit dependence of the function ρ on time t because we are interested in the study of equilibrium situations only. Evidently, the function $\rho(q, p)$ is a measure of the probability of finding a representative point in the vicinity of the phase point (q, p), which in turn depends upon the corresponding value $H(q, p)$ of the Hamiltonian of the system. In the canonical ensemble,

$$\rho(q, p) \propto \exp\{-\beta H(q, p)\}; \tag{2}$$

cf. eqn. (3.1.6). The expression for $\langle f \rangle$ then takes the form

$$\langle f \rangle = \frac{\int f(q, p) \exp(-\beta H)\, d\omega}{\int \exp(-\beta H)\, d\omega}. \tag{3}$$

where $d\omega\, (\equiv d^{3N}q\, d^{3N}p)$ denotes a volume element of the phase space. The denominator of this expression is directly related to the partition function of the system. However, to write the precise expression for the latter, we must take into account the relationship between a volume element in the phase space and the corresponding number of distinct quantum states of the system. This relationship was established in Secs 2.4 and 2.5, whereby an element of volume $d\omega$ in the phase space corresponds to

$$\frac{d\omega}{N!\, h^{3N}} \tag{4}$$

distinct quantum states of the system.[8] The appropriate expression for the partition function would, therefore, be

$$Q_N(V, T) = \frac{1}{N!\, h^{3N}} \int e^{-\beta H(q, p)}\, d\omega; \tag{5}$$

it is understood that the integration in (5) goes over the *whole* of the phase space.

As our first application of this formulation, we consider the example of an ideal gas. Here, we have a system of N identical molecules, assumed to be monatomic (so that there are no internal degrees of motion to be considered), confined to a space of volume V and in equilibrium at temperature T. Since there are no intermolecular interactions to be taken into account, the energy of the system is wholly kinetic:

$$H(q, p) = \sum_{i=1}^{N} (p_i^2/2m). \tag{6}$$

The partition function would then be

$$Q_N(V,T) = \frac{1}{N!h^{3N}} \int e^{-(\beta/2m)\sum_i p_i^2} \prod_{i=1}^{N} (d^3q_i\, d^3p_i). \tag{7}$$

Integrations over the space coordinates are rather trivial; they yield a factor of V^N. Integrations over the momentum coordinates are also quite easy, once we note that integral (7) is simply a product of N identical integrals. Thus, we get

$$Q_N(V,T) = \frac{V^N}{N!h^{3N}} \left[\int_0^\infty e^{-p^2/2mkT} (4\pi p^2\, dp) \right]^N \tag{8}$$

$$= \frac{1}{N!} \left[\frac{V}{h^3} (2\pi mkT)^{3/2} \right]^N ; \tag{9}$$

here, use has been made of eqn. (B.13a). The Helmholtz free energy is then given by, using Stirling's formula (B.29),

$$A(N,V,T) \equiv -kT \ln Q_N(V,T) = NkT \left[\ln \left\{ \frac{N}{V} \left(\frac{h^2}{2\pi mkT} \right)^{3/2} \right\} - 1 \right]. \tag{10}$$

The foregoing result is identical with eqn. (1.5.8), which was obtained by following a very different procedure. The simplicity of the present approach is, however, striking. Needless to say, the complete thermodynamics of the ideal gas can be derived from eqn. (10) in a straightforward manner. For instance,

$$\mu \equiv \left(\frac{\partial A}{\partial N} \right)_{V,T} = kT \ln \left\{ \frac{N}{V} \left(\frac{h^2}{2\pi mkT} \right)^{3/2} \right\}, \tag{11}$$

$$P \equiv -\left(\frac{\partial A}{\partial V} \right)_{N,T} = \frac{NkT}{V} \tag{12}$$

and

$$S \equiv -\left(\frac{\partial A}{\partial T} \right)_{N,V} = Nk \left[\ln \left\{ \frac{V}{N} \left(\frac{2\pi mkT}{h^2} \right)^{3/2} \right\} + \frac{5}{2} \right] \tag{13}$$

These results are identical with the ones derived previously, namely (1.5.7), (1.4.2) and (1.5.1a), respectively. In fact, the identification of formula (12) with the ideal-gas law, $PV = nRT$, establishes the identity of the (hitherto undetermined) constant k as the *Boltzmann constant*; see eqn. (3.3.6). We further obtain

$$U \equiv -\left[\frac{\partial}{\partial \beta} (\ln Q) \right]_{E_r} \equiv -T^2 \left[\frac{\partial}{\partial T} \left(\frac{A}{T} \right) \right]_{N,V} \equiv A + TS = \frac{3}{2} NkT, \tag{14}$$

and so on.

At this stage we have an important remark to make. Looking at the form of eqn. (8) and the manner in which it came about, we may write

$$Q_N(V,T) = \frac{1}{N!} [Q_1(V,T)]^N , \tag{15}$$

where $Q_1(V, T)$ may be regarded as the partition function of a *single* molecule in the system. A little reflection shows that this result obtains essentially from the fact that the basic constituents of our system are non-interacting (and hence the total energy of the system is simply the sum of their individual energies). Clearly, the situation will not be altered even if the molecules in the system had internal degrees of motion as well. What is essentially required for eqn. (15) to be valid is the absence of interactions among the basic constituents of the system (and, of course, the absence of quantum-mechanical correlations).

Going back to the ideal gas, we could as well have started with the density of states $g(E)$. From eqn. (1.4.17), and in view of the Gibbs correction factor, we have

$$g(E) = \frac{\partial \Sigma}{\partial E} \approx \frac{1}{N!}\left(\frac{V}{h^3}\right)^N \frac{(2\pi m)^{3N/2}}{\{(3N/2)-1\}!} E^{(3N/2)-1}. \tag{16}$$

Substituting this into eqn. (3.4.5), and noting that the integral involved is equal to $\{(3N/2)-1\}!/\beta^{3N/2}$, we readily obtain

$$Q_N(\beta) = \frac{1}{N!}\left(\frac{V}{h^3}\right)^N \left(\frac{2\pi m}{\beta}\right)^{3N/2}. \tag{17}$$

which is identical with (9). It may also be noted that if one starts with the single-particle density of states (2.4.7), namely

$$a(\varepsilon) \approx \frac{2\pi V}{h^3}(2m)^{3/2}\varepsilon^{1/2}. \tag{18}$$

computes the single-particle partition function,

$$Q_1(\beta) = \int_0^\infty e^{-\beta\varepsilon} a(\varepsilon)\, d\varepsilon = \frac{V}{h^3}\left(\frac{2\pi m}{\beta}\right)^{3/2} \tag{19}$$

and then makes use of formula (15), one arrives at the same result for $Q_N(V, T)$.

Lastly, we consider the question of determining the density of states, $g(E)$, from the expression for the partition function, $Q(\beta)$, assuming that the latter is already known; indeed, expression (9) for $Q(\beta)$ was derived without making use of any knowledge regarding the function $g(E)$. According to eqn. (3.4.7) and (9), we have

$$g(E) = \frac{V^N}{N!}\left(\frac{2\pi m}{h^2}\right)^{3N/2} \frac{1}{2\pi i} \int_{\beta'-i\infty}^{\beta'+i\infty} \frac{e^{\beta E}}{\beta^{3N/2}}\, d\beta \quad (\beta' > 0). \tag{20}$$

Noting that, for all positive n,

$$\frac{1}{2\pi i} \int_{s'-i\infty}^{s'+i\infty} \frac{e^{sx}}{s^{n+1}}\, ds = \begin{cases} \dfrac{x^n}{n!} & \text{for} \quad x \geq 0 \\ 0 & \text{for} \quad x \leq 0, \end{cases} \tag{21}$$[9]

eqn. (20) becomes

$$g(E) = \begin{cases} \dfrac{V^N}{N!}\left(\dfrac{2\pi m}{h^2}\right)^{3N/2} \dfrac{E^{(3N/2)-1}}{\{(3N/2)-1\}!} & \text{for} \quad E \geq 0 \\ 0 & \text{for} \quad E \leq 0, \end{cases} \tag{22}$$

which is indeed the correct result for the density of states of an ideal gas; cf. eqn. (16). The foregoing derivation may not appear particularly valuable because in the present case we already knew the expression for $g(E)$. However, cases do arise where the evaluation of the partition function of a given system and the consequent evaluation of its density of states turn out to be quite simple, whereas a direct evaluation of the density of states from first principles happens to be rather involved. In such cases, the method given here can indeed be useful; see, for example, Problem 3.15 in comparison with Problems 1.7 and 2.8.

3.6. Energy fluctuations in the canonical ensemble: correspondence with the microcanonical ensemble

In the canonical ensemble, a system can have energy anywhere between zero and infinity. On the other hand, the energy of a system in the microcanonical ensemble is restricted to a very narrow range. How, then, can we assert that the thermodynamic properties of a system derived through the formalism of the canonical ensemble would be the same as the ones derived through the formalism of the microcanonical ensemble? Of course, we do expect that the two formalisms yield identical results, for otherwise our whole scheme would be marred by inner inconsistency. And, indeed, in the case of an ideal classical gas the results obtained by following one approach were precisely the same as the ones obtained by following the other approach. What is the underlying reason for this equivalence?

The answer to this question is obtained by examining the *actual* extent of the range over which the energies of the systems (in the canonical ensemble) have a significant probability to spread; that will tell us the extent to which the canonical ensemble *really* differs from the microcanonical one. To do this, we write down the expression for the mean energy,

$$U \equiv \langle E \rangle = \frac{\sum\limits_r E_r \exp(-\beta E_r)}{\sum\limits_r \exp(-\beta E_r)}, \tag{1}$$

and differentiate it with respect to the parameter β, holding the energy values E_r constant; we obtain

$$\begin{aligned} \frac{\partial U}{\partial \beta} &= -\frac{\sum\limits_r E_r^2 \exp(-\beta E_r)}{\sum\limits_r \exp(-\beta E_r)} + \frac{\left[\sum\limits_r E_r \exp(-\beta E_r\right]^2}{\left[\sum\limits_r \exp(-\beta E_r)\right]^2} \\ &= -\langle E^2 \rangle + \langle E \rangle^2, \end{aligned} \tag{2}$$

whence it follows that

$$\langle(\Delta E)^2\rangle \equiv \langle E^2\rangle - \langle E\rangle^2 = -\left(\frac{\partial U}{\partial \beta}\right) = kT^2\left(\frac{\partial U}{\partial T}\right) = kT^2 C_V. \tag{3}$$

Note that we have here the specific heat *at constant volume*, because the partial differentiation in (2) was carried out with the E_r kept constant! For the relative root-mean-square fluctuation in E, eqn. (3) gives

$$\frac{\surd[\langle(\Delta E)^2\rangle]}{\langle E\rangle} = \frac{\surd(kT^2 C_V)}{U}, \tag{4}$$

which is $O(N^{-1/2})$, N being the number of particles in the system. Consequently, for large N (which is true for every statistical system) the relative r.m.s. fluctuation in the values of E is quite negligible! Thus, for all practical purposes, a system in the canonical ensemble has an energy equal to, or almost equal to, the mean energy U; the situation in this ensemble is, therefore, practically the same as in the microcanonical ensemble. That explains why the two ensembles lead to practically identical results.

For further understanding of the situation, we consider the manner in which energy is distributed among the various members of the (canonical) ensemble. To do this, we treat E as a continuous variable and start from expression (3.4.3), namely

$$P(E)dE \propto \exp(-\beta E)g(E)dE. \tag{3.4.3}$$

The probability density $P(E)$ is given by the product of two factors: (i) the Boltzmann factor, which monotonically decreases with E, and (ii) the density of states, which monotonically increases with E. The product, therefore, has an extremum at some value of E, say E^*.[10] The value E^* is determined by the condition

$$\frac{\partial}{\partial E}\left\{e^{-\beta E}g(E)\right\}\bigg|_{E=E^*} = 0,$$

that is, by

$$\frac{\partial \ln g(E)}{\partial E}\bigg|_{E=E^*} = \beta. \tag{5}$$

Recalling that

$$S = k\ln g \quad \text{and} \quad \left(\frac{\partial S(E)}{\partial E}\right)_{E=U} = \frac{1}{T} = k\beta.$$

the foregoing condition implies that

$$E^* = U. \tag{6}$$

This is a very interesting result, for it shows that, irrespective of the physical nature of the given system, the most probable value of its energy is identical with its mean value. Accordingly, if it is advantageous, we may use one instead of the other.

We now expand the logarithm of the probability density $P(E)$ around the value $E^* \approx U$; we get

$$\ln\left[e^{-\beta E}g(E)\right] = \left(-\beta U + \frac{S}{k}\right) + \frac{1}{2}\frac{\partial^2}{\partial E^2}\ln\left\{e^{-\beta E}g(E)\right\}\Bigg|_{E=U}(E-U)^2 +$$

$$= -\beta(U - TS) - \frac{1}{2kT^2C_V}(E-U)^2 + \cdots, \tag{7}$$

whence

$$P(E) \propto e^{-\beta E}g(E) \simeq e^{-\beta(U-TS)}\exp\left\{-\frac{(E-U)^2}{2kT^2C_V}\right\} \tag{8}$$

This is a *Gaussian* distribution in E, with mean value U and dispersion $\sqrt{(kT^2C_V)}$; cf. (3). In terms of the reduced variable E/U, the distribution is again Gaussian, with mean value unity and dispersion $\sqrt{(kT^2C_V)}/U$ {which is $O(N^{-1/2})$}; thus, for $N \gg 1$, we have an extremely sharp distribution which, as $N \to \infty$, approaches a delta-function!

It would be instructive here to consider once again the case of a classical ideal gas. Here, $g(E)$ is proportional to $E^{(3N/2-1)}$ and hence increases very fast with E; the factor $e^{-\beta E}$, of course, decreases with E. The product $g(E)\exp(-\beta E)$ exhibits a maximum at $E^* = (3N/2 - 1)\beta^{-1}$, which is practically the same as the mean value $U = (3N/2)\beta^{-1}$. For values of E significantly different from E^*, the product essentially vanishes (for smaller values of E, due to the relative paucity of the available energy states; for larger values of E, due to the relative depletion caused by the Boltzmann factor). The overall picture is shown in Fig. 3.3 where we have depicted the actual behavior of these functions in the special case: $N = 10$. The most probable value of E is now $\frac{14}{15}$ of the mean value; so, the distribution is somewhat asymmetrical. The effective width Δ can be readily calculated from (3) and turns out to be $(2/3N)^{1/2}U$ which, for $N = 10$, is about a quarter of U. We can see that, as N becomes large, both E^* and U increase (essentially linearly with N), the ratio E^*/U approaches unity and the distribution tends to become symmetrical about E^*. At the same time, the width Δ increases (but only as $N^{1/2}$); considered in the relative sense, it tends to vanish (as $N^{-1/2}$).

Let us finally look at the partition function $Q_N(V,T)$, as given by eqn. (3.4.5), with its integrand replaced by (8). We have

$$Q_N(V,T) \simeq e^{-\beta(U-TS)}\int_0^\infty e^{-(E-U)^2/2kT^2C_V}dE$$

$$\simeq e^{-\beta(U-TS)}\sqrt{(2kT^2C_V)}\int_{-\infty}^{\infty} e^{-x^2}\,dx$$

$$= e^{-\beta(U-TS)}\sqrt{(2\pi kT^2C_V)}.$$

so that

$$-kT\ln Q_N(V,T) \equiv A \simeq (U - TS) - \tfrac{1}{2}kT\ln(2\pi kT^2C_V). \tag{9}$$

The last term, being $O(\ln N)$, is negligible in comparison with the other terms, which are all $O(N)$. Hence,

$$A \approx U - TS. \tag{10}$$

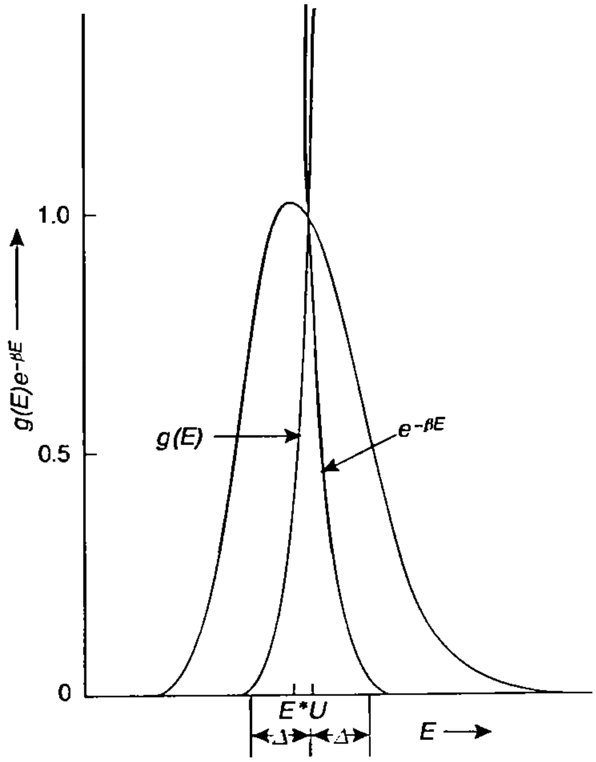


FIG. 3.3. The actual behavior of the functions $g(E)$, $e^{-\beta E}$ and $g(E)e^{-\beta E}$ for an ideal gas, with $N = 10$. The numerical values of the functions have been expressed as fractions of their respective values at $E = U$.

Note that the quantity A in this formula has come through the formalism of the canonical ensemble, while the quantity S has come through a definition belonging to the microcanonical ensemble. The fact that we finally end up with a consistent thermodynamic relationship establishes beyond doubt that the two approaches are, for all practical purposes, identical.

3.7. Two theorems — the "equipartition" and the "virial"

To arrive at these theorems, we determine the expectation value of the quantity $x_i(\partial H/\partial x_j)$, where $H(q, p)$ is the Hamiltonian of the system (assumed classical) while x_i and x_j are any of the $6N$ generalized coordinates (q, p). In the canonical ensemble,

$$\left\langle x_i \frac{\partial H}{\partial x_j} \right\rangle = \frac{\int \left(x_i \dfrac{\partial H}{\partial x_j} \right) e^{-\beta H}\, d\omega}{\int e^{-\beta H}\, d\omega} \qquad (d\omega = d^{3N}q\, d^{3N}p). \tag{1}$$

Let us consider the integral in the numerator. Integrating over x_j by parts, it becomes

$$\int \left[-\frac{1}{\beta} x_i e^{-\beta H} \Big|_{(x_j)_1}^{(x_j)_2} + \frac{1}{\beta} \int \left(\frac{\partial x_i}{\partial x_j} \right) e^{-\beta H} dx_j \right] d\omega_{(j)};$$

here, $(x_j)_1$ and $(x_j)_2$ are the "extreme" values of the coordinate x_j, while $d\omega_{(j)}$ denotes "$d\omega$ devoid of dx_j". The integrated part vanishes because whenever any of the coordinates takes an "extreme" value the Hamiltonian of the system becomes infinite.[11] In the integral that remains, the factor $\partial x_i/\partial x_j$, being equal to δ_{ij}, comes out of the integral sign and we are left with

$$\frac{1}{\beta}\delta_{ij}\int e^{-\beta H}\,d\omega.$$

Substituting this into (1), we arrive at the desired result:

$$\left\langle x_i \frac{\partial H}{\partial x_j}\right\rangle = \delta_{ij}kT. \tag{2}$$

which is independent of the precise form of the function H

In the special case $x_i = x_j = p_i$, eqn. (2) takes the form

$$\left\langle p_i \frac{\partial H}{\partial p_i}\right\rangle \equiv \langle p_i \dot{q}_i\rangle = kT, \tag{3}$$

while for $x_i = x_j = q_i$, it becomes

$$\left\langle q_i \frac{\partial H}{\partial q_i}\right\rangle \equiv -\langle q_i \dot{p}_i\rangle = kT. \tag{4}$$

Adding over all i, from $i = 1$ to $3N$, we obtain

$$\left\langle \sum_i p_i \frac{\partial H}{\partial p_i}\right\rangle \equiv \left\langle \sum_i p_i \dot{q}_i\right\rangle = 3NkT \tag{5}$$

and

$$\left\langle \sum_i q_i \frac{\partial H}{\partial q_i}\right\rangle \equiv -\left\langle \sum_i q_i \dot{p}_i\right\rangle = 3NkT. \tag{6}$$

Now, in several physical situations the Hamiltonian of the system happens to be a *quadratic* function of its coordinates; so, through a canonical transformation, it can be brought into the form

$$H = \sum_j A_j P_j^2 + \sum_j B_j Q_j^2, \tag{7}$$

where P_j and Q_j are the transformed, canonically conjugate, coordinates while A_j and B_j are certain constants of the problem. For such a system, we clearly have

$$\sum_j \left(P_j \frac{\partial H}{\partial P_j} + Q_j \frac{\partial H}{\partial Q_j}\right) = 2H; \tag{8}$$

accordingly, by eqns (3) and (4),

$$\langle H\rangle = \tfrac{1}{2} fkT, \tag{9}$$

where f is the number of nonvanishing coefficients in the expression (7). We therefore conclude that each harmonic term in the (transformed) Hamiltonian makes a contribution of $\frac{1}{2}kT$ towards the internal energy of the system and, hence, a contribution of $\frac{1}{2}k$ towards the specific heat C_V. This result embodies the classical theorem of *equipartition of energy* (among the various degrees of freedom of the system). It [illegible] be mentioned here that, for the distribution of kinetic energy alone, the equipartition theorem was first stated by Boltzmann (1871).

In our subsequent study we shall find that the equipartition theorem as stated here is not always valid; it applies only when the relevant degrees of freedom can be *freely* excited. At a given temperature T, there may be certain degrees of freedom which, due to the insufficiency of the energy available, are more or less "frozen". Such degrees of freedom do not make a significant contribution towards the internal energy of the system or towards its specific heat; see, for example, Secs 6.5, 7.3 and 8.3. Of course, the higher the temperature of the system the better the validity of this theorem.

We now consider the implications of formula (6). First of all, we note that this formula embodies the so-called *virial theorem* of Clausius (1870) for the quantity $\langle \sum_i q_i \dot{p}_i \rangle$, which is the expectation value of the sum of the products of the coordinates of the various particles in the system and the respective forces acting on them; this quantity is referred to as the *virial* of the system and is generally denoted by the symbol $\mathcal{V}$. The virial theorem then states that

$$\mathcal{V} = -3NkT. \qquad (10)$$

The relationship between the virial and other physical quantities of the system is best understood by first looking at a classical gas of non-interacting particles. In this case, the only forces that come into play are the ones arising from the walls of the container; these forces can be designated by an external pressure P that acts upon the system by virtue of the fact that it is bounded by the walls of the container. Consequently, we have here a force $-P\,d\mathbf{S}$ associated with an element of area $d\mathbf{S}$ of the walls; the negative sign appears because the force is directed *inward* while the vector $d\mathbf{S}$ is directed outward. The virial of the gas is then given by

$$\mathcal{V}_0 = \left\langle \sum_i q_i F_i \right\rangle_0 = -P \oint_S \mathbf{r} \cdot d\mathbf{S}, \qquad (11)$$[12]

where $\mathbf{r}$ is the position vector of a particle that happens to be in the (close) vicinity of the surface element $d\mathbf{S}$; accordingly, $\mathbf{r}$ may be considered to be the position vector of the surface element itself. By the divergence theorem, eqn. (11) becomes

$$\mathcal{V}_0 = -P \int_V (\operatorname{div} \mathbf{r})\, dV = -3PV. \qquad (12)$$

Comparing (12) with (10), we obtain the well-known result:

$$PV = NkT. \qquad (13)$$

The internal energy of the gas, which in this case is wholly kinetic, follows from the equipartition theorem (9) and is equal to $\frac{3}{2}NkT$, $3N$ being the number of degrees of freedom. Comparing this result with (10), we obtain the classical relationship

$$\mathcal{V} = -2K, \qquad (14)$$

where K denotes the average kinetic energy of the system.

It is straightforward to carry out an extension of this study to a system of particles interacting through a two-body potential $u(\mathbf{r}_j - \mathbf{r}_i)$. The virial $\mathcal{V}$ then draws a contribution from the interior as well. Assuming the interparticle potential to be central and denoting it by the symbol $u(r)$, where $r = |r_j - r_i|$, the contribution arising from the pair of particles i and j, with position vectors r_i and $\mathbf{r}_j$, is given by

$$r_i \cdot \left(-\frac{\partial u(r)}{\partial r_i}\right) + r_j \cdot \left(-\frac{\partial u(r)}{\partial r_j}\right)$$

$$= -\frac{\partial u(r)}{\partial r^2}\left\{r_i \cdot \frac{\partial |r_j - r_i|^2}{\partial r_i} + r_j \cdot \frac{\partial |r_j - r_i|^2}{\partial r_j}\right\} = -r\frac{\partial u(r)}{\partial r} \tag{15}$$

The net contribution, arising from all the $N(N-1)/2$ pairs of particles, would then be, for $N \gg 1$,

$$\frac{1}{2}N^2\left\langle -r\frac{\partial u(r)}{\partial r}\right\rangle = -\frac{N^2}{2}\iint\left\{r\frac{\partial u(r)}{\partial r}\right\}g(r_2 - r_1)\frac{dr_1\,dr_2}{V^2}$$

$$= -\frac{N^2}{2V}\int_0^\infty \left\{r\frac{\partial u(r)}{\partial r}\right\}g(r)(4\pi r^2\,dr), \tag{16}$$

where $g(r)$, the *pair distribution function* of the system, is a measure of the probability of finding a pair of particles in it separated by a distance r; as used here, $g(r) \to 1$ as $r \to \infty$. Combining (16) with (12) and equating the resulting sum with (10), we obtain for a classical interacting system

$$PV = NkT\left[1 - \frac{2\pi n}{3kT}\int_0^\infty \frac{\partial u(r)}{\partial r}g(r)r^3\,dr\right] \tag{17}$$

where n is the particle density in the system. The internal energy of the system can also be expressed in terms of the functions $u(r)$ and $g(r)$. Noting that the average kinetic energy is still given by the expression $\frac{3}{2}NkT$, we have for the total energy

$$U = \frac{3}{2}NkT + \frac{1}{2}N^2\iint u(r)g(r)\frac{dr_1\,dr_2}{V^2}$$

$$= \frac{3}{2}NkT\left[1 + \frac{4\pi n}{3kT}\int_0^\infty u(r)g(r)r^2\,dr\right] \tag{18}$$

the second term being the average potential energy of the system. Clearly, a knowledge of the functions $u(r)$ and $g(r)$, the latter itself depending on the nature of the former, is essential before we can make use of eqns (17) and (18). For further study along these lines, see Hill (1956), Chap. 6; see also Problem 3.23.

3.8. A system of harmonic oscillators

We shall now examine a system of N, practically independent, harmonic oscillators. This study will not only provide an interesting illustration of the canonical ensemble formulation but will also serve as a basis for some of our subsequent studies in this text. Two important problems in this line are (i) the theory of the black-body radiation (or the "statistical mechanics of photons") and (ii) the theory of lattice vibrations (or the "statistical mechanics of phonons"); see Secs 7.2 and 7.3 for details.

We start with the specialized situation when the oscillators can be treated *classically*. The Hamiltonian of any one of them (assumed to be one-dimensional) is then given by

$$H(q_i, p_i) = \frac{1}{2}m\omega^2 q_i^2 + \frac{1}{2m}p_i^2 \quad (i = 1, \ldots, N). \tag{1}$$

For the single-oscillator partition function, we readily obtain

$$\begin{aligned} Q_1(\beta) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left\{-\beta\left(\frac{1}{2}m\omega^2 q^2 + \frac{1}{2m}p^2\right)\right\}\frac{dq\,dp}{h} \\ &= \frac{1}{h}\left(\frac{2\pi}{\beta m\omega^2}\right)^{1/2}\left(\frac{2\pi m}{\beta}\right)^{1/2} = \frac{1}{\beta\hbar\omega}, \end{aligned} \tag{2}$$

where $\hbar = h/2\pi$. The partition function of the N-oscillator system would then be

$$Q_N(\beta) = [Q_1(\beta)]^N = (\beta\hbar\omega)^{-N}; \tag{3}$$

note that in writing (3) we have assumed the oscillators to be *distinguishable*. This is so because, as we shall see later, these oscillators are merely a representation of the energy levels available in the system; they are not particles (or even "quasiparticles"). It is actually photons in one case and phonons in the other, which distribute themselves over the various oscillator levels, that are *indistinguishable*!

The Helmholtz free energy of the system is now given by

$$A \equiv -kT\ln Q_N = NkT\ln\left(\frac{\hbar\omega}{kT}\right), \tag{4}$$

whence

$$\mu = kT\ln\left(\frac{\hbar\omega}{kT}\right) \tag{5}$$

$$P = 0. \tag{6}$$

$$S = Nk\left[\ln\left(\frac{kT}{\hbar\omega}\right) + 1\right], \tag{7}$$

$$U = NkT \tag{8}$$

and

$$C_P = C_V = Nk. \tag{9}$$

We note that the mean energy per oscillator is in complete agreement with the equipartition theorem, namely $2 \times \frac{1}{2}kT$, for we have here *two* independent quadratic terms in the single-oscillator Hamiltonian.

We may determine the density of states, $g(E)$, of this system from the expression (3) for its partition function. We have, in view of (3.4.7),

$$g(E) = \frac{1}{(\hbar\omega)^N} \frac{1}{2\pi i} \int\limits_{\beta'-i\infty}^{\beta'+i\infty} \frac{e^{\beta E}}{\beta^N} d\beta \quad (\beta' > 0),$$

that is,

$$g(E) = \begin{cases} \dfrac{1}{(\hbar\omega)^N} \dfrac{E^{N-1}}{(N-1)!} & \text{for} \quad E \geq 0 \\ 0 & \text{for} \quad E \leq 0. \end{cases} \tag{10}$$

To test its correctness, we may calculate the entropy of the system with the help of this formula. Taking $N \gg 1$ and making use of the Stirling approximation, we get

$$S(N, E) = k \ln g(E) \approx Nk \left[\ln \left(\frac{E}{N\hbar\omega} \right) + 1 \right], \tag{11}$$

which gives for the temperature of the system

$$T = \left(\frac{\partial S}{\partial E} \right)_N^{-1} = \frac{E}{Nk}. \tag{12}$$

Eliminating E between these two relations, we obtain precisely our earlier result (7) for the function $S(N, T)$.

We now take up the quantum-mechanical situation, according to which the energy eigenvalues of a one-dimensional harmonic oscillator are given by

$$\varepsilon_n = \left(n + \tfrac{1}{2}\right) \hbar\omega; \quad n = 0, 1, 2, . \tag{13}$$

Accordingly, we have for the single-oscillator partition function

$$\begin{aligned} Q_1(\beta) &= \sum_{n=0}^{\infty} e^{-\beta(n+1/2)\hbar\omega} = \frac{\exp\left(-\frac{1}{2}\beta\hbar\omega\right)}{1 - \exp(-\beta\hbar\omega)} \\ &= \left\{2 \sinh\left(\tfrac{1}{2}\beta\hbar\omega\right)\right\}^{-1} \end{aligned} \tag{14}$$

The N-oscillator partition function is then given by

$$\begin{aligned} Q_N(\beta) &= [Q_1(\beta)]^N = \left[2 \sinh\left(\tfrac{1}{2}\beta\hbar\omega\right)\right]^{-N} \\ &= e^{-(N/2)\beta\hbar\omega} \left\{1 - e^{-\beta\hbar\omega}\right\}^{-N} \end{aligned} \tag{15}$$

For the Helmholtz free energy of the system, we get

$$A = NkT \ln \left[2 \sinh\left(\tfrac{1}{2}\beta\hbar\omega\right)\right] = N \left[\tfrac{1}{2}\hbar\omega + kT \ln\left\{1 - e^{-\beta\hbar\omega}\right\}\right], \tag{16}$$

whence

$$\mu = A/N, \tag{17}$$

$$P = 0 \tag{18}$$

$$S = Nk \left[\tfrac{1}{2}\beta\hbar\omega \coth\left(\tfrac{1}{2}\beta\hbar\omega\right) - \ln\left\{2 \sinh\left(\tfrac{1}{2}\beta\hbar\omega\right)\right\}\right]$$

$$= Nk\left[\frac{\beta\hbar\omega}{e^{\beta\hbar\omega}-1} - \ln\left\{1 - e^{-\beta\hbar\omega}\right\}\right], \tag{19}$$

$$U = \frac{1}{2}N\hbar\omega\coth\left(\frac{1}{2}\beta\hbar\omega\right) = N\left[\frac{1}{2}\hbar\omega + \frac{\hbar\omega}{e^{\beta\hbar\omega}-1}\right] \tag{20}$$

and

$$C_P = C_V = Nk\left(\tfrac{1}{2}\beta\hbar\omega\right)^2 \operatorname{cosech}^2\left(\tfrac{1}{2}\beta\hbar\omega\right)$$

$$= Nk(\beta\hbar\omega)^2\frac{e^{\beta\hbar\omega}}{\left(e^{\beta\hbar\omega}-1\right)^2}. \tag{21}$$

Formula (20) is especially significant, for it shows that the quantum-mechanical oscillators do not obey the equipartition theorem. The mean energy per oscillator is different from the equipartition value kT; actually, it is always greater than kT; see curve 2 in Fig. 3.4. Only in the limit of high temperatures, where the thermal energy kT is much larger than the energy quantum $\hbar\omega$, does the mean energy per oscillator tend to the equipartition value. It should be noted here that if the zero-point energy $\frac{1}{2}\hbar\omega$ were not present, the limiting value of the mean energy would be $(kT - \frac{1}{2}\hbar\omega)$, and not kT–we may call such an oscillator the *Planck oscillator*; see curve 1 in the figure. In passing, we observe that the specific heat (21), which is the same for the Planck oscillator as for the Schrödinger oscillator, is temperature-dependent; moreover, it is always less than, and at high temperatures tends to, the classical value (9).

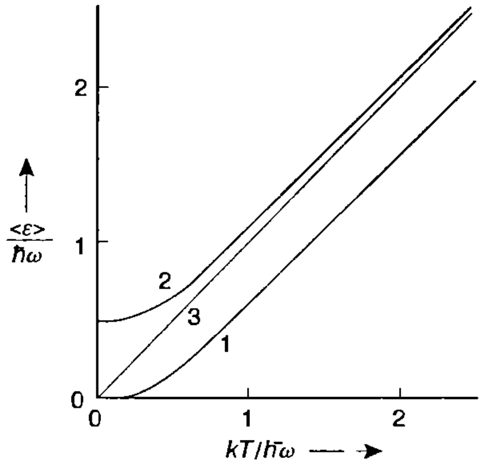


FIG. 3.4. The mean energy $\langle\varepsilon\rangle$ of a simple harmonic oscillator as a function of temperature. 1 — the Planck oscillator; 2 — the Schrödinger oscillator; 3 — the classical oscillator.

Indeed. for $kT \gg \hbar\omega$, formulae (14)–(21) go over to their classical counterparts, namely (2)–(9), respectively.

We shall now determine the density of states $g(E)$ of the N-oscillator system from its partition function (15). Carrying out the binomial expansion of this expression, we have

$$Q_N(\beta) = \sum_{R=0}^{\infty} \binom{N+R-1}{R} e^{-\beta(\frac{1}{2}N\hbar\omega + R\hbar\omega)} \tag{22}$$

Comparing this with the formula

$$Q_N(\beta) = \int_0^{\infty} g(E) e^{-\beta E}\, dE,$$

we conclude that

$$g(E) = \sum_{R=0}^{\infty} \binom{N+R-1}{R} \delta\left(E - \left\{R + \tfrac{1}{2}N\right\}\hbar\omega\right), \tag{23}$$

where $\delta(x)$ denotes the Dirac delta function. Equation (23) implies that there are $(N+R-1)!/R!(N-1)!$ microstates available to the system when its energy E has the discrete value $(R + \frac{1}{2}N)\hbar\omega$, where $R = 0, 1, 2, \ldots$, and that no microstate is available for other values of E. This is hardly surprising, but it is instructive to look at this result from a slightly different point of view.

We consider the following problem which arises naturally in the microcanonical ensemble theory. Given an energy E for distribution among a set of N harmonic oscillators, each of which can be in any one of the eigenstates (13), what is the total number of *distinct* ways in which the process of distribution can be carried out? Now, in view of the form of the eigenvalues ε_n, it makes sense to give away, right in the beginning, the zero-point energy $\frac{1}{2}\hbar\omega$ to each of the N oscillators and convert the rest of it into quanta (of energy $\hbar\omega$). Let R be the number of these quanta; then

$$R = \left(E - \tfrac{1}{2}N\hbar\omega\right)/\hbar\omega. \tag{24}$$

Clearly, R must be an integer; by implication, E must be of the form $(R + \frac{1}{2}N)\hbar\omega$. The problem then reduces to determining the number of *distinct* ways of allotting R quanta to N oscillators, such that an oscillator may have 0 or 1 or 2 ... quanta; in other words, we have to determine the number of *distinct* ways of putting R *indistinguishable* balls into N *distinguishable* boxes, such that a box may receive 0 or 1 or 2 ... balls. A little reflection will show that this is precisely the number of permutations that can be realized by shuffling R balls, placed along a row, with $(N-1)$ partitioning lines (that divide the given space into N boxes); see Fig. 3.5. The answer clearly is

$$\frac{(R+N-1)!}{R!(N-1)!}, \tag{25}$$

which agrees with (23).

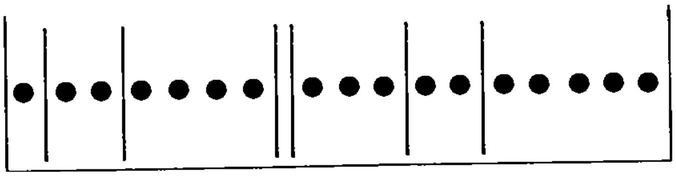


FIG. 3.5. Distributing 17 *indistinguishable* balls among *tinguishable* boxes. The arrangement shown here represents one of the 23!/17!6! inct ways of carrying out the distribution.

We can determine the entropy of the system from the number (25). Since $N \gg 1$, we have

$$S \approx k\{\ln (R+N)! - \ln R! - \ln N!\}$$
$$\approx k\{(R+N)\ln (R+N) - R\ln R - N\ln N\}; \tag{26}$$

the number R is, of course, a measure of the energy of the system, see (24). For the temperature of the system, we obtain

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_N = \left(\frac{\partial S}{\partial R}\right)_N \frac{1}{\hbar\omega} = \frac{k}{\hbar\omega} \ln\left(\frac{R+N}{R}\right) = \frac{k}{\hbar\omega} \ln\left(\frac{E+\frac{1}{2}N\hbar\omega}{E-\frac{1}{2}N\hbar\omega}\right) \tag{27}$$

whence

$$\frac{E}{N} = \frac{1}{2}\hbar\omega \frac{\exp(\hbar\omega/kT)+1}{\exp(\hbar\omega/kT)-1}, \tag{28}$$

which is identical with (20). It can be further checked that, by eliminating R between (26) and (27), we obtain precisely the formula (19) for $S(N, T)$. Thus, once again, we find that the results obtained by following the microcanonical approach and the canonical approach are asymptotically the same.

Finally, we may consider the classical limit when E/N, the mean energy per oscillator, is much larger than the energy quantum $\hbar\omega$, i.e. when $R \gg N$. The expression (25) may, in that case, be replaced by

$$\frac{(R+N-1)(R+N-2)\dots(R+1)}{(N-1)!} \approx \frac{R^{N-1}}{(N-1)!}, \tag{25a}$$

with

$$R \approx E/\hbar\omega.$$

The corresponding expression for the entropy turns out to be

$$S \approx k\{N\ln(R/N) + N\} \approx Nk\left\{\ln\left(\frac{E}{N\hbar\omega}\right) + 1\right\} \tag{26a}$$

whence

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_N \approx \frac{Nk}{E}, \tag{27a}$$

so that

$$\frac{E}{N} \approx kT. \tag{28a}$$

These results are identical with the ones derived earlier in the section.

3.9. The statistics of paramagnetism

Next, we study a system of N magnetic dipoles, each having a magnetic moment μ. In the presence of an external magnetic field H, the dipoles will experience a torque tending to align them in the direction of the field. If there were nothing else to check this tendency, the dipoles would align themselves precisely in this direction and we would achieve a complete magnetization of the system. In reality, however, thermal agitation in the system offers resistance to this tendency and, in equilibrium, we obtain only a *partial* magnetization. Clearly, as $T \to 0$ K, the thermal agitation becomes ineffective and the system exhibits a complete orientation of the dipole moments, whatever the strength of the applied field; at the other extreme, as $T \to \infty$, we approach a state of complete randomization of the dipole moments, which implies a vanishing magnetization. At intermediate temperatures, the situation is governed by the parameter $(\mu H/kT)$.

The model adopted for this study consists of N identical, localized (and, hence, distinguishable), practically static, mutually non-interacting and freely orientable dipoles. It is obvious that the only energy we need to consider here is the potential energy of the dipoles which arises from the presence of the external field $\boldsymbol{H}$ and is determined by the orientations of the dipoles with respect to the direction of the field:

$$E = \sum_{i=1}^{N} E_i = -\sum_{i=1}^{N} \boldsymbol{\mu}_i \cdot \boldsymbol{H} = -\mu H \sum_{i=1}^{N} \cos\theta_i. \tag{1}$$

The partition function of the system is then given by

$$Q_N(\beta) = [Q_1(\beta)]^N. \tag{2}$$

where

$$Q_1(\beta) = \sum_{\theta} \exp(\beta\mu H \cos\theta). \tag{3}$$

The mean magnetic moment M of the system will obviously be in the direction of the field $\boldsymbol{H}$; for its magnitude we shall have

$$M_z = N\langle\mu\cos\theta\rangle = N\frac{\sum_{\theta} \mu\cos\theta \exp(\beta\mu H\cos\theta)}{\sum_{\theta} \exp(\beta\mu H\cos\theta)}$$

$$= \frac{N}{\beta}\frac{\partial}{\partial H} \ln Q_1(\beta). \tag{4}$$

Thus, to determine the degree of magnetization in the system all we have to do is to evaluate the single-dipole partition function (3).

First, we do it classically (after Langevin, 1905a,b). Using $(\sin\theta \, d\theta \, d\phi)$ as the elemental solid angle representing a small range of orientations of the dipole, we get

$$Q_1(\beta) = \int_0^{2\pi}\int_0^{\pi} e^{\beta\mu H\cos\theta} \sin\theta \, d\theta \, d\phi = 4\pi\frac{\sinh(\beta\mu H)}{\beta\mu H}, \tag{5}$$

whence

$$\overline{\mu}_z \equiv \frac{M_z}{N} = \mu\left\{\coth(\beta\mu H) - \frac{1}{\beta\mu H}\right\} = \mu L(\beta\mu H), \tag{6}$$

where $L(x)$ is the so-called *Langevin function*

$$L(x) = \coth x - \frac{1}{x}; \tag{7}$$

we note that the parameter $\beta\mu H$ denotes the strength of the (magnetic) potential energy μH against the (thermal) kinetic energy kT. A plot of the Langevin function is shown in Fig. 3.6.

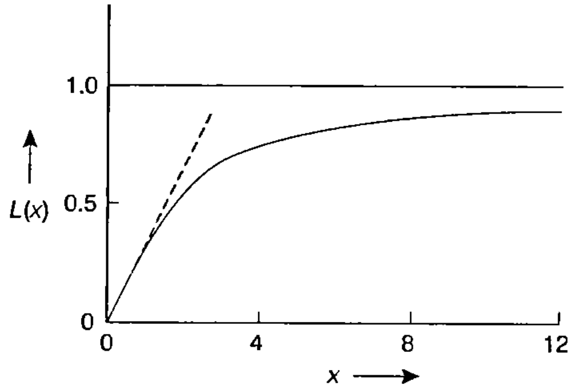


FIG. 3.6. The Langevin function $L(x)$.

If we have N_0 dipoles per unit volume in the system, then the magnetization of the system, viz. the mean magnetic moment per unit volume, is given by

$$M_{z0} = N_0\overline{\mu}_z = N_0\mu L(x) \quad (x = \beta\mu H). \tag{8}$$

For magnetic fields so strong (or temperatures so low) that the parameter $x \gg 1$, the function $L(x)$ is almost equal to 1; the system then acquires a state of magnetic saturation:

$$\overline{\mu}_z \simeq \mu \quad \text{and} \quad M_{z0} \simeq N_0\mu. \tag{9}$$

For temperatures so high (or magnetic fields so weak) that the parameter $x \ll 1$, the function $L(x)$ may be written as

$$\frac{x}{3} - \frac{x^3}{45} + \cdots \tag{10}$$

which, in the lowest approximation, gives

$$M_{z0} \simeq \frac{N_0\mu^2}{3kT}H. \tag{11}$$

The high-temperature susceptibility of the system is, therefore, given by

$$\chi = \operatorname*{Lim}_{H\to 0}\left(\frac{\partial M_{z0}}{\partial H}\right) \simeq \frac{N_0\mu^2}{3kT} = \frac{C}{T}, \text{ say.} \tag{12}$$

Equation (12) is the *Curie law* of paramagnetism, the parameter C being the *Curie constant* of the system. Figure 3.7 shows a plot of the susceptibility of a powdered sample of copper–potassium sulphate hexahydrate as a function of T^{-1}; the fact that the plot is linear and passes almost through the origin vindicates the Curie law for this particular salt.

We shall now treat the problem of paramagnetism quantum-mechanically. The major modification here arises from the fact that the magnetic dipole moment μ and its component μ_z in the direction of the applied field cannot have *arbitrary* values.

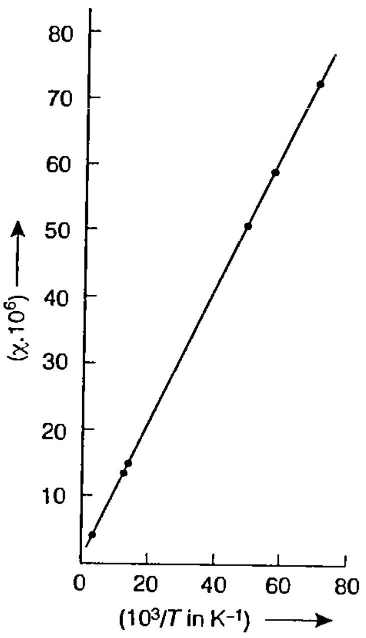


FIG. 3.7. χ vs. $1/T$ plot for a powdered sample of copper–potassium sulphate hexahydrate (after Hupse, 1942).

Quite generally, we have a direct relationship between the magnetic moment μ of a given dipole and its angular momentum l:

$$\mu = \left(g\frac{e}{2mc}\right) l, \tag{13}$$

with

$$l^2 = J(J+1)\hbar^2; \quad J = \tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \quad \text{or} \quad 0, 1, 2, \tag{14}$$

The quantity $g(e/2mc)$ is referred to as the *gyromagnetic ratio* of the dipole while the number g is known as *Lande's g-factor*. If the net angular momentum of the dipole is due solely to electron spins, then $g = 2$; on the other hand, if it is due solely to orbital motions, then $g = 1$. In general, however, its origin is mixed; g is then given by the formula

$$g = \frac{3}{2} + \frac{S(S+1) - L(L+1)}{2J(J+1)}, \tag{15}$$

S and L being, respectively, the spin and the orbital quantum numbers of the dipole. Note that there is no upper or lower bound for the values that g can have!

Combining (13) and (14), we can write

$$\mu^2 = g^2\mu_B^2 J(J+1), \tag{16}$$

where $\mu_B(= e\hbar/2mc)$ is the *Bohr magneton*. The component μ_z of the magnetic moment in the direction of the applied field is, on the other hand, given by

$$\mu_z = g\mu_B m, \quad m = -J, -J+1, \ldots, J-1, J. \tag{17}$$

Thus, a dipole whose magnetic moment μ conforms to expression (16) can have no other orientations with respect to the applied field except the ones conforming to the values (17) of the component μ_z; obviously, the number of allowed orientations, for a given value of J, is $(2J+1)$. In view of this, the single-dipole partition function $Q_1(\beta)$ is now given by, see (3),

$$Q_1(\beta) = \sum_{m=-J}^{J} \exp(\beta g \mu_B m H). \tag{18}$$

Introducing a parameter x, defined by

$$x = \beta(g\mu_B J)H, \tag{19}$$

eqn. (18) becomes

$$\begin{aligned} Q_1(\beta) &= \sum_{m=-J}^{J} e^{mx/J} = \frac{e^{-x}\{e^{(2J+1)x/J} - 1\}}{e^{x/J} - 1} \\ &= \frac{e^{(2J+1)x/2J} - e^{-(2J+1)x/2J}}{e^{x/2J} - e^{-x/2J}} \\ &= \sinh\left\{\left(1 + \frac{1}{2J}\right)x\right\} \Big/ \sinh\left\{\frac{1}{2J}x\right\}. \end{aligned} \tag{20}$$

The mean magnetic moment of the system is then given by, see (4),

$$\begin{aligned} M_z = N\bar{\mu}_z &= \frac{N}{\beta}\frac{\partial}{\partial H}\ln Q_1(\beta) \\ &= N(g\mu_B J)\left[\left(1 + \frac{1}{2J}\right)\coth\left\{\left(1 + \frac{1}{2J}\right)x\right\} - \frac{1}{2J}\coth\left\{\frac{1}{2J}x\right\}\right]. \end{aligned} \tag{21}$$

Thus

$$\bar{\mu}_z = (g\mu_B J)B_J(x), \tag{22}$$

where $B_J(x)$ is the *Brillouin function* (of order J):

$$B_J(x) = \left(1 + \frac{1}{2J}\right)\coth\left\{\left(1 + \frac{1}{2J}\right)x\right\} - \frac{1}{2J}\coth\left\{\frac{1}{2J}x\right\}. \tag{23}$$

In Fig. 3.8 we have plotted the function $B_J(x)$ for some typical values of the quantum number J.

We shall now consider a few special cases. First of all, we note that for strong fields and low temperatures ($x \gg 1$), the function $B_J(x) \simeq 1$ *for all* J, which corresponds to a state of magnetic saturation. On the other hand, for high temperatures and weak fields ($x \ll 1$), the function $B_J(x)$ may be written as

$$\tfrac{1}{3}(1 + 1/J)x + \tag{24}$$

so that

$$\bar{\mu}_z \simeq \frac{(g\mu_B J)^2}{3kT}\left(1 + \frac{1}{J}\right)H = \frac{g^2\mu_B^2 J(J+1)}{3kT}H. \tag{25}$$

The Curie law, $\chi \propto 1/T$, is again obeyed; however, the Curie constant is now given by

$$C_J = \frac{N_0 g^2 \mu_B^2 J(J+1)}{3k} = \frac{N_0 \mu^2}{3k}; \tag{26}$$

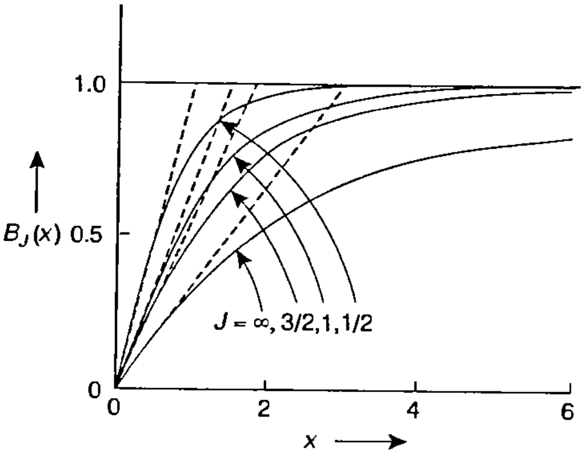


FIG. 3.8. The Brillouin function $B_J(x)$ for various values of J.

see (16). It is indeed interesting that the high-temperature results, (25) and (26), directly involve the eigenvalues of the operator μ^2.

Let us now look a little more closely at the dependence of the foregoing results on the quantum number J. First of all, we consider the extreme case $J \to \infty$, with the understanding that simultaneously $g \to 0$ such that the value of μ stays constant. From (23), we readily observe that, in this limit, the Brillouin function $B_J(x)$ tends to become (i) independent of J and (ii) identical with the Langevin function $L(x)$. This is not surprising because, in this limit, the number of allowed orientations for a magnetic dipole becomes infinitely large, with the result that the problem essentially reduces to its classical counterpart (where one allows all

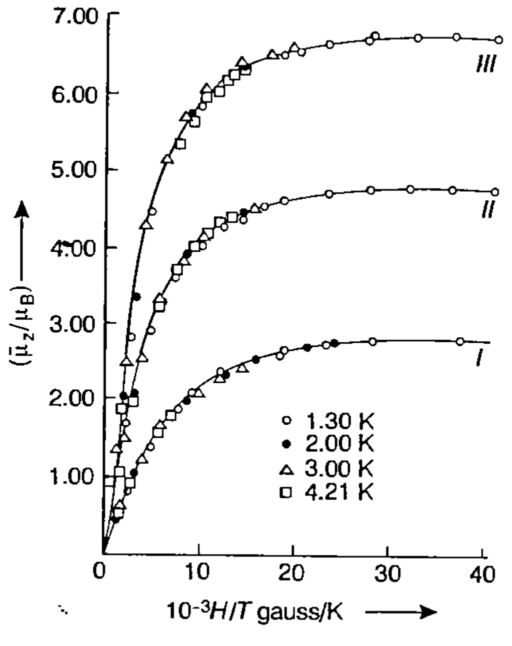


FIG. 3.9. Plots of $\bar{\mu}_z/\mu_B$ as a function of H/T. The solid curves represent the theoretical results, while the points mark the experimental findings of Henry (1952). Curve I is for potassium chromium alum ($J = \frac{3}{2}, g = 2$), curve II for iron ammonia alum ($J = \frac{5}{2}, g = 2$) and curve III for gadolinium sulphate octahydrate ($J = \frac{7}{2}, g = 2$).

conceivable orientations). At the other extreme, we have the case $J = \frac{1}{2}$, which allows only two orientations. The results in this case are very different from the ones for $J \gg 1$. We now have, with $g = 2$,

$$\overline{\mu}_z = \mu_B B_{1/2}(x) = \mu_B \tanh x. \tag{27}$$

For $x \gg 1$, μ_z is very nearly equal to μ_B. For $x \ll 1$, however, $\overline{\mu}_z \simeq \mu_B x$, which corresponds to the Curie constant

$$C_{1/2} = \frac{N_0 \mu_B^2}{k}. \tag{28}$$

In Fig. 3.9 we have reproduced the experimental values of $\overline{\mu}_z$ (in terms of μ_B) as a function of the quantity H/T, for three paramagnetic salts; the corresponding theoretical plots, viz. the curves $gJB_J(x)$, are also included in the figure. The agreement between theory and experiment is indeed good. In passing, we note that, at a temperature of 1.3 K, a field of about 50,000 gauss is sufficient to produce over 99 per cent of saturation in these salts.

3.10. Thermodynamics of magnetic systems: negative temperatures

For the purpose of this section, it will suffice to consider a system of dipoles with $J = \frac{1}{2}$. Each dipole then has a choice of two orientations, the corresponding energies being $-\mu_B H$ and $+\mu_B H$; let us call these energies $-\varepsilon$ and $+\varepsilon$, respectively. The partition function of the system is then given by

$$Q_N(\beta) = (e^{\beta\varepsilon} + e^{-\beta\varepsilon})^N = \{2\cosh(\beta\varepsilon)\}^N; \tag{1}$$

cf. the general expression (3.9.20). Accordingly, the Helmholtz free energy is given by

$$A = -NkT \ln\{2\cosh(\varepsilon/kT)\}, \tag{2}$$

whence

$$S = -\left(\frac{\partial A}{\partial T}\right)_H = Nk\left[\ln\left\{2\cosh\left(\frac{\varepsilon}{kT}\right)\right\} - \frac{\varepsilon}{kT}\tanh\left(\frac{\varepsilon}{kT}\right)\right], \tag{3}$$

$$U = A + TS = -N\varepsilon\tanh\left(\frac{\varepsilon}{kT}\right), \tag{4}$$

$$M = -\left(\frac{\partial A}{\partial H}\right)_T = N\mu_B\tanh\left(\frac{\varepsilon}{kT}\right) \tag{5}$$

and, finally,

$$C = \left(\frac{\partial U}{\partial T}\right)_H = Nk\left(\frac{\varepsilon}{kT}\right)^2 \operatorname{sech}^2\left(\frac{\varepsilon}{kT}\right). \tag{6}$$

Equation (5) is essentially the same as (3.9.27); moreover, as expected, $U = -MH$.

The temperature dependence of the quantities S, U, M and C is shown in Figs 3.10–13. We note that the entropy of the system is vanishingly small for $kT \ll \varepsilon$; it rises rapidly when kT is of the order of ε and approaches the limiting value $Nk \ln 2$ for $kT \gg \varepsilon$. (This limiting value of S corresponds to the fact th at high temperatures the orientation of the dipoles assumes a completely random character, with the result that the system now has 2^N *equally likely* microstate available to it.) The energy of the system attains its lowest value, $-N\varepsilon$, as $T \rightarrow$ 0 K; this clearly corresponds to a state of magnetic saturation and, hence, to a state of perfect order in the system. Towards high temperatures, the energy tends to vanish,[13] implying a purely random orientation of the dipoles and hence a complete loss of magnetic order. These features are re-emphasized in Fig. 3.12 which depicts the temperature dependence of the magnetization M. The specific heat of the system is vanishingly small at low temperatures but, in view of the fac that the energy of the system tends to a constant value as $T \rightarrow \infty$, the specific he

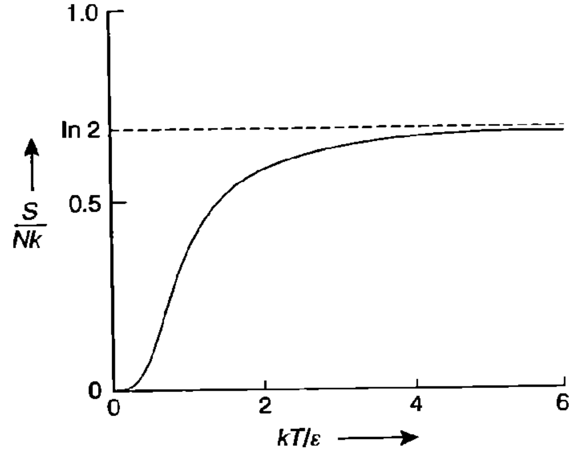


FIG. 3.10. The entropy of a system of magnetic dipoles (with $J = \frac{1}{2}$) as a function of temperature

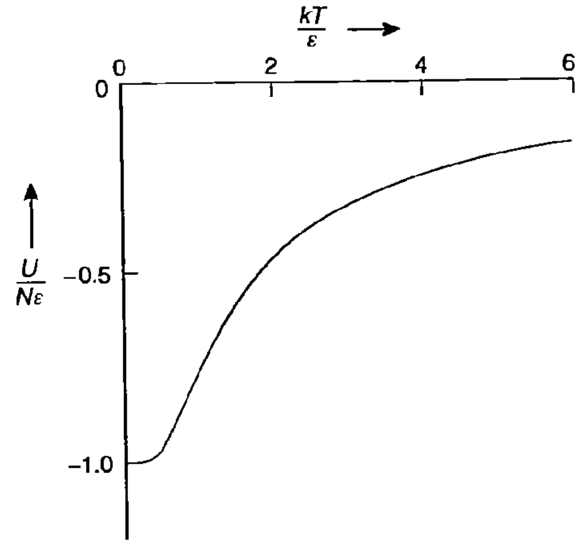


FIG. 3.11. The energy of a system of magnetic dipoles (with $J = \frac{1}{2}$) as a function of temperature

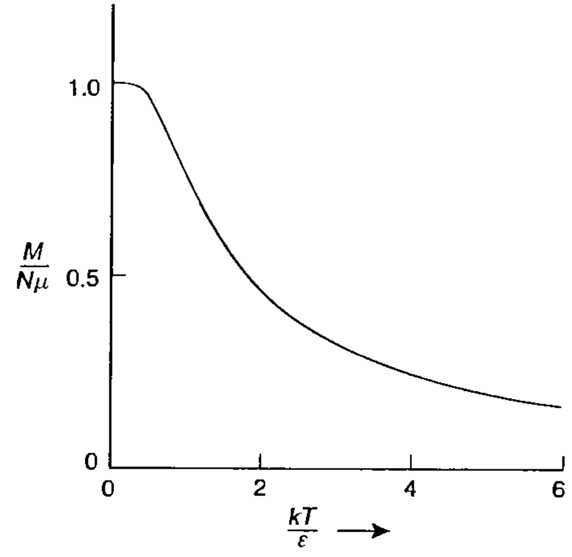


FIG. 3.12. The magnetization of a system of magnetic dipoles (with $J = \frac{1}{2}$) as a function of temperature.

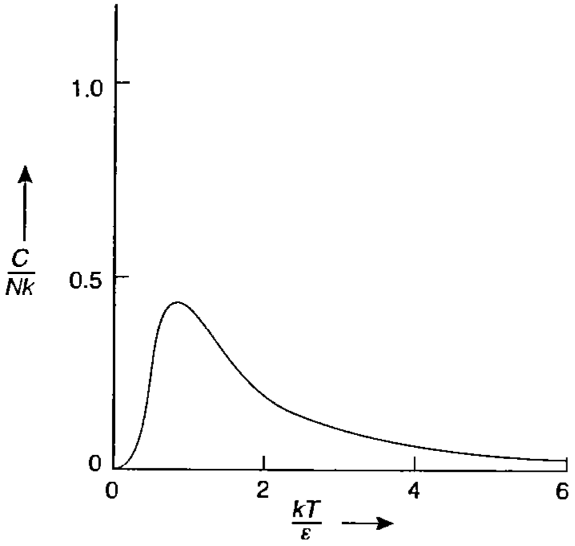


FIG. 3.13. The specific heat of a system of magnetic dipoles (with $J = \frac{1}{2}$) as a function of temperature.

vanishes at high temperatures as well. Somewhere around $T = \varepsilon/k$, it displays a maximum. Writing Δ for the energy difference between the two allowed states of the dipole, the formula for the specific heat can be written as

$$C = Nk\left(\frac{\Delta}{kT}\right)^2 e^{\Delta/kT}\left(1 + e^{\Delta/kT}\right)^{-2} \tag{7}$$

A specific heat of this form is generally known as the *Schottky specific heat*; characterized by an anomalous peak, it is observed in all systems with an excitation gap Δ.

Now, throughout our study so far we have considered only those cases for which $T > 0$. For *normal* systems, this is essential, for otherwise we have to contend with canonical distributions that blow up as the energy of the system is indefinitely increased. If, however, the energy of a system is bounded from above, then there is no compelling reason to exclude the possibility of negative temperatures. Such specialized situations do indeed arise, and the system of magnetic dipoles provides a good example thereof. From eqn. (4) we note that, so long as $U < 0, T > 0$–and that is the only range we covered in Figs 3.10 to 3.13. However, the same equation tells us that if $U > 0$ then $T < 0$, which prompts us to examine the matter a little more closely. For this we may consider the variation of the temperature T and the entropy S with energy U, namely

$$\frac{1}{T} = -\frac{k}{\varepsilon}\tanh^{-1}\left(\frac{U}{N\varepsilon}\right) = \frac{k}{2\varepsilon}\ln\left(\frac{N\varepsilon - U}{N\varepsilon + U}\right) \tag{8}$$

and

$$\frac{S}{Nk} = -\frac{N\varepsilon + U}{2N\varepsilon}\ln\left(\frac{N\varepsilon + U}{2N\varepsilon}\right) - \frac{N\varepsilon - U}{2N\varepsilon}\ln\left(\frac{N\varepsilon - U}{2N\varepsilon}\right); \tag{9}$$

these expressions follow straightforwardly from eqns (3) and (4), and are shown graphically in Figs 3.14 and 3.15. We note that for $U = -N\varepsilon$, both S and T vanish. As U increases, they too increase, until we reach the special situation where $U = 0$. The entropy is then seen to have attained its maximum value $Nk \ln 2$, while the temperature has reached infinity. Throughout this range, the entropy was a monotonically increasing function of energy, so T was positive. Now, as U becomes 0_+, (dS/dU) becomes 0_- and T becomes $-\infty$. With a further increase in U, the entropy monotonically decreases; as a result, the temperature continues to be negative, though its magnitude steadily decreases. Finally, we reach the largest value of U, namely $+N\varepsilon$, where the entropy is once again zero and $T = -0$.

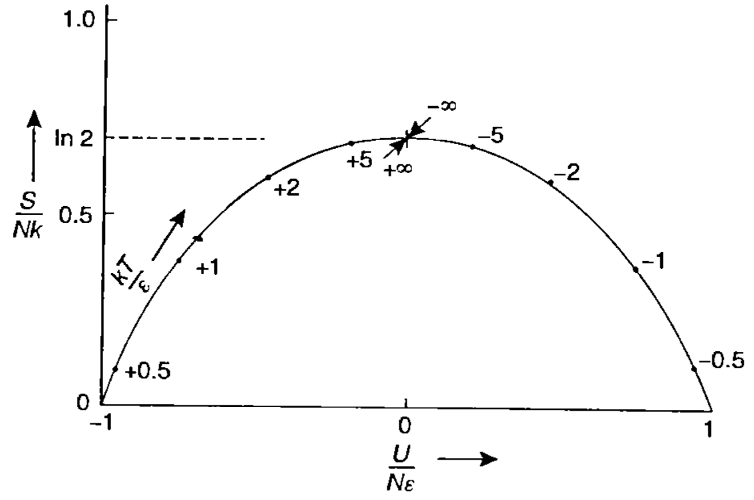


Fig. 3.14. The entropy of a system of magnetic dipoles (with $J = \frac{1}{2}$) as a function of energy. Some values of the parameter kT/ε are also shown in the figure.

The region where $U > 0$ (and hence $T < 0$) is indeed abnormal because it corresponds to a magnetization *opposite* in direction to that of the applied field! Nevertheless, it can be realized experimentally in the system of nuclear moments of

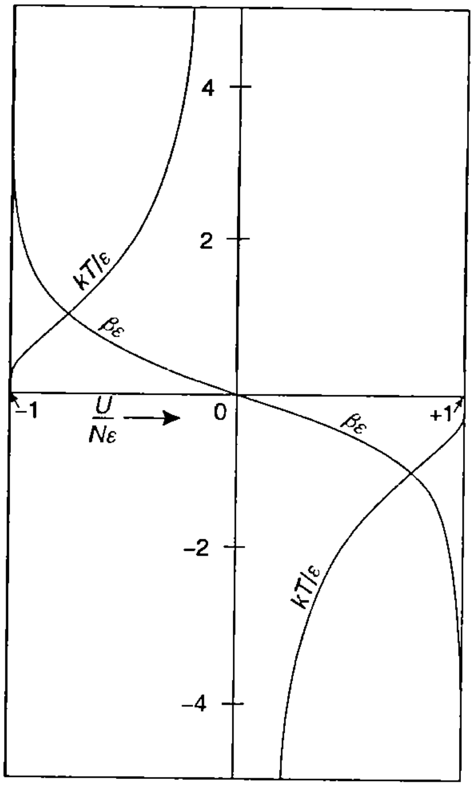


FIG. 3.15. The temperature parameter kT/ε, and its reciprocal $\beta\varepsilon$, for a system of magnetic dipoles (with $J = \frac{1}{2}$) as a function of energy.

a crystal in which the relaxation time t_1 for mutual interaction among nuclear spins is very small in comparison with the relaxation time t_2 for interaction between the spins and the lattice. Let such a crystal be magnetized in a strong magnetic field, and then the field reversed so quickly that the spins are unable to follow the switch-over. This will leave the system in a non-equilibrium state, with energy higher than the new equilibrium value U. During a period of order t_1, the sub-system of the nuclear spins should be able to attain a state of *internal* equilibrium; this state will have a negative magnetization and will, therefore, correspond to a negative temperature. The sub-system of the lattice, which involves energy parameters that are in principle unbounded, will indeed be at a positive temperature. During a period of order t_2, the two sub-systems would attain a state of mutual equilibrium, which again will have a positive temperature.[14] An experiment of this kind was successfully performed by Purcell and Pound (1951) with a crystal of LiF; in this case, t_1 was of order 10^{-5} sec while t_2 was of order 5 min. A state of negative temperature for the sub-system of spins was indeed attained and was found to persist for a period of several minutes; see Fig. 3.16.

Before we close this discussion, a few general remarks seem in order. First of all, we should clearly note that the onset of negative temperatures is possible only if there exists an upper limit on the energy of the given system. In most physical systems this is not the case, simply because most physical systems possess kinetic energy of motion which is obviously unbounded. By the same token, the onset of positive temperatures is related to the existence of a lower limit on the energy of a system; this, however, does not present any problems because, if nothing else,

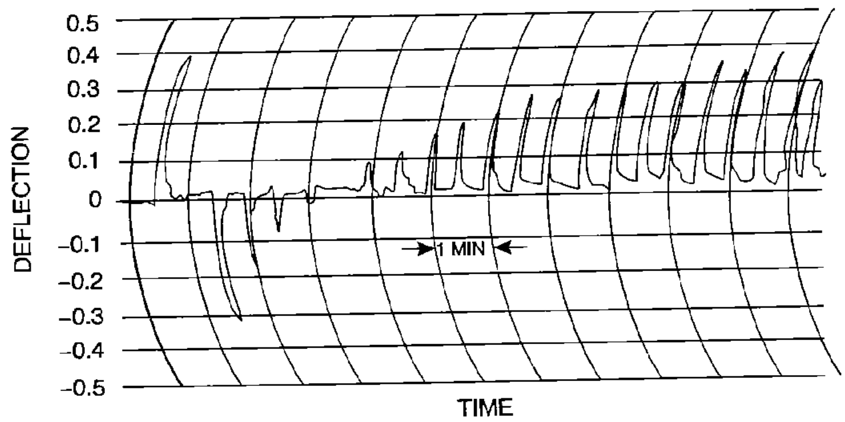


FIG. 3.16. A typical record of the reversed nuclear magnetization (after Purcell and Pound, 1951). On the left we have a deflection corresponding to normal, equilibrium magnetization ($T \sim 300$ K); it is followed by the reversed deflection (corresponding to $T \sim -350$ K), which decays through zero deflection (corresponding to a passage from $T = -\infty$ to $T = +\infty$) towards the new equilibrium state which again has a positive T.

the uncertainty principle alone is sufficient to set such a limit for *every* physical system. Thus, it is quite normal for a system to be at a positive temperature whereas it is very unusual for one to be at a negative temperature. Now, suppose that we have a system whose energy *cannot* assume unlimited high values. Then we can surely visualize a temperature T such that the quantity NkT is much large than any admissible value, E_r, of the energy. At such a high temperature, the mutual interactions of the microscopic entities constituting the system may be regarded as negligible; accordingly, one may write for the partition function o the system

$$Q_N(\beta) \simeq \left[\sum_n e^{-\beta\varepsilon_n}\right]^N . \qquad \text{(1}$$

Since, by assumption, all $\beta\varepsilon_n \ll 1$, we have

$$Q_N(\beta) \simeq \left[\sum_n \left\{1 - \beta\varepsilon_n + \tfrac{1}{2}\beta^2\varepsilon_n^2\right\}\right]^N . \qquad \text{(11}$$

Let g denote the number of possible orientations of a microscopic constituent the system with respect to the direction of the external field; then, the quantiti $\sum_n \varepsilon_n^\alpha (\alpha = 0, 1, 2)$ may be replaced by the quantities $g\overline{\varepsilon^\alpha}$. So, we get

$$\begin{aligned} \ln Q_N(\beta) &\simeq N\left[\ln g + \ln\left(1 - \beta\bar{\varepsilon} + \tfrac{1}{2}\beta^2\overline{\varepsilon^2}\right)\right] \\ &\simeq N\left[\ln g - \beta\bar{\varepsilon} + \tfrac{1}{2}\beta^2\left(\overline{\varepsilon^2} - \bar{\varepsilon}^2\right)\right] . \qquad \text{(1} \end{aligned}$$

The Helmholtz free energy of the system is then given by

$$A(N, \beta) \simeq -\frac{N}{\beta}\ln g + N\bar{\varepsilon} - \frac{N}{2}\beta\overline{(\varepsilon - \bar{\varepsilon})^2}, \qquad \text{(1}$$

whence

$$S(N, \beta) \simeq Nk \ln g - \frac{Nk}{2}\beta^2\overline{(\varepsilon - \bar{\varepsilon})^2} \tag{14}$$

$$U(N, \beta) \simeq N\bar{\varepsilon} - N\beta\overline{(\varepsilon - \bar{\varepsilon})^2}, \tag{15}$$

and

$$C(N, \beta) \simeq Nk\beta^2\overline{(\varepsilon - \bar{\varepsilon})^2}. \qquad (16)^{15}$$

Formulae (12)–(16) determine the thermodynamic properties of the system for $\beta \simeq 0$. The important thing to note here is that they do so not only for $\beta \gtrsim 0$ but also for $\beta \lesssim 0$. In fact, these formulae hold in the vicinity of, and on *both* sides of, the maximum in the S–U curve; see Fig. 3.14. Quite expectedly, the maximum value of S is given by $Nk \ln g$, and it occurs at $\beta = \pm 0$; S here decreases both ways, whether U decreases ($\beta > 0$) or increases ($\beta < 0$). It will be noted that the specific heat of the system is positive in either case.

It is not difficult to show that if two systems, characterized by the temperature parameters β_1 and β_2, are brought into thermal contact, then energy will flow from the system with the smaller value of β to the system with the larger value of β; this will continue until the two systems acquire a common value of this parameter. What is more important to note is that this result remains *literally* true even if one or both of the β are negative. Thus, if β_1 is −ve while β_2 is +ve, then energy will flow from system 1 to system 2, i.e. from the system at negative temperature to the one at positive temperature. In this sense, systems at negative temperatures are *hotter* than the ones at positive temperatures; indeed, all negative temperatures are above +∞, not below zero!

For further discussion of this topic, reference may be made to a paper by Ramsey (1956).

Problems

3.1. (a) Derive formula (3.2.36) from eqns (3.2.14) and (3.2.35).
(b) Derive formulae (3.2.39) and (3.2.40) from eqns (3.2.37) and (3.2.38).

3.2. Prove that the quantity $g''(x_0)$, see eqn. (3.2.25), is equal to $\langle (E - U)^2 \rangle \exp(2\beta)$. Hence show that eqn. (3.2.28) is physically equivalent to eqn. (3.6.9).

3.3. Using the fact that $(1/n!)$ is the coefficient of x^n in the power expansion of the function $\exp(x)$, derive an asymptotic formula for this coefficient by the method of saddle-point integration. Compare your result with the Stirling formula for $(n!)$.

3.4. Verify that the quantity $(k/N) \ln \Gamma$, where

$$\Gamma(N, U) = \sum_{\{n_r\}}{}' W\{n_r\},$$

is equal to the (mean) entropy of the given system. Show that we obtain essentially the same result for $\ln \Gamma$ if we take, in the foregoing summation, *only* the largest term of the sum, viz. the term $W\{n_r^*\}$ which corresponds to the *most probable* distribution set.

[Surprised? Well, note the following example:

For all N, $$\sum_{r=0}^{N} {}^N C_r = 2^N;$$

therefore,
$$\ln\left\{\sum_{r=0}^{N} {}^N C_r\right\} = N\ln 2. \tag{a}$$

Now, the largest term in the sum corresponds to $r \simeq N/2$; so, for large N, the logarithm of the largest term is very nearly equal to

$$\begin{aligned}&\ln\{N!\} - 2\ln\{(N/2)!\}\\ &\approx N\ln N - 2\frac{N}{2}\ln\frac{N}{2} = N\ln 2,\end{aligned} \tag{b}$$

which agrees with (a).]

3.5. Making use of the fact that the Helmholtz free energy $A(N, V, T)$ of a thermodynamic system is an *extensive* property of the system, show that

$$N\left(\frac{\partial A}{\partial N}\right)_{V,T} + V\left(\frac{\partial A}{\partial V}\right)_{N,T} = A.$$

[Note that this result implies the well-known relationship: $N\mu = A + PV(\equiv G)$.]

3.6. (a) Assuming that the total number of microstates accessible to a given statistical system is Ω, show that the entropy of the system, as given by eqn. (3.3.13), is maximum when all Ω states are equally likely to occur.

(b) If, on the other hand, we have an ensemble of systems sharing energy (with mean value $\bar{E}$), then show that the entropy, as given by the same formal expression, is maximum when $P_r \propto \exp(-\beta E_r)$, β being a constant to be determined by the given value of $\bar{E}$.

(c) Further, if we have an ensemble of systems sharing energy (with mean value $\bar{E}$) and also sharing particles (with mean value $\bar{N}$), then show that the entropy, given by a similar expression, is maximum when $P_{r,s} \propto \exp(-\alpha N_r - \beta E_s)$, α and β being constants to be determined by the given values of $\bar{N}$ and $\bar{E}$.

3.7. Prove that, quite generally,

$$C_P - C_V = -k\frac{\left[\frac{\partial}{\partial T}\left\{T\left(\frac{\partial \ln Q}{\partial V}\right)_T\right\}\right]_V^2}{\left(\frac{\partial^2 \ln Q}{\partial V^2}\right)_T} > 0.$$

Verify that the value of this quantity for an ideal classical gas is Nk.

3.8. Show that, for an ideal gas,

$$\frac{S}{Nk} = \ln\left(\frac{Q_1}{N}\right) + T\left(\frac{\partial \ln Q_1}{\partial T}\right)_P.$$

3.9. If an ideal monatomic gas is expanded *adiabatically* to twice its initial volume, what will the ratio of the final pressure to the initial pressure be? If during the process some heat is added to the system, will the final pressure be higher or lower than in the preceding case? Support your answer by deriving the relevant formula for the ratio P_f/P_i.

3.10. (a) The volume of a sample of helium gas is increased by withdrawing the piston of the containing cylinder. The final pressure P_f is found to be equal to the initial pressure P_i times $(V_i/V_f)^{1.2}$, V_i and V_f being the initial and final volumes. Assuming that the product PV is always equal to $\frac{2}{3}U$, will (i) the energy and (ii) the entropy of the gas increase, remain constant or decrease during the process?

(b) If the process were reversible, how much would the work done and the heat added in doubling the volume of the gas be? Take $P_i = 1$ atm and $V_i = 1$ m^3.

3.11. Determine the work done on a gas and the amount of heat absorbed by it during a compression from volume V_1 to volume V_2, following the law $PV^n = \text{const}$.

3.12. If the "free volume" $\bar{V}$ of a classical system is defined by the equation

$$\bar{V}^N = \int e^{\{\bar{U}-U(q_i)\}/kT}\prod_{i=1}^{N} d^3q_i,$$

where $\bar{U}$ is the average potential energy of the system and $U(\mathbf{q}_i)$ the actual potential energy as a function of the molecular configuration, then show that

$$S = Nk\left[\ln\left\{\frac{\bar{V}}{N}\left(\frac{2\pi mkT}{h^2}\right)^{3/2}\right\} + \frac{5}{2}\right].$$

In what sense is it justified to refer to the quantity $\bar{V}$ as the "free volume" of the system? Substantiate your answer by considering a particular case, e.g. the case of a hard sphere gas.

3.13. (a) Evaluate the partition function and the major thermodynamic properties of an ideal gas consisting of N_1 molecules of mass m_1 and N_2 molecules of mass m_2, confined to a space of volume V at temperature T. Assume that the molecules of a given kind are mutually indistinguishable, while those of one kind are distinguishable from those of the other kind.

(b) Compare your results with the ones pertaining to an ideal gas consisting of $(N_1 + N_2)$ molecules, *all of one kind*, of mass m, such that $m(N_1 + N_2) = m_1N_1 + m_2N_2$.

3.14. Consider an ideal-gas mixture of atoms A, atoms B and molecules AB, undergoing the reaction $AB \Leftrightarrow A + B$. If n_A, n_B and n_{AB} denote their respective concentrations (that is, the number densities of respective atoms/molecules), then show that, in equilibrium,

$$\frac{n_{AB}}{n_An_B} = V\frac{f_{AB}}{f_Af_B} = K(T) \quad \text{(the law of mass action).}$$

Here, V is the volume of the system while the f_i are the respective single-particle partition functions; the quantity $K(T)$ is generally referred to as the *equilibrium constant* of the reaction.

3.15. Show that the partition function $Q_N(V,T)$ of an *extreme* relativistic gas consisting of N monatomic molecules with energy–momentum relationship $\varepsilon = pc$, c being the speed of light, is given by

$$Q_N(V,T) = \frac{1}{N!}\left\{8\pi V\left(\frac{kT}{hc}\right)^3\right\}^N$$

Study the thermodynamics of this system, checking in particular that

$$PV = \tfrac{1}{3}U, \quad U/N = 3kT \quad \text{and} \quad \gamma = \tfrac{4}{3}.$$

Next, using the inversion formula (3.4.7), derive an expression for the density of states $g(E)$ of this system.

3.16. Consider a system similar to the one in the preceding problem, but consisting of $3N$ particles moving in one dimension. Show that the partition function in this case is given by

$$Q_{3N}(L,T) = \frac{1}{(3N)!}\left[2L\left(\frac{kT}{hc}\right)\right]^{3N}$$

L being the "length" of the space available. Compare the thermodynamics and the density of states of this system with the corresponding quantities obtained in the preceding problem.

3.17. If we take the function $f(q,p)$ in eqn. (3.5.3) to be $U - H(q,p)$, then clearly $\langle f\rangle = 0$; formally, this would mean

$$\int [U - H(q,p)]e^{-\beta H(q,p)}\,d\omega = 0.$$

Derive from this equation expression (3.6.3) for the mean-square fluctuation in the energy of a system embedded in the canonical ensemble.

3.18. Show that for a system in the canonical ensemble

$$\langle(\Delta E)^3\rangle = k^2\left\{T^4\left(\frac{\partial C_V}{\partial T}\right)_V + 2T^3C_V\right\}.$$

Verify that for an ideal gas

$$\left\langle\left(\frac{\Delta E}{U}\right)^2\right\rangle = \frac{2}{3N} \text{ and } \left\langle\left(\frac{\Delta E}{U}\right)^3\right\rangle = \frac{8}{9N^2}.$$

3.19. Consider the long-time averaged behavior of the quantity dG/dt, where

$$G = \sum_i q_ip_i,$$

and show that the validity of eqn. (3.7.5) implies the validity of eqn. (3.7.6), and vice versa.

3.20. Show that, for a statistical system in which the interparticle potential energy $u(\boldsymbol{r})$ is a homogeneous function (of degree n) of the particle coordinates, the *virial* $\mathcal{V}$ is given by

$$\mathcal{V} = -3PV - nU$$

and, hence, the *mean kinetic energy* K by

$$K = -\frac{1}{2}\mathcal{V} = \frac{1}{2}(3PV + nU) = \frac{1}{(n+2)}(3PV + nE);$$

here, U denotes the *mean potential energy* of the system and $E = K + U$. Note that this result holds not only for a classical system but for a quantum-mechanical one as well.

3.21. (a) Calculate the time-averaged kinetic energy and potential energy of a one-dimensional harmonic oscillator, both classically and quantum-mechanically, and show that the results obtained are consistent with the result established in the preceding problem (with $n = +2$).

(b) Consider, similarly, the case of the hydrogen atom ($n = -1$) on the basis of (i) the Bohr–Sommerfeld model and (ii) the Schrödinger model.

(c) Finally, consider the case of a planet moving in (i) a circular orbit or (ii) an elliptic orbit around the sun.

3.22. The restoring force of an anharmonic oscillator is proportional to the cube of the displacement. Show that the mean kinetic energy of the oscillator is *twice* its mean potential energy.

3.23. (a) For a dilute gas, the pair distribution function $g(r)$ may be approximated as

$$g(r) \simeq \exp\{-u(r)/kT\}.$$

Show that, under this approximation, eqn. (3.7.17) takes the form

$$\frac{PV}{NkT} \simeq 1 - 2\pi n \int_0^\infty f(r) r^2\, dr,$$

where $f(r)[= \exp\{-u(r)/kT\} - 1]$ is the so-called *Mayer function*; see eqn. (9.1.6).

(b) What form will this result take for a gas of hard spheres, for which

$$\begin{aligned} u(r) &= \infty \quad \text{for } r \le \sigma \\ &\;\; 0 \quad\;\; \text{otherwise?} \end{aligned}$$

Compare your result with that of Problem 1.4.

3.24. Show that in the relativistic case the equipartition theorem takes the form

$$\langle m_0 u^2 (1 - u^2/c^2)^{-1/2} \rangle = 3kT,$$

where m_0 is the rest mass of a particle and u its speed. Check that in the extreme relativistic case the mean thermal energy per particle is twice its value in the nonrelativistic case.

3.25. Develop a *kinetic* argument to show that in a non-interacting system the average value of the quantity $\sum_i p_i \dot{q}_i$ is precisely equal to $3PV$. Hence show that, regardless of relativistic considerations, $PV = NkT$.

3.26. The energy eigenvalues of an s-dimensional harmonic oscillator can be written as

$$\varepsilon_j = (j + s/2)\hbar\omega; \qquad j = 0, 1, 2, \ldots$$

Show that the jth energy level has a multiplicity $(j + s - 1)!/j!(s - 1)!$. Evaluate the partition function, and the major thermodynamic properties, of a system of N such oscillators, and compare your results with a corresponding system of sN one-dimensional oscillators. Show, in particular, that the chemical potential $\mu_s = s\mu_1$.

3.27. Obtain an asymptotic expression for the quantity $\ln g(E)$ for a system of N quantum-mechanical harmonic oscillators by using the inversion formula (3.4.7) and the partition function (3.8.15). Hence show that

$$\frac{S}{Nk} = \left(\frac{E}{N\hbar\omega} + \frac{1}{2}\right)\ln\left(\frac{E}{N\hbar\omega} + \frac{1}{2}\right) - \left(\frac{E}{N\hbar\omega} - \frac{1}{2}\right)\ln\left(\frac{E}{N\hbar\omega} - \frac{1}{2}\right).$$

[*Hint:* Employ the Darwin–Fowler method.]

3.28. (a) When a system of N oscillators with total energ, is in thermal equilibrium, what is the probability p_n that a particular oscillator among them is in the quantum state n?

[*Hint:* Use expression (3.8.25).] Show that, for $N \gg 1$ and $R \gg n$, $p_n \approx (\bar{n})^n/(\bar{n}+1)^{n+1}$, where $\bar{n} = R/N$.

(b) When an ideal gas of N monatomic molecules with total energy E is in thermal equilibrium, show that the probability of a particular molecule having an energy in the neighborhood of ε is proportional to $\exp(-\beta\varepsilon)$, where $\beta = 3N/2E$.
[*Hint:* Use expression (3.5.16) and assume that $N \gg 1$ and $E \gg \varepsilon$.]

3.29. The potential energy of a one-dimensional, *anharmonic* oscillator may be written as

$$V(q) = cq^2 - gq^3 - fq^4,$$

where c, g and f are positive constants; quite generally, g and f may be assumed to be very small in value. Show that the leading contribution of anharmonic terms to the heat capacity of the oscillator, assumed classical, is given by

$$\frac{3}{2}k^2\left(\frac{f}{c^2} + \frac{5}{4}\frac{g^2}{c^3}\right)T$$

and, to the same order, the mean value of the position coordinate q is given by

$$\frac{3}{4}\frac{gkT}{c^2}.$$

3.30. The energy levels of a quantum-mechanical, one-dimensional, *anharmonic* oscillator may be approximated as

$$\varepsilon_n = \left(n + \tfrac{1}{2}\right)\hbar\omega - x\left(n + \tfrac{1}{2}\right)^2\hbar\omega; \qquad n = 0, 1, 2, \ldots$$

The parameter x, usually $\ll 1$, represents the degree of anharmonicity. Show that, to the first order in x and the fourth order in $u(\equiv \hbar\omega/kT)$, the specific heat of a system of N such oscillators is given by

$$C = Nk\left[\left(1 - \frac{1}{12}u^2 + \frac{1}{240}u^4\right) + 4x\left(\frac{1}{u} + \frac{1}{80}u^3\right)\right].$$

Note that the correction term *increases* with temperature.

3.31. Study, along the lines of Sec. 3.8, the statistical mechanics of a system of N "Fermi oscillators" which are characterized by only two eigenvalues, namely 0 and ε.

3.32. The quantum states available to a given physical system are (i) a group of g_1 *equally likely* states, with a common energy ε_1 and (ii) a group of g_2 *equally likely* states, with a common energy $\varepsilon_2 > \varepsilon_1$. Show that the entropy of the system is given by

$$S = -k[p_1 \ln(p_1/g_1) + p_2 \ln(p_2/g_2)],$$

where p_1 and p_2 are, respectively, the probabilities of the system being in a state belonging to group 1 or to group 2: $p_1 + p_2 = 1$.

(a) Assuming that the p_i are given by a canonical distribution, show that

$$S = k\left[\ln g_1 + \ln\{1 + (g_2/g_1)e^{-x}\} + \frac{x}{1 + (g_1/g_2)e^x}\right],$$

where $x = (\varepsilon_2 - \varepsilon_1)/kT$, assumed positive. Compare the special case $g_1 = g_2 = 1$ with that of the Fermi oscillator of the preceding problem.

(b) Verify the foregoing expression for S by deriving it from the partition function of the system.

(c) Check that at $T \to 0$, $S \to k \ln g_1$. Interpret this result physically.

3.33. Gadolinium sulphate obeys Langevin's theory of paramagnetism down to a few degrees Kelvin. Its molecular magnetic moment is 7.2×10^{-23} amp-m^2. Determine the degree of magnetic saturation in this salt at a temperature of 2 K in a field of flux density 2 weber/m^2.

3.34. Oxygen is a paramagnetic gas obeying Langevin's theory of paramagnetism. Its susceptibility per unit volume, at 293 K and at atmospheric pressure, is 1.80×10^{-6} mks units. Determine its molecular magnetic moment and compare it with the Bohr magneton (which is very nearly equal to 9.27×10^{-24} amp-m^2).

3.35. (a) Consider a gaseous system of N non-interacting, diatomic molecules, each having an electric dipole moment μ, placed in an external electric field of strength E. The energy of a molecule will be given by the kinetic energy of rotation as well as of translation *plus* the potential energy of orientation in the applied field:

$$\varepsilon = \frac{p^2}{2m} + \left\{ \frac{p_\theta^2}{2I} + \frac{p_\phi^2}{2I\sin^2\theta} \right\} - \mu E \cos\theta,$$

where I is the moment of inertia of the molecule. Study the thermodynamics of this system, including the electric polarization and the dielectric constant. Assume that (i) the system is a classical one and (ii) $|\mu E| \ll kT$.[16]

(b) The molecule H_2O has an electric dipole moment of 1.85×10^{-18} e.s.u. Calculate, on the basis of the preceding theory, the dielectric constant of steam at 100°C and at atmospheric pressure.

3.36. Consider a pair of electric dipoles μ and μ', oriented in the directions (θ, ϕ) and (θ', ϕ'), respectively; the distance R between their centers is assumed to be fixed. The potential energy in this orientation is given by

$$-\frac{\mu\mu'}{R^3}\{2\cos\theta\cos\theta' - \sin\theta\sin\theta'\cos(\phi - \phi')\}.$$

Now, consider the pair to be in thermal equilibrium, their orientations being governed by a canonical distribution. Show that the mean force between the dipoles, at high temperatures, is given by

$$-2\frac{\mu\mu'}{kT}\frac{\hat{R}}{R^7},$$

$\hat{R}$ being a unit vector in the direction of the line of centers.

3.37. Evaluating the high-temperature approximation of the partition function of a system of magnetic dipoles, show that the Curie constant C_J is given by

$$C_J = \frac{N_0 g^2 \mu_B^2}{k}\overline{m^2}.$$

Hence derive the formula (3.9.26).

3.38. Replacing the sum in (3.9.18) by an integral, evaluate $Q_1(\beta)$ of the given magnetic dipole and study the thermodynamics following from it. Compare these results with the ones following from the Langevin theory.

3.39. Atoms of silver vapor, each having a magnetic moment $\mu_B(g = 2, J = \frac{1}{2})$, align themselves either parallel or antiparallel to the direction of an applied magnetic field. Determine the respective fractions of atoms aligned parallel and antiparallel to a field of flux density 0.1 weber/m^2 at a temperature of 1000 K.

3.40. (a) Show that, for any magnetizable material, the heat capacities at constant field H and at constant magnetization M are connected by the relation

$$C_H - C_M = -T\left(\frac{\partial H}{\partial T}\right)_M \left(\frac{\partial M}{\partial T}\right)_H.$$

(b) Show that for a paramagnetic material obeying Curie's law

$$C_H - C_M = CH^2/T^2,$$

where C on the right-hand side of this equation denotes the *Curie constant* of the given sample.

3.41. A system of N spins at a negative temperature ($E > 0$) is brought into contact with an ideal-gas thermometer consisting of N' molecules. What will the nature of their state of mutual equilibrium be? Will their common temperature be negative or positive, and in what manner will it be affected by the ratio N'/N?

3.42. Consider the system of N magnetic dipoles, studied in Sec. 3.10, in the microcanonical ensemble. Enumerate the number of microstates, $\Omega(N, E)$, accessible to the system at energy E, and evaluate the quantities $S(N, E)$ and $T(N, E)$. Compare your results with eqns (3.10.8) and (3.10.9).

3.43. Consider a system of charged particles (not dipoles), obeying classical mechanics and classical statistics. Show that the magnetic susceptibility of this system is identically zero (Bohr–van Leeuwen theorem).

[Note that the Hamiltonian of this system in the presence of a magnetic field $H (= \nabla \times A)$ will be a function of the quantities $p_j + (e_j/c)A(r_j)$, and not of the p_j as such. One has now to show that the partition function of the system is independent of the applied field.]

Notes

[1] In what follows, the energy levels E_r appear as purely *mechanical* quantities–independent of the temperature of the system. For a treatment involving "temperature-dependent energy levels", see Elcock and Landsberg (1957).

[2] For the method of Lagrange multipliers, see ter Haar and Wergeland (1966), Appendix C.1.

[3] Actually, this is not a serious restriction at all, for a common divisor, if any, can be removed by selecting the unit of energy correspondingly larger.

[4] This too is not serious, for by doing so we are merely shifting the zero of the energy scale; the mean energy U then becomes $U - E_0$, but we can agree to call it U again.

[5] This can be seen by noting that (i) an analytic function must possess a *unique* derivative everywhere (so, in our case, it must be zero, irrespective of the direction in which we pass through the point x_0), and (ii) by the Cauchy–Riemann conditions of analyticity, the second derivative of the function with respect to y must be equal and opposite to the second derivative with respect to x.

[6] It is indeed true that, for large $\mathcal{N}$, the contribution from the rest of the circle is negligible. The intuitive reason for this is that the terms $(\omega_r z^{E_r})$, which constitute the function $f(z)$, "reinforce" one another *only* at the point $z = x_0$; elsewhere, there is bound to be disagreement among their phases, so that at *all* other points along the circle, $|f(z)| < f(x_0)$. Now, the factor that actually governs the relative contributions is $[|f(z)|/f(x_0)]^{\mathcal{N}}$; for $\mathcal{N} \gg 1$, this will clearly be negligible. For a rigorous demonstration of this point, see Schrödinger (1960), pp. 31–33.

[7] Of course, if the ground state of the system is *degenerate* (with a multiplicity Ω_0), then the ground-state entropy is nonzero and is given by the expression $k \ln \Omega_0$; see eqn. (14).

[8] Ample justification has already been given for the factor h^{3N}. The factor $N!$ comes from the considerations of Secs 1.5 and 1.6; it arises essentially from the fact that the particles constituting the given system are after all *indistinguishable* For a complete proof, see Sec. 5.5.

[9] For the details of this evaluation, see Kubo (1965), pp. 165–8.

[10] Subsequently we shall see that this extremum is actually a maximum–and that too an extremely sharp one.

[11] We observe that if x_j is a space coordinate, then its extreme values will correspond to "locations at the walls of the container"; accordingly, the potential energy of the system would become infinite. If, on the other hand, x_j is a momentum coordinate, then its extreme values will themselves be $\pm\infty$, in which case the kinetic energy of the system would become infinite.

[12] It will be noted that the summation over the various particles of the system, which appears in the definition of the virial, has been replaced by an integration over the surface of the container, for the simple reason that no contribution arises from the interior of the container.

[13] Note that in the present study we are completely disregarding the kinetic energy of the dipoles.

[14] Note that in the latter process, during which the spins realign themselves (now more favorably in the new direction of the field), the energy will flow from the sub-system of the spins to that of the lattice, and not vice versa. This is in agreement with the fact that negative temperatures are *hotter* than positive ones; see the subsequent discussion in the text.

[15] Compare this result with eqn. (3.6.3).

[16] The electric dipole moments of molecules are generally of order 10^{-18} e.s.u. (or a *Debye unit*). In a field of 1 e.s.u. (= 300 volts/cm) and at a temperature of 300 K, the parameter $\beta\mu E = O(10^{-4})$.

CHAPTER 4

THE GRAND CANONICAL ENSEMBLE

IN THE preceding chapter we developed the formalism of the canonical ensemble and established a scheme of operations for deriving the various thermodynamic properties of a given physical system. The effectiveness of that approach became clear from the examples discussed there; it will become even more vivid in the subsequent studies carried out in this text. However, for a number of problems, both physical and chemical, the usefulness of the canonical ensemble formalism turns out to be rather limited and it appears that a further generalization of the formalism is called for. The motivation that brings about this generalization is physically of the same nature as the one that led us from the microcanonical to the canonical ensemble–it is just the next natural step from there. It comes from the realization that not only the energy of a system but the number of particles as well is hardly ever measured in a "direct" manner; we only estimate it through an indirect probing into the system. Conceptually, therefore, we may regard both N and E as *variables* and identify their expectation values, $\langle N \rangle$ and $\langle E \rangle$, with the corresponding thermodynamic quantities.

The procedure for studying the statistics of the variables N and E is self-evident. We may *either* (i) consider the given system A to be immersed in a large reservoir A' with which it can exchange both energy and particles *or* (ii) regard it as a member of what we may call a *grand canonical ensemble*, which consists of the given system A and a large number of (mental) copies thereof, the members of the ensemble carrying out a mutual exchange of both energy and particles. The end results, in either case, are asymptotically the same.

4.1. Equilibrium between a system and a particle–energy reservoir

We consider the given system A as immersed in a large reservoir A', with which it can exchange both energy and particles; see Fig. 4.1. After some time has elapsed, the system and the reservoir are supposed to attain a state of mutual equilibrium. Then, according to Sec. 1.3, the system and the reservoir will have a common temperature T and a common chemical potential μ. The fraction of the total number of particles $N^{(0)}$ and the fraction of the total energy $E^{(0)}$ which the system A can have at any time t are, howe[illegible] variables (whose values, in principle, can lie anywhere between zero and unity). If, at a particular instant of time, the system A happens to be in *one* of its states characterized by the number N_r of particles and the amount E_s of energy, then the number of particles in the reservoir would be N'_r and its energy E'_s, such that

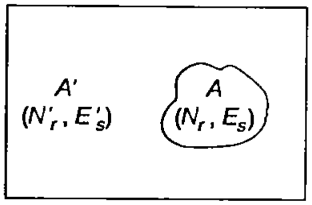


FIG. 4.1. A statistical system immersed in a particle–energy reservoir.

$$N_r + N'_r = N^{(0)} = \text{const.} \tag{1}$$

and

$$E_s + E'_s = E^{(0)} = \text{const.} \tag{2}$$

Again, since the reservoir is supposed to be much larger than the given system, the values of N_r and E_s, which are going to be of practical importance, will be very small fractions of the total magnitudes $N^{(0)}$ and $E^{(0)}$, respectively; therefore, for all practical purposes,[1]

$$\frac{N_r}{N^{(0)}} = \left(1 - \frac{N'_r}{N^{(0)}}\right) \ll 1 \tag{3}$$

and

$$\frac{E_s}{E^{(0)}} = \left(1 - \frac{E'_s}{E^{(0)}}\right) \ll 1. \tag{4}$$

Now, in the manner of Sec. 3.1, the probability $P_{r,s}$ that, at any time t, the system A is found to be in an (N_r, E_s)-state would be directly proportional to the number of microstates $\Omega'(N'_r, E'_s)$ which the reservoir can have for the corresponding macrostate (N'_r, E'_s). Thus

$$P_{r,s} \propto \Omega'(N^{(0)} - N_r, E^{(0)} - E_s). \tag{5}$$

Again, in view of (3) and (4), we can write

$$\begin{aligned}
\ln \Omega'(N^{(0)} - N_r, E^{(0)} - E_s) &= \ln \Omega'(N^{(0)}, E^{(0)}) \\
&\quad + \left(\frac{\partial \ln \Omega'}{\partial N'}\right)_{N'=N^{(0)}} (-N_r) + \left(\frac{\partial \ln \Omega'}{\partial E'}\right)_{E'=E^{(0)}} (-E_s) + \cdots \\
&\simeq \ln \Omega'(N^{(0)}, E^{(0)}) + \frac{\mu'}{kT'} N_r - \frac{1}{kT'} E_s;
\end{aligned} \tag{6}$$

see eqns (1.2.3), (1.2.7), (1.3.3) and (1.3.5). Here, μ' and T' are, respectively, the chemical potential and the temperature of the reservoir (and hence of the given system as well). From (5) and (6), we obtain the desired result:

$$P_{r,s} \propto \exp(-\alpha N_r - \beta E_s), \tag{7}$$

with

$$\alpha = -\mu/kT, \quad \beta = 1/kT. \tag{8}$$

On normalization, it becomes

$$P_{r,s} = \frac{\exp(-\alpha N_r - \beta E_s)}{\sum\limits_{r,s} \exp(-\alpha N_r - \beta E_s)}; \tag{9}$$

the summation in the denominator goes over *all* the (N_r, E_s)-states accessible to the system A. Note that our final expression for $P_{r,s}$ is independent of the choice of the reservoir.

We shall now examine the same problem from the ensemble point of view.

4.2. A system in the grand canonical ensemble

We now visualize an ensemble of $\mathcal{N}$ identical systems (which, of course, can be labelled as $1, 2, \ldots, \mathcal{N}$) mutually sharing a total number of particles[2] $\mathcal{N}\bar{N}$ and a total energy $\mathcal{N}\bar{E}$. Let $n_{r,s}$ denote the number of systems that have, at any time t, the number N_r of particles and the amount E_s of energy ($r, s = 0, 1, 2, \ldots$); then, obviously,

$$\sum_{r,s} n_{r,s} = \mathcal{N}, \tag{1a}$$

$$\sum_{r,s} n_{r,s} N_r = \mathcal{N}\bar{N} \tag{1b}$$

and

$$\sum_{r,s} n_{r,s} E_s = \mathcal{N}\bar{E}. \tag{1c}$$

Any set $\{n_{r,s}\}$, of the numbers $n_{r,s}$, which satisfies the restrictive conditions (1), represents one of the possible modes of distribution of particles and energy among the members of our ensemble. Furthermore, any such mode of distribution can be realized in $W\{n_{r,s}\}$ different ways, where

$$W\{n_{r,s}\} = \frac{\mathcal{N}!}{\prod\limits_{r,s}(n_{r,s}!)}. \tag{2}$$

We may now define the *most probable* mode of distribution, $\{n^*_{r,s}\}$, as the one that maximizes expression (2), satisfying at the same time the restrictive conditions (1). Going through the conventional derivation, see Sec. 3.2, we obtain for a large ensemble

$$\frac{n^*_{r,s}}{\mathcal{N}} = \frac{\exp(-\alpha N_r - \beta E_s)}{\sum\limits_{r,s} \exp(-\alpha N_r - \beta E_s)}; \tag{3}$$

cf. the corresponding eqn. (3.2.10) for the canonical ensemble. Alternatively, we may define the *expectation* (or *mean*) values of the numbers $n_{r,s}$, namely

$$\langle n_{r,s}\rangle = \frac{\sum'_{\{n_{r,s}\}} n_{r,s} W\{n_{r,s}\}}{\sum'_{\{n_{r,s}\}} W\{n_{r,s}\}}, \tag{4}$$

where the primed summations go over all distribution sets that conform to conditions (1). An asymptotic expression for $\langle n_{r,s}\rangle$ can be derived by using the method of Darwin and Fowler–the only difference from the corresponding derivation in Sec. 3.2 being that in the present case we have to work with functions of more than one (complex) variable. The derivation, however, runs along similar lines, with the result

$$\operatorname*{Lim}_{\mathcal{N}\to\infty} \frac{\langle n_{r,s}\rangle}{\mathcal{N}} \simeq \frac{n^*_{r,s}}{\mathcal{N}} = \frac{\exp(-\alpha N_r - \beta E_s)}{\sum_{r,s} \exp(-\alpha N_r - \beta E_s)}, \tag{5}$$

in agreement with eqn. (4.1.9). The parameters α and β, so far undetermined, are eventually determined by the equations

$$\overline{N} = \frac{\sum_{r,s} N_r \exp(-\alpha N_r - \beta E_s)}{\sum_{r,s} \exp(-\alpha N_r - \beta E_s)} \equiv -\frac{\partial}{\partial\alpha}\left\{\ln \sum_{r,s} \exp(-\alpha N_r - \beta E_s)\right\} \tag{6}$$

and

$$\overline{E} = \frac{\sum_{r,s} E_s \exp(-\alpha N_r - \beta E_s)}{\sum_{r,s} \exp(-\alpha N_r - \beta E_s)} \equiv -\frac{\partial}{\partial\beta}\left\{\ln \sum_{r,s} \exp(-\alpha N_r - \beta E_s)\right\}, \tag{7}$$

where the quantities $\overline{N}$ and $\overline{E}$ here are supposed to be preassigned.

4.3. Physical significance of the various statistical quantities

In order to establish a connection between the statistics of the grand canonical ensemble and the thermodynamics of the system under study, we introduce a quantity q, defined by

$$q \equiv \ln\left\{\sum_{r,s} \exp(-\alpha N_r - \beta E_s)\right\}; \tag{1}$$

the quantity q is a function of the parameters α and β and of all the E_s. Taking the differential of q and making use of eqns (4.2.5), (4.2.6) and (4.2.7), we get

$$dq = -\overline{N}\,d\alpha - \overline{E}\,d\beta - \frac{\beta}{\mathcal{N}} \sum_{r,s} \langle n_{r,s}\rangle\, dE_s, \tag{2}$$

so that

$$d(q + \alpha\overline{N} + \beta\overline{E}) = \beta\left(\frac{\alpha}{\beta}\,d\overline{N} + d\overline{E} - \frac{1}{\mathcal{N}} \sum_{r,s} \langle n_{r,s}\rangle\, dE_s\right). \tag{3}$$

To interpret the terms appearing on the right-hand side of this equation, we compare the expression enclosed within the parentheses with the statement of the first law of thermodynamics, viz.

$$\delta Q = d\overline{E} + \delta W - \mu \, d\overline{N}, \tag{4}$$

where the various symbols have their usual meanings. The following correspondence seems inevitable:

$$\delta W = -\frac{1}{\beta} \sum_{r,s} \langle n_{r,s} \rangle \, dE_s, \quad \mu = -\alpha/\beta, \tag{5}$$

with the result that

$$d(q + \alpha\overline{N} + \beta\overline{E}) = \beta \, \delta Q. \tag{6}$$

The parameter β, being the integrating factor for the heat δQ, must be equivalent to the reciprocal of the absolute temperature; so, we may write

$$\beta = 1/kT \tag{7}$$

and, hence,

$$\alpha = -\mu/kT \tag{8}$$

The quantity $(q + \alpha\overline{N} + \beta\overline{E})$ would then be identified with the thermodynamic variable S/k. Accordingly,

$$q = \frac{S}{k} - \alpha\overline{N} - \beta\overline{E} = \frac{TS + \mu\overline{N} - \overline{E}}{kT}. \tag{9}$$

However, $\mu\overline{N}$ is identically equal to G, the Gibbs free energy of the system, and hence to $(\overline{E} - TS + PV)$. So, finally,

$$q \equiv \ln \left\{ \sum_{r,s} \exp(-\alpha N_r - \beta E_s) \right\} = \frac{PV}{kT}. \tag{10}$$

Equation (10) provides the essential link between the thermodynamics of the given system and the statistics of the corresponding grand canonical ensemble. It is, therefore, a relationship of central importance in the formalism developed in this chapter.

To derive further results, we prefer to introduce a parameter z, defined by the relation

$$z \equiv e^{-\alpha} = e^{\mu/kT}; \tag{11}$$

the parameter z is generally referred to as the *fugacity* of the system. In terms of z, the q-potential takes the form

$$q \equiv \ln \left\{ \sum_{r,s} z^{N_r} e^{-\beta E_s} \right\} \tag{12}$$

$$= \ln \left\{ \sum_{N_r=0}^{\infty} z^{N_r} Q_{N_r}(V, T) \right\} \quad \text{(with } Q_0 \equiv 1\text{)}, \tag{13}$$

so that we may write

$$q(z, V, T) \equiv \ln \mathcal{Q}(z, V, T), \tag{14}$$

where

$$\mathcal{Q}(z, V, T) \equiv \sum_{N_r=0}^{\infty} z^{N_r} Q_{N_r}(V, T) \quad (\text{with } Q_0 \equiv 1). \tag{15}$$

Note that, in going from expression (12) to (13), we have (mentally) carried out a summation over the energy values E_s, with N_r kept fixed, thus giving rise to the partition function $Q_{N_r}(V, T)$; of course, the dependence of Q_{N_r} on V comes from the dependence of the E on V. In going from (13) to (14), we have (again mentally) carried out a summation over all the numbers N_r, viz. $0, 1, 2, \ldots, \infty$, thus giving rise to the *grand partition function* $\mathcal{Q}(z, V, T)$ of the system. The q-potential, which we have already identified with PV/kT, is, therefore, the logarithm of the grand partition function.

It appears that in order to evaluate the grand partition function $\mathcal{Q}(z, V, T)$ we have to pass through the routine of evaluating the partition function $Q(N, V, T)$. In principle, this is indeed true. In practice, however, we find that on many occasions an explicit evaluation of the partition function is extremely hard while considerable progress can be made in the evaluation of the grand partition function. This is particularly true when we deal with systems in which the influence of quantum statistics and/or interparticle interactions is important; see Secs 6.2 and 9.1. The formalism of the grand canonical ensemble then proves to be of considerable value.

We are now in a position to write down the full recipe for deriving the leading thermodynamic quantities of a given system from its q-potential. We have first of all, for the pressure of the system,

$$P(z, V, T) = \frac{kT}{V} q(z, V, T) \equiv \frac{kT}{V} \ln \mathcal{Q}(z, V, T). \tag{16}$$

Next, writing N for $\overline{N}$ and U for $\overline{E}$, we obtain with the help of eqns (4.2.6), (4.2.7) and (11)

$$N(z, V, T) = z \left[\frac{\partial}{\partial z} q(z, V, T)\right]_{V,T} = kT \left[\frac{\partial}{\partial \mu} q(\mu, V, T)\right]_{V,T} \tag{17}$$

and

$$U(z, V, T) = -\left[\frac{\partial}{\partial \beta} q(z, V, T)\right]_{z,V} = kT^2 \left[\frac{\partial}{\partial T} q(z, V, T)\right]_{z,V} \tag{18}$$

Eliminating z between eqns (16) and (17), one obtains the equation of state, i.e. the (P, V, T)-relationship, of the system. On the other hand, eliminating z between eqns (17) and (18), one obtains U as a function of N, V and T, which readily leads to the specific heat at constant volume as $(\partial U/\partial T)_{N,V}$. The Helmholtz free energy is given by the formula

$$\begin{aligned} A &= N\mu - PV = NkT \ln z - kT \ln \mathcal{Q}(z, V, T) \\ &= -kT \ln \frac{\mathcal{Q}(z, V, T)}{z^N}, \end{aligned} \tag{19}$$

which may be compared with the canonical ensemble formula $A = -kT \ln Q(N, V, T)$; see also Problem 4.2. Finally, we have for the entropy of the system

$$S = \frac{U - A}{T} = kT \left(\frac{\partial q}{\partial T}\right)_{z,V} - Nk \ln z + kq. \tag{20}$$

4.4. Examples

We shall now study a couple of simple problems, with the explicit purpose of demonstrating how the method of the q-potential works. This is not intended to be a demonstration of the power of this method, for we shall consider here only those problems which can be solved equally well by the methods of the preceding chapters. The real power of the new method will become apparent only when we study problems involving quantum-statistical effects and effects arising from interparticle interactions; many such problems will appear in the remainder of the text.

The first problem we propose to consider here is that of the classical ideal gas. In Sec. 3.5 we showed that the partition function $Q_N(V,T)$ of this system could be written as

$$Q_N(V,T)=\frac{[Q_1(V,T)]^N}{N!}, \tag{1}$$

where $Q_1(V,T)$ may be regarded as the partition function of a single particle in the system. First of all, we should note that eqn. (1) does not imply any restrictions on the particles having *internal* degrees of motion; these degrees of motion, if present, would affect the results only through Q_1. Secondly, we should recall that the factor $N!$ in the denominator arises from the fact that the particles constituting the gas are, in fact, *indistinguishable*. Closely related to the indistinguishability of the particles is the fact that they are *nonlocalized*, for otherwise we could distinguish them through their very sites; compare, for instance, the system of harmonic oscillators, which was studied in Sec. 3.8. Now, since our particles are nonlocalized they can be *anywhere* in the space available to them; consequently, the function Q_1 will be directly proportional to V:

$$Q_1(V,T)=Vf(T). \tag{2}$$

where $f(T)$ is a function of temperature alone. We thus obtain for the grand partition function of the gas

$$\begin{aligned}\mathcal{Q}(z,V,T)&=\sum_{N_r=0}^{\infty} z^{N_r}Q_{N_r}(V,T)=\sum_{N_r=0}^{\infty}\frac{\{zVf(T)\}^{N_r}}{N_r!}\\&=\exp\{zVf(T)\},\end{aligned} \tag{3}$$

whence

$$q(z,V,T)=zVf(T). \tag{4}$$

Formulae (4.3.16)–(4.3.20) then lead to the following results:

$$P=zkTf(T), \tag{5}$$

$$N=zVf(T), \tag{6}$$

$$U=zVkT^2f'(T), \tag{7}$$

$$A=NkT\ln z-zVkTf(T) \tag{8}$$

and

$$S=-Nk\ln z+zVk\{Tf'(T)+f(T)\}. \tag{9}$$

Eliminating' tween (5) and (6), we obtain the equation of state of the system:

$$PV = NkT. \tag{10}$$

We note that eqn. (10) holds irrespective of the form of the function $f(T)$. Next, eliminating z between (6) and (7), we obtain

$$U = NkT^2 f'(T)/f(T), \tag{11}$$

whence

$$C_V = Nk\frac{2Tf(T)f'(T) + T^2\{f(T)f''(T) - [f'(T)]^2\}}{[f(T)]^2}. \tag{12}$$

In simple cases, the function $f(T)$ turns out to be directly proportional to a certain power of T. Supposing that $f(T) \propto T^n$, eqns (11) and (12) become

$$U = n(NkT) \tag{11'}$$

and

$$C_V = n(Nk). \tag{12'}$$

Accordingly, the pressure in such cases is directly proportional to the energy density of the gas, the constant of proportionality being $1/n$. The reader will recall that the case $n = 3/2$ corresponds to a nonrelativistic gas while $n = 3$ corresponds to an extreme relativistic one.

Finally, eliminating z between eqn. (6) and eqns (8,9), we obtain A and S as functions of N, V and T. This essentially completes our study of the classical ideal gas.

The next problem to be considered here is that of a system of independent, *localized* particles — a model which, in some respects, approximates a solid. Mathematically, the problem is similar to that of a system of harmonic oscillators. In either case, the microscopic entities constituting the system are mutually *distinguishable*. The partition function $Q_N(V, T)$ of such a system can be written as

$$Q_N(V, T) = [Q_1(V, T)]^N. \tag{13}$$

At the same time, in view of the localized nature of the particles, the single-particle partition function $Q_1(V, T)$ is essentially independent of the volume occupied by the system. Consequently, we may write

$$Q_1(V, T) = \phi(T), \tag{14}$$

where $\phi(T)$ is a function of temperature alone. We then obtain for the grand partition function of the system

$$\mathcal{Q}(z, V, T) = \sum_{N_r=0}^{\infty} [z\phi(T)]^{N_r} = [1 - z\phi(T)]^{-1}: \tag{15}$$

clearly, the quantity $z\phi(T)$ must stay below unity, so that the summation over N_r is convergent.

The thermodynamics of the system follows straightforwardly from eqn. (15). We have, to begin with,

$$P \equiv \frac{kT}{V}q(z, T) = -\frac{kT}{V}\ln\{1 - z\phi(T)\}. \tag{16}$$

Since both z and T are intensive variables, the right-hand side of (16) vanishes as $V \to \infty$. Hence, in the thermodynamic limit, $P = 0$.[4] For other quantities of interest, we obtain, with the help of eqns (4.3.17)–(4.3.20),

$$N = \frac{z\phi(T)}{1 - z\phi(T)}, \tag{17}$$

$$U = \frac{zkT^2\phi'(T)}{1 - z\phi(T)}, \tag{18}$$

$$A = NkT \ln z + kT \ln \{1 - z\phi(T)\} \tag{19}$$

and

$$S = -Nk \ln z - k \ln \{1 - z\phi(T)\} + \frac{zkT\phi'(T)}{1 - z\phi(T)}. \tag{20}$$

From (17), we get

$$z\phi(T) = \frac{N}{N+1} \simeq 1 - \frac{1}{N} \quad (N \gg 1). \tag{21}$$

It follows that

$$1 - z\phi(T) = \frac{1}{N+1} \simeq \frac{1}{N}. \tag{22}$$

Equations (17)–(20) now give

$$U/N = kT^2\phi'(T)/\phi(T), \tag{18$'$}$$

$$A/N = -kT \ln \phi(T) + O\left(\frac{\ln N}{N}\right) \tag{19$'$}$$

and

$$S/Nk = \ln \phi(T) + T\phi'(T)/\phi(T) + O\left(\frac{\ln N}{N}\right) \tag{20$'$}$$

Substituting

$$\phi(T) = [2 \sinh (\hbar\omega/2kT)]^{-1} \tag{23}$$

into these formulae, we obtain results pertaining to a system of *quantum-mechanical*, one-dimensional harmonic oscillators. The substitution

$$\phi(T) = kT/\hbar\omega, \tag{24}$$

on the other hand, leads to results pertaining to a system of *classical*, one-dimensional harmonic oscillators.

As a corollary, we examine here the problem of the *solid–vapor equilibrium*. Consider a single-component system, having two phases–solid and vapor–in equilibrium, contained in a closed vessel of volume V at temperature T. Since the phases are free to exchange particles, a state of mutual equilibrium would imply that their chemical potentials are equal; this, in turn, means that they have a common fugacity as well. Now, the fugacity z_g of the gaseous phase is given by, see eqn. (6),

$$z_g = \frac{N_g}{V_g f(T)}, \tag{25}$$

where N_g is the number of particles in the gaseous phase and V_g the volume occupied by them; in a typical case, $V_g \simeq V$. The fugacity z_s of the solid phase, on the other hand, is given by eqn. (21):

$$z_s \simeq \frac{1}{\phi(T)}. \tag{26}$$

Equating (25) and (26), we obtain for the *equilibrium particle density* in the vapor phase

$$N_g/V_g = f(T)/\phi(T). \tag{27}$$

Now, if the density in the vapor phase is sufficiently low and the temperature of the system sufficiently high, the vapor pressure P would be given by

$$P = \frac{N_g}{V_g}kT = kT\frac{f(T)}{\phi(T)}. \tag{28}$$

To be specific, we may assume the vapor to be monatomic; the function $f(T)$ is then of the form

$$f(T) = (2\pi mkT)^{3/2}/h^3. \tag{29}$$

On the other hand, if the solid phase can be approximated by a set of three-dimensional harmonic oscillators characterized by a single frequency ω (the *Einstein* model), the function $\phi(T)$ would be

$$\phi(T) = [2\sinh(\hbar\omega/2kT)]^{-3}. \tag{30}$$

However, there is one important difference here. An atom in a solid is energetically more stabilized than an atom which is free–that is why a certain threshold energy is required to transform a solid into separate atoms. Let ε denote the value of this energy per atom, which in a way implies that the zeros of the energy spectra ε_g and ε_s, which led to the functions (29) and (30) respectively, are displaced with respect to one another by an amount ε. A *true* comparison between the functions $f(T)$ and $\phi(T)$ must take this into account. As a result, we obtain for the vapor pressure

$$P = kT\left(\frac{2\pi mkT}{h^2}\right)^{3/2}[2\sinh(\hbar\omega/2kT)]^3 e^{-\varepsilon/kT} \tag{31}$$

In passing, we note that eqn. (27) also gives us the necessary condition for the formation of the solid phase. The condition clearly is:

$$N > V\frac{f(T)}{\phi(T)}, \tag{32}$$

where N is the total number of particles in the system. Alternatively, we require that

$$T < T_c, \tag{33}$$

where T_c is a *characteristic* temperature, determined by the implicit relationship

$$\frac{f(T_c)}{\phi(T_c)} = \frac{N}{V}. \tag{34}$$

Once the two phases appear, the number $N_g(T)$ will have a value determined by eqn. (27) while the remainder, $N - N_g$, will constitute the solid phase.

4.5. Density and energy fluctuations in the grand canonical ensemble: correspondence with other ensembles

In a grand canonical ensemble, the variables N and E, for any member of the ensemble, can lie anywhere between zero and infinity. Therefore, on the face of it, the grand canonical ensemble appears to be very different from its predecessors–the canonical and the microcanonical ensembles. However, as far as thermodynamics is concerned, the results obtained from this ensemble turn out to be identical with the ones obtained from the other two. Thus, in spite of strong facial differences, the overall behavior of a given physical system is practically the same whether it belongs to one kind of ensemble or another. The basic reason for this is that the "relative fluctuations" in the values of the quantities that vary from member to member in an ensemble are practically negligible. Therefore, in spite of the different surroundings which different ensembles provide to a given physical system, the overall behavior of the system is not significantly affected.

To appreciate this point, we shall evaluate the relative fluctuations in the particle density n and the energy E of a given physical system in the grand canonical ensemble. Recalling that

$$\overline{N} = \frac{\sum\limits_{r,s} N_r e^{-\alpha N_r - \beta E_s}}{\sum\limits_{r,s} e^{-\alpha N_r - \beta E_s}}, \tag{1}$$

it readily follows that

$$\left(\frac{\partial \overline{N}}{\partial \alpha}\right)_{\beta, E_s} = -\overline{N^2} + \overline{N}^2. \tag{2}$$

Thus

$$\overline{(\Delta N)^2} \equiv \overline{N^2} - \overline{N}^2 = -\left(\frac{\partial \overline{N}}{\partial \alpha}\right)_{T,V} = kT\left(\frac{\partial \overline{N}}{\partial \mu}\right)_{T,V} \tag{3}$$

From (3), we obtain for the relative mean-square fluctuation in the particle density $n(= N/V)$

$$\frac{\overline{(\Delta n)^2}}{\overline{n}^2} = \frac{\overline{(\Delta N)^2}}{\overline{N}^2} = \frac{kT}{\overline{N}^2}\left(\frac{\partial \overline{N}}{\partial \mu}\right)_{T,V} \tag{4}$$

In terms of the variable v – $(= V/\overline{N})$, we may write

$$\frac{\overline{(\Delta n)^2}}{\overline{n}^2} = \frac{kTv^2}{V^2}\left(\frac{\partial (V/v)}{\partial \mu}\right)_{T,V} = -\frac{kT}{V}\left(\frac{\partial v}{\partial \mu}\right)_T \tag{5}$$

To put this result into a more practical form, we recall the thermodynamic relationship,

$$d\mu = v\,dP - s\,dT, \tag{6}$$

according t[illegible]ich $d\mu$ (at constant T) = $v\,dP$. Equation (5) then takes the form

$$\frac{\overline{(\Delta n)^2}}{\bar{n}^2} = -\frac{kT}{V}\frac{1}{v}\left(\frac{\partial v}{\partial P}\right)_T = \frac{kT}{V}\kappa_T, \tag{7}$$

where κ_T is the isothermal compressibility of the system.

Thus, the relative root-mean-square fluctuation in the particle density of the given system is *ordinarily* $O(N^{-1/2})$ and, hence, negligible. However, there are exceptions, like the ones met with in situations accompanying *phase transitions*. In those situations, the compressibility of a given system can become excessively large, as is evidenced by an almost "flattening" of the isotherms. Under these circumstances, the derivative $(\partial v/\partial P)_T$, and hence the quantity κ_T, can very well be $O(N)$; the relative root-mean-square fluctuation in the particle density n can then be $O(1)$. Thus, in the region of phase transitions, especially at the critical points, we expect to encounter unusually large fluctuations in the particle density of the system. Such fluctuations indeed exist and account for phenomena like *critical opalescence*. It is clear that under these circumstances the formalism of the grand canonical ensemble could, in principle, lead to results which are not necessarily identical with the ones following from the corresponding canonical ensemble. In such cases, it is the formalism of the grand canonical ensemble which will have to be preferred because only this one will provide a correct picture of the actual physical situation.

We shall now examine fluctuations in the energy of the system. Following the usual procedure, we obtain

$$\overline{(\Delta E)^2} \equiv \overline{E^2} - \bar{E}^2 = -\left(\frac{\partial \bar{E}}{\partial \beta}\right)_{z,V} = kT^2\left(\frac{\partial U}{\partial T}\right)_{z,V} \tag{8}$$

To put expression (8) into a more comprehensible form, we write

$$\left(\frac{\partial U}{\partial T}\right)_{z,V} = \left(\frac{\partial U}{\partial T}\right)_{N,V} + \left(\frac{\partial U}{\partial N}\right)_{T,V}\left(\frac{\partial N}{\partial T}\right)_{z,V} \tag{9}$$

where the symbol N is being used interchangeably for $\bar{N}$. Now, in view of the fact that

$$N = -\left(\frac{\partial}{\partial \alpha}\ln \mathcal{Q}\right)_{\beta,V}, \quad U = -\left(\frac{\partial}{\partial \beta}\ln \mathcal{Q}\right)_{\alpha,V}, \tag{10}$$

we have

$$\left(\frac{\partial N}{\partial \beta}\right)_{\alpha,V} = \left(\frac{\partial U}{\partial \alpha}\right)_{\beta,V} \tag{11}$$

and, hence,

$$\left(\frac{\partial N}{\partial T}\right)_{z,V} = \frac{1}{T}\left(\frac{\partial U}{\partial \mu}\right)_{T,V} \tag{12}$$

Substituting expressions (9) and (12) into (8) and remembering that the quantity $(\partial U/\partial T)_{N,V}$ is the familiar C_V, we get

$$\overline{(\Delta E)^2} = kT^2 C_V + kT\left(\frac{\partial U}{\partial N}\right)_{T,V}\left(\frac{\partial U}{\partial \mu}\right)_{T,V}. \tag{13}$$

Invoking eqns (3.6.3) and (3), we finally obtain

$$\overline{(\Delta E)^2} = \langle(\Delta E)^2\rangle_{\text{can}} + \left\{\left(\frac{\partial U}{\partial N}\right)_{T,V}\right\}^2 \overline{(\Delta N)^2}. \tag{14}$$

Formula (14) is highly instructive; it tells us that the mean-square fluctuation in the energy E of a system in the grand canonical ensemble is equal to the value it would have in the canonical ensemble *plus* a contribution arising from the fact that now the particle number N is also fluctuating. Again, under ordinary circumstances, the relative root-mean-square fluctuation in the energy density of the system would be practically negligible. However, in the region of phase transitions, unusually large fluctuations in the values of this variable can arise by virtue of the second term in the formula.

Problems

4.1. Show that the entropy of a system in the grand canonical ensemble can be written as

$$S = -k \sum_{r,s} P_{r,s} \ln P_{r,s},$$

where $P_{r,s}$ is given by eqn. (4.1.9).

4.2. "In the thermodynamic limit (when the extensive properties of the system become infinitely large, while the intensive ones remain constant), the q-potential of the system may be calculated by taking only the largest term in the sum

$$\sum_{N_r=0}^{\infty} z^{N_r} Q_{N_r}(V,T)."$$

Verify this statement and interpret the result physically.

4.3. A vessel of volume $V^{(0)}$ contains $N^{(0)}$ molecules. Assuming that there is no correlation whatsoever between the locations of the various molecules, calculate the probability, $p(N, V)$, that a region of volume V (located anywhere in the vessel) contains exactly N molecules.

- (i) Show that $\overline{N} = N^{(0)}p$ and $(\Delta N)_{\text{r.m.s.}} = \{N^{(0)}p(1-p)\}^{1/2}$, where $p = V/V^{(0)}$.
- (ii) Show that if both $N^{(0)}p^{(0)}$ and $N(1-p)$ are large numbers, the function $p(N, V)$ assumes a Gaussian form.
- (iii) Further, if $p \ll 1$ and $N \gg N^{(0)}$, show that the function $p(N, V)$ assumes the form of a Poisson distribution:

$$p(N) = e^{-\overline{N}} \frac{(\overline{N})^N}{N!}.$$

4.4. "The probability that a system in the grand canonical ensemble has *exactly* N particles is given by

$$p(N) = \frac{z^N Q_N(V,T)}{\mathcal{Q}(z,V,T)}."$$

Verify this statement and show that in the case of a classical, ideal gas the distribution of particles among the members of a grand canonical ensemble is identically a Poisson distribution. Calculate the root-mean-square value of (ΔN) for this system both from the general formula (4.5.3) and from the Poisson distribution, and show that the two results are the same.

4.5. Show that expression (4.3.20) for the entropy of a system in the grand canonical ensemble can also be written as

$$S = k\left[\frac{\partial}{\partial T}(Tq)\right]_{\mu,V}.$$

4.6. Define the function $Y(N, \gamma, T)$ as

$$Y(N, \gamma, T) = \int_0^\infty Q(N, V, T) e^{\gamma V}\, dV.$$

What physical meaning must be given to the parameter γ so that the function Y assumes some direct relevance to thermodynamics? Set up a scheme to derive the various thermodynamic quantities of a given physical system from the function Y, and illustrate it by considering a few simple examples.

4.7. Consider a classical system of non-interacting, diatomic molecules enclosed in a box of volume V at temperature T. The Hamiltonian of a single molecule is given by

$$H(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{p}_1, \boldsymbol{p}_2) = \frac{1}{2m}(p_1^2 + p_2^2) + \frac{1}{2}K|\boldsymbol{r}_1 - \boldsymbol{r}_2|^2.$$

Study the thermodynamics of this system, including the dependence of the quantity $\langle r_{12}^2 \rangle$ on T.

4.8. Determine the grand partition function of a gaseous system of "magnetic" atoms (with $J = \frac{1}{2}$ and $g = 2$) which can have, in addition to the kinetic energy, a magnetic potential energy equal to $\mu_B H$ or $-\mu_B H$, depending upon their orientation with respect to an applied magnetic field H. Derive an expression for the magnetization of the system, and calculate how much heat will be given off by the system when the magnetic field is reduced from H to zero at constant volume and constant temperature.

4.9. Study the problem of the solid–vapor equilibrium (Sec. 4.4) by setting up the grand partition function of the system.

4.10. A surface with N_0 adsorption centers has $N(\le N_0)$ gas molecules adsorbed on it. Show that the chemical potential of the adsorbed molecules is given by

$$\mu = kT \ln \frac{N}{(N_0 - N)a(T)},$$

where $a(T)$ is the partition function of a single adsorbed molecule. Solve the problem by constructing the grand partition function as well as the partition function of the system.

[Neglect the intermolecular interaction among the adsorbed molecules.]

4.11. Study the state of equilibrium between a gaseous phase and an adsorbed phase in a single-component system. Show that the pressure in the gaseous phase is given by the Langmuir equation

$$P_g = \frac{\theta}{1-\theta} \times \text{(a certain function of temperature)},$$

where θ is the equilibrium fraction of the adsorption sites that *are* occupied by the adsorbed molecules.

4.12. Show that for a system in the grand canonical ensemble

$$\{\overline{(NE)} - \overline{N}\,\overline{E}\} = \left(\frac{\partial U}{\partial N}\right)_{T,V} \overline{(\Delta N)^2}.$$

4.13. Define a quantity J as

$$J = E - N\mu = TS - PV.$$

Show that for a system in the grand canonical ensemble

$$\overline{(\Delta J)^2} = kT^2 C_V + \left\{\left(\frac{\partial U}{\partial N}\right)_{T,V} - \mu\right\}^2 \overline{(\Delta N)^2}.$$

Notes

[1] Note that A here could be a relatively small "part" of a given system $A^{(0)}$, while A' represents the "rest" of $A^{(0)}$. That gives a truly practical perspective to the grand canonical formalism.

[2] For simplicity, we shall henceforth use the symbols $\overline{N}$ and $\overline{E}$ instead of $\langle N \rangle$ and $\langle E \rangle$.

[3] This quantity [illegible] first introduced by Kramers, who called it the *q-potential*.

[4] It will be see [illegible] the sequel that P actually vanishes like $(\ln N)/N$.

CHAPTER 5

FORMULATION OF QUANTUM STATISTICS

THE scope of the ensemble theory developed in Chapters 2–4 is extremely general, though the applications considered so far were confined either to classical systems or to quantum-mechanical systems composed of *distinguishable* entities. When it comes to quantum-mechanical systems composed of *indistinguishable* entities, as most physical systems are, considerations of the preceding chapters have to be applied with care. One finds that in this case it is advisable to rewrite ensemble theory in a language that is more natural to a quantum-mechanical treatment, namely the language of the operators and the wave functions. Insofar as statistics are concerned, this rewriting of the theory may not seem to introduce any new physical ideas as such; nonetheless, it provides us with a tool which is highly suited for studying typical quantum systems. And once we set out to study these systems in detail, we encounter a stream of new, and altogether different, physical concepts. In particular, we find that the behavior of even a non-interacting system, such as the ideal gas, departs considerably from the pattern set by the classical treatment. In the presence of interactions, the pattern becomes even more complicated. Of course, in the limit of high temperatures and low densities, the behavior of all physical systems tends *asymptotically* to what we expect on classical grounds. In the process of demonstrating this point, we automatically obtain a criterion which tells us whether a given physical system may or may not be treated classically. At the same time, we obtain rigorous evidence in support of the procedure, employed in the previous chapters, for computing the number, Γ, of microstates (corresponding to a given macrostate) of a given system from the volume, ω, of the relevant region of its phase space, viz. $\Gamma \approx \omega/h^f$, where f is the number of "degrees of freedom" in the problem.

5.1. Quantum-mechanical ensemble theory: the density matrix

We consider an ensemble of $\mathcal{N}$ identical systems, where $\mathcal{N} \gg 1$. These systems are characterized by a (common) Hamiltonian, which may be denoted by the operator $\hat{H}$. At time t, the physical states of the various systems in the ensemble will be characterized by the wave functions $\psi(r_i, t)$, where r_i denote the position coordinates relevant to the system under study. Let $\psi^k(r_i, t)$ denote the (normalized) wave function characterizing the physical state in which the kth system of the ensemble happens to be at time t; naturally, $k = 1, 2, \ldots, \mathcal{N}$. The time variation of the function $\psi^k(t)$ will be determined by the Schrödinger equation[1]

$$\hat{H}\psi^k(t) = i\hbar\dot{\psi}^k(t). \tag{1}$$

Introducing a complete set of orthonormal functions ϕ_n, the wave functions $\psi^k(t)$ may be written as

$$\psi^k(t) = \sum_n a_n^k(t)\phi_n, \tag{2}$$

where

$$a_n^k(t) = \int \phi_n^* \psi^k(t)\, d\tau; \tag{3}$$

here, ϕ_n^* denotes the complex conjugate of ϕ_n while $d\tau$ denotes the volume element of the coordinate space of the given system. Clearly, the physical state of the kth system can be described equally well in terms of the coefficients $a_n^k(t)$. The time variation of these coefficients will be given by

$$\begin{aligned} i\hbar\dot{a}_n^k(t) &= i\hbar \int \phi_n^* \dot{\psi}^k(t)\, d\tau = \int \phi_n^* \hat{H} \psi^k(t)\, d\tau \\ &= \int \phi_n^* \hat{H} \left\{ \sum_m a_m^k(t)\phi_m \right\} d\tau \\ &= \sum_m H_{nm} a_m^k(t), \end{aligned} \tag{4}$$

where

$$H_{nm} = \int \phi_n^* \hat{H} \phi_m\, d\tau. \tag{5}$$

The physical significance of the coefficients $a_n^k(t)$ is evident from eqn. (2). They are the *probability amplitudes* for the various systems of the ensemble to be in the various states ϕ_n; to be practical, the number $|a_n^k(t)|^2$ represents the probability that a measurement at time t finds the kth system of the ensemble to be in the particular state ϕ_n. Clearly, we must have

$$\sum_n |a_n^k(t)|^2 = 1 \quad \text{(for all } k\text{)}. \tag{6}$$

We now introduce the *density operator* $\hat{\rho}(t)$, as defined by the matrix elements

$$\rho_{mn}(t) = \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} \{a_m^k(t) a_n^{k*}(t)\}; \tag{7}$$

Clearly, the matrix element $\rho_{mn}(t)$ is the ensemble average of the quantity $a_m(t)a_n^*(t)$ which, as a rule, varies from member to member in the ensemble. In particular, the diagonal element $\rho_{nn}(t)$ is the ensemble average of the probability $|a_n(t)|^2$, the latter itself being a (quantum-mechanical) average. Thus, we encounter here a double-averaging process–once due to the probabilistic aspect of the wave functions and again due to the statistical aspect of the ensemble. The quantity $\rho_{nn}(t)$ now represents the probability that a system, chosen *at random* from the ensemble, at time t, is found to be in the state ϕ_n. In view of eqns (6) and (7),

$$\sum_n \rho_{nn} = 1. \tag{8}$$

We shall now determine the equation of motion for the density matrix $\rho_{mn}(t)$. We obtain, with the help of the foregoing equations,

$$
\begin{aligned}
i\hbar\dot{\rho}_{mn}(t) &= \frac{1}{\mathcal{N}}\sum_{k=1}^{\mathcal{N}}\left[i\hbar\left\{\dot{a}_m^k(t)a_n^{k*}(t) + a_m^k(t)\dot{a}_n^{k*}(t)\right\}\right]\\
&= \frac{1}{\mathcal{N}}\sum_{k=1}^{\mathcal{N}}\left[\left\{\sum_l H_{ml}a_l^k(t)\right\}a_n^{k*}(t) - a_m^k(t)\left\{\sum_l H_{nl}^{*}a_l^{k*}(t)\right\}\right]\\
&= \sum_l\left\{H_{ml}\rho_{ln}(t) - \rho_{ml}(t)H_{ln}\right\}\\
&= (\hat{H}\hat{\rho} - \hat{\rho}\hat{H})_{mn};
\end{aligned}
\tag{9}
$$

here, use has been made of the fact that, in view of the Hermitian character of the operator $\hat{H}$, $H_{nl}^{*} = H_{ln}$. Using the commutator notation, eqn. (9) may be written as

$$i\hbar\dot{\hat{\rho}} = [\hat{H}, \hat{\rho}]_{-} \tag{10}$$

Equation (10) is the quantum-mechanical analogue of the classical equation (2.2.10) of Liouville. As expected in going from a classical equation of motion to its quantum-mechanical counterpart, the Poisson bracket $[\rho, H]$ has given place to the commutator $(\hat{\rho}\hat{H} - \hat{H}\hat{\rho})/i\hbar$.

If the given system is known to be in a state of equilibrium, the corresponding ensemble must be *stationary*, i.e. $\dot{\rho}_{mn} = 0$. Equations (9) and (10) then tell us that, for this to be the case, (i) the density operator $\hat{\rho}$ must be an explicit function of the Hamiltonian operator $\hat{H}$ (for then the two operators will necessarily commute) and (ii) the Hamiltonian must not depend explicitly on time, i.e., we must have (i) $\hat{\rho} = \hat{\rho}(\hat{H})$ and (ii) $\dot{\hat{H}} = 0$. Now, if the basis functions ϕ_n were the eigenfunctions of the Hamiltonian itself, then the matrices H and ρ would be diagonal:

$$H_{mn} = E_n\delta_{mn}, \qquad \rho_{mn} = \rho_n\delta_{mn}. \tag{11^2}$$

The diagonal element ρ_n, being a measure of the probability that a system, chosen *at random* (and at *any* time) from the ensemble, is found to be in the eigenstate ϕ_n, will naturally depend upon the corresponding eigenvalue E_n of the Hamiltonian; the precise nature of this dependence is, however, determined by the "kind" of ensemble we wish to construct.

In any other representation, the density matrix may or may not be diagonal. However, quite generally, it will be symmetric:

$$\rho_{mn} = \rho_{nm}. \tag{13}$$

The physical reason for this symmetry is that, in statistical equilibrium, the tendency of a physical system to switch from one state (in the new representation) to another must be counterbalanced by an equally strong tendency to switch between the same states in the reverse direction. This condition of *detailed balancing* is essential for the maintenance of an equilibrium distribution within the ensemble.

Finally, we consider the expectation value of a physical quantity G, which is dynamically represented by an operator $\hat{G}$. This will be given by

$$\langle G\rangle = \frac{1}{\mathcal{N}}\sum_{k=1}^{\mathcal{N}}\int \psi^{k*}\hat{G}\psi^{k}\,d\tau. \tag{14}$$

In terms of the coefficients a_n^k,

$$\langle G\rangle = \frac{1}{\mathcal{N}}\sum_{k=1}^{\mathcal{N}}\left[\sum_{m,n} a_n^{k*}a_m^k G_{nm}\right], \tag{15}$$

where

$$G_{nm} = \int \phi_n^* \hat{G}\phi_m\,d\tau. \tag{16}$$

Introducing the density matrix ρ, eqn. (15) becomes

$$\langle G\rangle = \sum_{m,n}\rho_{mn}G_{nm} = \sum_{m}(\hat{\rho}\hat{G})_{mm} = \mathrm{Tr}\,(\hat{\rho}\hat{G}). \tag{17}$$

Taking $\hat{G} = \hat{1}$, where $\hat{1}$ is the unit operator, we have

$$\mathrm{Tr}\,(\hat{\rho}) = 1, \tag{18}$$

which is identical with (8). It should be noted here that if the original wave functions ψ^k were not normalized then the expectation value $\langle G\rangle$ would be given by the formula

$$\langle G\rangle = \frac{\mathrm{Tr}\,(\hat{\rho}\hat{G})}{\mathrm{Tr}\,(\hat{\rho})} \tag{19}$$

instead. In view of the mathematical structure of the formulae (17) and (19), the expectation value of any given physical quantity G is *manifestly* independent of the choice of the basis $\{\phi_n\}$, as it indeed should be.

5.2. Statistics of the various ensembles

A. *The microcanonical ensemble*

The construction of the microcanonical ensemble is based on the premise that the systems constituting the ensemble are characterized by a fixed number of particles N, a fixed volume V and an energy lying within the interval $(E - \frac{1}{2}\Delta, E + \frac{1}{2}\Delta)$, where $\Delta \ll E$. The total number of distinct microstates accessible to a system is then denoted by the symbol $\Gamma(N, V, E; \Delta)$ and, by assumption, any one of these microstates is as likely to occur as any other. This assumption enters into our theory in the nature of a postulate, and is often referred to as the postulate of *equal a priori probabilities* for the various accessible states. Accordingly, the density matrix (which, in the energy representation, must be a diagonal matrix) will be of the n

$$\rho_{mn} = \rho_n\delta_{mn}, \tag{1}$$

with

$$\rho_n = \begin{cases} 1/\Gamma & \text{for each of the accessible states,} \\ 0 & \text{for all other states;} \end{cases} \tag{2}$$

the normalization condition (5.1.18) is clearly satisfied. As we already know, the thermodynamics of the system is completely determined from the expression for its entropy which, in turn, is given by

$$S = k \ln \Gamma. \tag{3}$$

Since Γ, the total number of distinct, accessible states, is supposed to be computed quantum-mechanically, taking due account of the indistinguishability of the particles right from the beginning, no paradox, such as Gibbs', is now expected to arise. Moreover, if the quantum state of the system turns out to be unique ($\Gamma = 1$), the entropy of the system would identically vanish. This provides us with a sound theoretical basis for the hitherto empirical theorem of Nernst (also known as the *third law of thermodynamics*).

The situation corresponding to the case $\Gamma = 1$ is usually referred to as a *pure* case. In such a case, the construction of an ensemble is essentially superfluous, because every system in the ensemble has got to be in one and the same state. Accordingly, there is only one diagonal element ρ_{nn} which is nonzero (actually equal to unity), while all others are zero. The density matrix, therefore, satisfies the condition

$$\rho^2 = \rho. \tag{4}$$

In a different representation, the pure case will correspond to

$$\rho_{mn} = \frac{1}{\Gamma} \sum_{k=1}^{\Gamma} a_m^k a_n^{k*} = a_m a_n^* \tag{5}$$

because all values of k are now literally equivalent. We then have

$$\begin{aligned} \rho^2_{mn} &= \sum_l \rho_{ml}\rho_{ln} = \sum_l a_m a_l^* a_l a_n^* \\ &= a_m a_n^* \quad \left(\text{because } \sum_l a_l^* a_l = 1\right) \\ &= \rho_{mn}. \end{aligned} \tag{6}$$

Condition (4) thus holds in all representations.

A situation in which $\Gamma > 1$ is usually referred to as a *mixed* case. The density matrix, in the energy representation, is then given by eqns (1) and (2). If we now change over to any other representation, the general form of the density matrix should remain the same, namely (i) the off-diagonal elements should continue to be zero, while (ii) the diagonal elements (over the allowed range) should continue to be equal to one another. Now, had we constructed our ensemble on a representation other than the energy representation right from the beginning, how could we have possibly anticipated *ab initio* property (i) of the density matrix, though property (ii) could have been easily invoked through a postulate of *equal a priori probabilities*? To ensure that property (i), as well as property (ii), holds in *every* representation, we must invoke yet another postulate, viz. the postulate of *random a priori phases* for the probability amplitudes a_n^k, which in turn implies that the wave function ψ^k, for all k, is an *incoherent* superposition of the basis $\{\phi_n\}$. As a consequence of this postulate, coupled with the postulate of equal *a priori* probabilities, we would have in any representation

$$\begin{aligned}\rho_{mn} \equiv \frac{1}{\mathcal{N}}\sum_{k=1}^{\mathcal{N}} a_m^k a_n^{k*} &= \frac{1}{\mathcal{N}}\sum_{k=1}^{\mathcal{N}} |a|^2 e^{i(\theta_m^k-\theta_n^k)}\\ &= c\langle e^{i(\theta_m^k-\theta_n^k)}\rangle\\ &= c\delta_{mn},\end{aligned}\tag{7}$$

as it should be for a microcanonical ensemble.

Thus, contrary to what might have been expected on customary thought, to secure the physical situation corresponding to a microcanonical ensemble, we require in general two postulates instead of one! The second postulate arises solely from quantum-mechanics and is intended to ensure noninterference (and hence a complete absence of correlations) among the member systems; this, in turn, enables us to form a mental picture of each system, one at a time, completely disentangled from other systems in the ensemble.

B. The canonical ensemble

In this ensemble the macrostate of a member system is defined through the parameters N, V and T; the energy E is now a variable quantity. The probability that a system, chosen *at random* from the ensemble, possesses an energy E_r is determined by the Boltzmann factor $\exp(-\beta E_r)$, where $\beta = 1/kT$; see Secs 3.1 and 3.2. The density matrix in the energy representation is, therefore, taken as

$$\rho_{mn} = \rho_n \delta_{mn},\tag{8}$$

with

$$\rho_n = C\exp(-\beta E_n); \quad n = 0, 1, 2, \ldots\tag{9}$$

The constant C is determined by the normalization condition (5.1.18), whence

$$C = \frac{1}{\sum_n \exp(-\beta E_n)} = \frac{1}{Q_N(\beta)},\tag{10}$$

where $Q_N(\beta)$ is the *partition function* of the system. In view of eqn. (5.1.12), see Note 2 the density operator in this ensemble may be written as

$$\begin{aligned}\hat{\rho} &= \sum_n |\phi_n\rangle \frac{1}{Q_N(\beta)} e^{-\beta E_n} \langle\phi_n|\\ &= \frac{1}{Q_N(\beta)} e^{-\beta\hat{H}} \sum_n |\phi_n\rangle\langle\phi_n|\\ &= \frac{1}{Q_N(\beta)} e^{-\beta\hat{H}} = \frac{e^{-\beta\hat{H}}}{\mathrm{Tr}\,(e^{-\beta\hat{H}})},\end{aligned}\tag{11}$$

for the operator $\sum_n |\phi_n\rangle\langle\phi_n|$ is identically the unit operator. It is understood th the operator $\exp(-\beta\hat{H})$ in eqn. (11) stands for the sum

$$\sum_{j=0}^{\infty}(-1)^j\frac{(\beta\hat{H})^j}{j!}. \qquad (1$$

The expectation value $\langle G\rangle_N$ of a physical quantity G, which is represented by · operator $\hat{G}$, is now given by

$$\begin{aligned}\langle G\rangle_N = \mathrm{Tr}\,(\hat{\rho}\hat{G}) &= \frac{1}{Q_N(\beta)}\mathrm{Tr}\,(\hat{G}e^{-\beta\hat{H}})\\ &= \frac{\mathrm{Tr}\,(\hat{G}e^{-\beta\hat{H}})}{\mathrm{Tr}\,(e^{-\beta\hat{H}})};\end{aligned} \qquad (13$$

the suffix N here emphasizes the fact that the averaging is being done over · ensemble *with N fixed.*

C. The grand canonical ensemble

In this ensemble the density operator $\hat{\rho}$ operates on a Hilbert space with indefinite number of particles. The density operator must therefore commute n only with the Hamiltonian operator $\hat{H}$ but also with a number operator $\hat{n}$ who eigenvalues are $0, 1, 2, \ldots$. The precise form of the density operator can now b obtained by a straightforward generalization of the preceding case, with the res

$$\hat{\rho} \propto e^{-\beta\hat{H}-\alpha\hat{n}} = \frac{1}{\mathcal{Q}(\mu, V, T)}e^{-\beta(\hat{H}-\mu\hat{n})}, \qquad (1$$

where

$$\mathcal{Q}(\mu, V, T) = \sum_{r,s} e^{-\beta(E_r-\mu N_s)} = \mathrm{Tr}\left\{e^{-\beta(\hat{H}-\mu\hat{n})}\right\}. \qquad (1$$

The ensemble average $\langle G\rangle$ is now given by

$$\begin{aligned}\langle G\rangle &= \frac{1}{\mathcal{Q}(\mu, V, T)}\mathrm{Tr}\left(\hat{G}e^{-\beta\hat{H}}e^{\beta\mu\hat{n}}\right)\\ &= \frac{\sum_{N=0}^{\infty} z^N\langle G\rangle_N Q_N(\beta)}{\sum_{N=0}^{\infty} z^N Q_N(\beta)},\end{aligned} \qquad (1$$

where $z\,(\equiv e^{\beta\mu})$ is the *fugacity* of the system while $\langle G\rangle_N$ is the canonical-ensem average, as given by eqn. (13). The quantity $\mathcal{Q}(\mu, V, T)$ appearing in th formulae is, clearly, the *grand partition function* of the system.

5.3. Examples

A. An electron in a magnetic field

Let us consider, for illustration, the case of a single electron which possesses an intrinsic spin $\frac{1}{2}\hbar\hat{\sigma}$ and a magnetic moment μ_B, where $\hat{\sigma}$ is the Pauli spin operator and $\mu_B = e\hbar/2mc$. The spin of the electron can have two possible orientations, ↑ or ↓, with respect to an applied magnetic field $\boldsymbol{B}$. If the applied field is taken to be in the direction of the z-axis, the configurational Hamiltonian of the spin takes the form

$$\hat{H} = -\mu_B(\hat{\sigma}\cdot\boldsymbol{B}) = -\mu_B B\hat{\sigma}_z. \tag{1}$$

In the representation that makes $\hat{\sigma}_z$ diagonal, namely

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2}$$

the density matrix in the canonical ensemble would be

$$\begin{aligned}(\hat{\rho}) &= \frac{(e^{-\beta\hat{H}})}{\mathrm{Tr}\,(e^{-\beta\hat{H}})} \\ &= \frac{1}{e^{\beta\mu_B B} + e^{-\beta\mu_B B}}\begin{pmatrix} e^{\beta\mu_B B} & 0 \\ 0 & e^{-\beta\mu_B B} \end{pmatrix}\end{aligned} \tag{3}$$

We thus obtain for the expectation value σ_z

$$\langle\sigma_z\rangle = \mathrm{Tr}\,(\hat{\rho}\hat{\sigma}_z) = \frac{e^{\beta\mu_B B} - e^{-\beta\mu_B B}}{e^{\beta\mu_B B} + e^{-\beta\mu_B B}} = \tanh\,(\beta\mu_B B), \tag{4}$$

in agreement with the findings of Secs 3.9 and 3.10.

B. A free particle in a box

We now consider the case of a free particle, of mass m, in a cubical box of side L. The Hamiltonian of the particle is given by

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right), \tag{5}$$

while the eigenfunctions of the Hamiltonian, which satisfy periodic boundary conditions

$$\begin{aligned}\phi(x+L, y, z) &= \phi(x, y+L, z) = \phi(x, y, z+L) \\ &= \phi(x, y, z),\end{aligned} \tag{6}$$

are given by

$$\phi_E(\boldsymbol{r}) = \frac{1}{L^{3/2}}\exp\,(i\boldsymbol{k}\cdot\boldsymbol{r}), \tag{7}$$

the corresponding eigenvalues E being

$$E = \frac{\hbar^2 k^2}{2m}, \tag{8}$$

with

$$k \equiv (k_x, k_y, k_z) = \frac{2\pi}{L}(n_x, n_y, n_z); \tag{9}$$

the quantum numbers n_x, n_y and n_z must be integers (positive, negative or zero). Symbolically, we may write for the wave vector $\boldsymbol{k}$

$$\boldsymbol{k} = \frac{2\pi}{L}\boldsymbol{n}, \tag{10}$$

where $\boldsymbol{n}$ is a vector with integral components 0, ±1, ±2, ...

We now proceed to evaluate the density matrix ($\hat{\rho}$) of this system in the canonical ensemble; we shall do so in the *coordinate* representation. In view of eqn. (5.2.11), we have

$$\begin{aligned}\langle \boldsymbol{r}|e^{-\beta\hat{H}}|\boldsymbol{r}'\rangle &= \sum_E \langle \boldsymbol{r}|E\rangle e^{-\beta E}\langle E|\boldsymbol{r}'\rangle \\ &= \sum_E e^{-\beta E}\phi_E(\boldsymbol{r})\phi_E^*(\boldsymbol{r}'). \end{aligned} \tag{11}$$

Substituting from (7) and making use of (8) and (10), we obtain

$$\begin{aligned}\langle \boldsymbol{r}|e^{-\beta\hat{H}}|\boldsymbol{r}'\rangle &= \frac{1}{L^3}\sum_{\boldsymbol{k}} \exp\left[-\frac{\beta\hbar^2}{2m}k^2 + i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')\right] \\ &\approx \frac{1}{(2\pi)^3}\int \exp\left[-\frac{\beta\hbar^2}{2m}k^2 + i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')\right] d^3k \\ &= \left(\frac{m}{2\pi\beta\hbar^2}\right)^{3/2} \exp\left[-\frac{m}{2\beta\hbar^2}|\boldsymbol{r}-\boldsymbol{r}'|^2\right]; \end{aligned} \tag{12}$$

see eqns (B.41, 42). It follows that

$$\begin{aligned}\mathrm{Tr}\left(e^{-\beta\hat{H}}\right) &= \int \langle \boldsymbol{r}|e^{-\beta\hat{H}}|\boldsymbol{r}\rangle\, d^3r \\ &= V\left(\frac{m}{2\pi\beta\hbar^2}\right)^{3/2}. \end{aligned} \tag{13}$$

Expression (13) is indeed the partition function, $Q_1(\beta)$, of a single particle confined to a box of volume V; see eqn. (3.5.19). Combining (12) and (13), we obtain for the density matrix in the coordinate representation

$$\langle \boldsymbol{r}|\hat{\rho}|\boldsymbol{r}'\rangle = \frac{1}{V}\exp\left[-\frac{m}{2\beta\hbar^2}|\boldsymbol{r}-\boldsymbol{r}'|^2\right]. \tag{14}$$

As expected, the matrix $\rho_{\boldsymbol{r},\boldsymbol{r}'}$ is symmetric between the states $\boldsymbol{r}$ and $\boldsymbol{r}'$. Moreover, the diagonal element $\langle \boldsymbol{r}|\rho|\boldsymbol{r}\rangle$, which represents the *probability density* for the particle to be in the neighborhood of the point $\boldsymbol{r}$, is independent of $\boldsymbol{r}$; this means that, in the case of a single free particle, *all* positions within the box are equally likely to obtain. A nondiagonal element $\langle \boldsymbol{r}|\rho|\boldsymbol{r}'\rangle$, on the other hand, is a measure of the probability of "spontaneous transition" between the position coordinates $\boldsymbol{r}$ and $\boldsymbol{r}'$ and is therefore a measure of the relative "intensity" of the wave packet (associated with the particle) at a distance $|\boldsymbol{r} - \boldsymbol{r}'|$ from the centre of the packet. The spatial extent of the wave packet, which in turn is a measure of the uncertainty involved in locating the position of the particle, is clearly of order $\hbar/(mkT)^{1/2}$; the latter is also a measure of the *mean thermal wavelength* of the particle. The spatial spread found here is a purely quantum-mechanical effect; quite expectedly, it tends to vanish at high temperatures. In fact, as $\beta \to 0$, the behavior of the matrix element (14) approaches that of a delta function, which implies a return to the classical picture of a *point* particle.

Finally, we determine the expectation value of the Hamiltonian itself. From eqns (5) and (14), we obtain

$$\begin{aligned}\langle H\rangle = \operatorname{Tr}(\hat{H}\hat{\rho}) &= -\frac{\hbar^2}{2mV}\int\left\{\nabla^2 \exp\left[-\frac{m}{2\beta\hbar^2}|\boldsymbol{r}-\boldsymbol{r}'|^2\right]\right\}_{\boldsymbol{r}=\boldsymbol{r}'} d^3r \\ &= \frac{1}{2\beta V}\int\left\{\left[3-\frac{m}{\beta\hbar^2}|\boldsymbol{r}-\boldsymbol{r}'|^2\right]\exp\left[-\frac{m}{2\beta\hbar^2}|\boldsymbol{r}-\boldsymbol{r}'|^2\right]\right\}_{\boldsymbol{r}=\boldsymbol{r}'} d^3r \\ &= \frac{3}{2\beta} = \frac{3}{2}kT,\end{aligned} \tag{15}$$

which was indeed expected. Otherwise, too,

$$\langle H\rangle = \frac{\operatorname{Tr}(\hat{H}e^{-\beta\hat{H}})}{\operatorname{Tr}(e^{-\beta\hat{H}})} = -\frac{\partial}{\partial\beta}\ln\operatorname{Tr}(e^{-\beta\hat{H}}) \tag{16}$$

which, on combination with (13), leads to the same result.

C. A linear harmonic oscillator

Next, we consider the case of a linear harmonic oscillator, whose Hamiltonian is given by

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial q^2} + \frac{1}{2}m\omega^2 q^2, \tag{17}$$

with eigenvalues

$$E_n = \left(n + \tfrac{1}{2}\right)\hbar\omega; \qquad n = 0, 1, 2, \ldots \tag{18}$$

and eigenfunctions

$$\phi_n(q) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\frac{H_n(\xi)}{(2^n n!)^{1/2}}e^{-(1/2)\xi^2}, \tag{19}$$

where

$$\xi = \left(\frac{m\omega}{\hbar}\right)^{1/2} q \tag{20}$$

and

$$H_n(\xi) = (-1)^n e^{\xi^2}\left(\frac{d}{d\xi}\right)^n e^{-\xi^2}. \tag{21}$$

The matrix elements of the operator exp $(-\beta\hat{H})$ in the q-representation are given b

$$\langle q|e^{-\beta\hat{H}}|q'\rangle = \sum_{n=0}^{\infty} e^{-\beta E_n}\phi_n(q)\phi_n(q')$$

$$= \left(\frac{m\omega}{\pi\hbar}\right)^{1/2} e^{-(1/2)(\xi^2+\xi'^2)} \sum_{n=0}^{\infty}\left\{e^{-(n+1/2)\beta\hbar\omega}\frac{H_n(\xi)H_n(\xi')}{2^n n!}\right\}. \quad (2$$

The summation over n is somewhat difficult to evaluate; nevertheless, the fin result is[3]

$$\langle q|e^{-\beta\hat{H}}|q'\rangle = \left[\frac{m\omega}{2\pi\hbar\sinh(\beta\hbar\omega)}\right]^{1/2}$$

$$\times \exp\left[-\frac{m\omega}{4\hbar}\left\{(q+q')^2\tanh\left(\frac{\beta\hbar\omega}{2}\right) + (q-q')^2\coth\left(\frac{\beta\hbar\omega}{2}\right)\right\}\right], \quad (2$$

whence

$$\mathrm{Tr}\,(e^{-\beta\hat{H}}) = \int_{-\infty}^{\infty}\langle q|e^{-\beta\hat{H}}|q\rangle\,dq$$

$$= \left[\frac{m\omega}{2\pi\hbar\sinh(\beta\hbar\omega)}\right]^{1/2}\int_{-\infty}^{\infty}\exp\left[-\frac{m\omega q^2}{\hbar}\tanh\left(\frac{\beta\hbar\omega}{2}\right)\right]dq$$

$$= \frac{1}{2\sinh\left(\frac{1}{2}\beta\hbar\omega\right)} = \frac{e^{-(1/2)\beta\hbar\omega}}{1-e^{-\beta\hbar\omega}}. \quad (2$$

Expression (24) is indeed the *partition function* of a linear harmonic oscillat see eqn. (3.8.14). At the same time, we find that the *probability density* for oscillator coordinate to be in the vicinity of the value q is given by

$$\langle q|\hat{\rho}|q\rangle = \left[\frac{m\omega\tanh\left(\frac{1}{2}\beta\hbar\omega\right)}{\pi\hbar}\right]^{1/2}\exp\left[-\frac{m\omega q^2}{\hbar}\tanh\left(\frac{\beta\hbar\omega}{2}\right)\right]; \quad (2$$

we note that this is a Gaussian distribution in q, with mean value zero and r mean-square deviation

$$q_{\mathrm{r.m.s.}} = \left[\frac{\hbar}{2m\omega\tanh\left(\frac{1}{2}\beta\hbar\omega\right)}\right]^{1/2} \quad ($$

The probability distribution (25) was first derived by Bloch in 1932. In the c sical limit ($\beta\hbar\omega \ll 1$), the distribution becomes *purely thermal*—free from quan effects:

$$\langle q|\hat{\rho}|q\rangle \approx \left(\frac{m\omega^2}{2\pi kT}\right)^{1/2}\exp\left[-\frac{m\omega^2 q^2}{2kT}\right],$$

with dispersion $(kT/m\omega^2)^{1/2}$. At the other extreme $(\beta\hbar\omega \gg 1)$, the distribution becomes *purely quantum-mechanical*—free from thermal effects:

$$\langle q|\hat{\rho}|q\rangle \approx \left(\frac{m\omega}{\pi\hbar}\right)^{1/2} \exp\left[-\frac{m\omega q^2}{\hbar}\right] \tag{28}$$

with dispersion $(\hbar/2m\omega)^{1/2}$. Note that the limiting distribution (28) is precisely the one expected for an oscillator in its ground state $(n = 0)$, viz. one with probability density $\phi_0^2(q)$; see eqns (19)–(21).

In view of the fact that the mean energy of the oscillator is given by

$$\langle H\rangle = -\frac{\partial}{\partial\beta}\ln \mathrm{Tr}\,(e^{-\beta\hat{H}}) = \tfrac{1}{2}\hbar\omega\coth\left(\tfrac{1}{2}\beta\hbar\omega\right), \tag{29}$$

we observe that the temperature dependence of the distribution (25) is *solely* determined through the expectation value $\langle H\rangle$. Actually, we can write

$$\langle q|\hat{\rho}|q\rangle = \left(\frac{m\omega^2}{2\pi\langle H\rangle}\right)^{1/2} \exp\left[-\frac{m\omega^2 q^2}{2\langle H\rangle}\right] \tag{30}$$

with

$$q_{\text{r.m.s.}} = \left(\frac{\langle H\rangle}{m\omega^2}\right)^{1/2} \tag{31}$$

It is now straightforward to see that the mean value of the potential energy $(\frac{1}{2}m\omega^2 q^2)$ of the oscillator is $\frac{1}{2}\langle H\rangle$; accordingly, the mean value of the kinetic energy $(p^2/2m)$ must also be the same.

5.4. Systems composed of indistinguishable particles

We shall now formulate the quantum-mechanical description of a system of N identical particles. To fix ideas, we consider a gas of *non-interacting* particles: the findings of this study will be of considerable relevance to other systems as well.

Now, the Hamiltonian of a system of N non-interacting particles is simply a sum of the individual single-particle Hamiltonians:

$$\hat{H}(q,p) = \sum_{i=1}^{N} \hat{H}_i(q_i, p_i); \tag{1}$$

here, (q_i, p_i) are the coordinates and momenta of the ith particle while $\hat{H}_i$ is its Hamiltonian.[4] Since the particles are identical, the Hamiltonians $\hat{H}_i (i = 1, 2, \ldots, N)$ are *formally* the same; they only differ in the values of their arguments. The time-independent Schrödinger equation for the system is

$$\hat{H}\psi_E(q) = E\psi_E(q), \tag{2}$$

where E is an eigenvalue of the Hamiltonian and $\psi_E(q)$ the corresponding eigenfunction. In view of (1), we can write a straightforward solution of the Schrödinger equation, namely

$$\psi_E(q) = \prod_{i=1}^{N} u_{\varepsilon_i}(q_i), \tag{3}$$

with

$$E = \sum_{i=1}^{N} \varepsilon_i; \tag{4}$$

the factor $u_{\varepsilon_i}(q_i)$ in (3) is an eigenfunction of the single-particle Hamiltonian $\hat{H}_i(q_i, p_i)$, with eigenvalue ε_i:

$$\hat{H}_i u_{\varepsilon_i}(q_i) = \varepsilon_i u_{\varepsilon_i}(q_i). \tag{5}$$

Thus, a stationary state of the given system may be described in terms of the single-particle states of the constituent particles. In general, we may do so by specifying the set of numbers $\{n_i\}$ to represent a particular state of the system; this would imply that there are n_i particles in the eigenstate characterized by the energy value ε_i. Clearly, the distribution set $\{n_i\}$ must conform to the conditions

$$\sum_i n_i = N \tag{6}$$

and

$$\sum_i n_i \varepsilon_i = E. \tag{7}$$

Accordingly, the wave function of this state may be written as

$$\psi_E(q) = \prod_{m=1}^{n_1} u_1(m) \prod_{m=n_1+1}^{n_1+n_2} u_2(m) \ldots, \tag{8}$$

where the symbol $u_i(m)$ stands for the single-particle wave function $u_{\varepsilon_i}(q_m)$.

Now, suppose we effect a permutation among the coordinates appearing on the right-hand side of (8); as a result, the coordinates $(1, 2, \ldots, N)$ get replaced by $(P1, P2, \ldots, PN)$, say. The resulting wave function, which we shall call $P\psi_E(q)$, will be

$$P\psi_E(q) = \prod_{m=1}^{n_1} u_1(Pm) \prod_{m=n_1+1}^{n_1+n_2} u_2(Pm) \ldots. \tag{9}$$

In classical physics, where the particles of a given system, even though identical, are regarded as mutually *distinguishable*, any permutation that brings about an interchange of particles in two *different* single-particle states is recognized to have led to a *new, physically distinct* microstate of the system. For example, classical physics regards a microstate in which the so-called 5th particle is in the state u_i and the so-called 7th particle in the state $u_j (j \neq i)$ as distinct from a microstate in which the 7th particle is in the state u_i and the 5th in the state u_j. This leads to

$$\frac{N!}{n_1! n_2! \ldots} \tag{10}$$

(supposedly distinct) microstates of the system, corresponding to a given mode of distribution $\{n_i\}$. The number (10) would then be ascribed as a "statistical weight factor" to the distribution set $\{n_i\}$. Of course, the "correction" applied by Gibbs, which has been discussed in Secs 1.5 and 1.6, reduces this weight factor to

$$W_c\{n_i\} = \frac{1}{n_1!n_2!\dots}. \tag{11}$$

And the only way one could understand the physical basis of that "correction" was in terms of the inherent *indistinguishability* of the particles.

According to quantum physics, however, the situation remains unsatisfactory even after the Gibbs correction has been incorporated, for, strictly speaking, an interchange among identical particles, *even* if they are in different single-particle states, should not lead to a new microstate of the system! Thus, if we want to take into account the indistinguishability of the particles properly, we must not regard a microstate in which the "5th" particle is in the state u_i and the "7th" in the state u_j as distinct from a microstate in which the "7th" particle is in the state u_i and the "5th" in the state u_j (even if $i \neq j$), for the labeling of the particles as No. 1, No. 2, etc. (which one often resorts to) is at most a matter of convenience–it is not a matter of reality. In other words, all that matters in the description of a particular state of the given system is the set of numbers n_i which tell us "*how many* particles there are in the various single-particle states u_i"; the question "*which* particle is in *which* single-particle state" has no relevance at all. Accordingly, the microstates resulting from any permutation P among the N particles (so long as the numbers n_i remain the same) must be regarded as *one and the same* microstate. For the same reason, the weight factor associated with a distribution set $\{n_i\}$, provided that the set is not disallowed on some other physical grounds, should be identically equal to unity, whatever the values of the numbers n_i:

$$W_q\{n_i\} \equiv 1 \qquad (12)^{5}$$

Indeed, if for some physical reason the set $\{n_i\}$ is disallowed, the weight factor W_q for that set should be identically equal to zero; see, for instance, eqn. (19).

At the same time, a wave function of the type (8), which we may call *Boltzmannian* and denote by the symbol $\psi_{\text{Boltz}}(q)$, is inappropriate for describing the state of a system composed of indistinguishable particles because an interchange of arguments among the factors u_i and u_j, where $i \neq j$, would lead to a wave function which is both mathematically and physically different from the one we started with. Now, since a mere interchange of the particle coordinates must not lead to a new microstate of the system, the wave function $\psi_E(q)$ must be constructed in such a way that, for all practical purposes, it is insensitive to any interchange among its arguments. The simplest way to do this is to set up a linear combination of all the $N!$ functions of the type (9) which obtain from (8) by all possible permutations among its arguments; of course, the combination must be such that if a permutation of coordinates is carried out in it, then the wave functions ψ and $P\psi$ must satisfy the property

$$|P\psi|^2 = |\psi|^2. \tag{13}$$

This leads to the following possibilities:

(i)
$$P\psi = \psi \quad \text{for all } P, \tag{14}$$

which means that the wave function is symmetric in all its arguments, or

(ii)
$$P\psi = \begin{cases} +\psi & \text{if } P \text{ is an } \textit{even} \text{ permutation,} \\ -\psi & \text{if } P \text{ is an } \textit{odd} \text{ permutation,} \end{cases} \tag{15}$$[6]

which means that the wave function is *antisymmetric* in its arguments. We call these wave functions ψ_S and ψ_A, respectively; their mathematical structure is given by

$$\psi_S(q) = \text{const.} \sum_P P\psi_{\text{Boltz}}(q) \tag{16}$$

and

$$\psi_A(q) = \text{const.} \sum_P \delta_P P\psi_{\text{Boltz}}(q). \tag{17}$$

where δ_P in the expression for ψ_A is $+1$ or -1 according as the permutation P is even or odd.

We note that the function $\psi_A(q)$ can be written in the form of a *Slater determinant:*

$$\psi_A(q) = \text{const.} \begin{vmatrix} u_i(1) & u_i(2) & & u_i(N) \\ u_j(1) & u_j(2) & & u_j(N) \\ & & & \\ & & & \\ u_l(1) & u_l(2) & \ldots & u_l(N) \end{vmatrix}, \tag{18}$$

where the leading diagonal is precisely the Boltzmannian wave function while the other terms of the expansion are the various permutations thereof; positive and negative signs in the combination (17) appear automatically as we expand the determinant. On interchanging a pair of arguments (which amounts to interchanging the corresponding columns of the determinant), the wave function ψ_A merely changes its sign, as it indeed should. However, if two or more particles happen to be in the same single-particle state, then the corresponding rows of the determinant become identical and the wave function vanishes.[7] Such a state is physically impossible to realize. We therefore conclude that if a system composed of indistinguishable particles is characterized by an antisymmetric wave function, then the particles of the system must all be in different single-particle states–a result equivalent to *Pauli's exclusion principle* for electrons. Conversely, a statistical system composed of particles obeying an exclusion principle must be described by a wave function which is antisymmetric in its arguments. The statistics governing the behavior of such particles is called *Fermi–Dirac*, or simply *Fermi*, statistics and the constituent particles themselves are referred to as *fermions*. The statistical weight factor $W_{\text{F.D.}}\{n_i\}$ for such a system is unity so long as the n_i in the distribution set are either 0 or 1; otherwise, it is zero:

$$W_{\text{F.D.}}\{n_i\} = \begin{cases} 1 & \text{if} \quad \sum_i n_i^2 = N, \\ 0 & \text{if} \quad \sum_i n_i^2 > N. \end{cases} \tag{19}^8$$

o such problems arise for systems characterized by symmetric wave functions; particular, we have no restriction whatsoever on the values of the numbers n_i. he statistics governing the behavior of such systems is called *Bose–Einstein*, or mply *Bose*, statistics and the constituent particles themselves are referred to as *sons*. The weight factor $W_{\text{B.E.}}\{n_i\}$ is identically equal to 1, whatever the values the numbers n_i:

$$W_{\text{B.E.}}\{n_i\} = 1; \quad n_i = 0, 1, 2, \ldots. \tag{20}$$

It should be pointed out here that there exists an intimate connection between e statistics governing a particular species of particles and the intrinsic spin of particles. For instance, particles with an integral spin (in units of $\hbar$, of course) ey Bose–Einstein statistics, while particles with a half-odd integral spin obey rmi–Dirac statistics. Examples in the first category are photons, phonons, π-esons, gravitons, He^4-atoms, etc., while those in the second category are elecons, nucleons (protons and neutrons), μ-mesons, neutrinos, He^3-atoms, etc.

Finally, it must be emphasized that, although we have derived our conclusions re on the basis of a study of non-interacting systems, the basic results hold for teracting systems as well. In general, the desired wave function $\psi(q)$ will not expressible in terms of the single-particle wave functions $u_i(q_m)$: nonetheless, will have to be either of the kind $\psi_S(q)$, satisfying eqn. (14), or of the kind ${}_A(q)$, satisfying eqn. (15).

5.5. The density matrix and the partition function of a system of free particles[9]

Suppose that the given system, which is composed of N indistinguishable, non-teracting particles confined to a cubical box of volume V, is a member of a anonical ensemble characterized by the temperature parameter β. The *density trix* of the system in the coordinate representation will be given by

$$\langle \boldsymbol{r}_1, \ldots, \boldsymbol{r}_N | \hat{\rho} | \boldsymbol{r}'_1, \ldots, \boldsymbol{r}'_N \rangle = \frac{1}{Q_N(\beta)} \langle \mathbf{r}_1, \ldots, \boldsymbol{r}_N | e^{-\beta \hat{H}} | \boldsymbol{r}'_1, \ldots, \boldsymbol{r}'_N \rangle, \tag{1}$$

here $Q_N(\beta)$ is the *partition function* of the system:

$$Q_N(\beta) = \operatorname{Tr}\left(e^{-\beta \hat{H}}\right) = \int \langle \boldsymbol{r}_1, \ldots, \boldsymbol{r}_N | e^{-\beta \hat{H}} | \boldsymbol{r}_1, \ldots, \boldsymbol{r}_N \rangle \, d^{3N} r. \tag{2}$$

r brevity, we denote the vector $\boldsymbol{r}_i$ by the letter i and the primed vector $\boldsymbol{r}'_i$ by i'. rther, let $\psi_E(1, \ldots, N)$ denote the eigenfunctions of the Hamiltonian, the suffix representing the corresponding eigenvalues. We then have

$$\langle 1, \ldots, N | e^{-\beta \hat{H}} | 1', \ldots, N' \rangle = \sum_E e^{-\beta E} \left[\psi_E(1, \ldots, N) \psi_E^*(1', \ldots, N') \right], \tag{3}$$

here the summation goes over all possible values of E; cf. eqn. (5.3.11).

Since the particles constituting the given system are non-interacting, we may express the eigenfunctions $\psi_E(1, \ldots, N)$ and the eigenvalues E in terms of the single-particle wave functions $u_i(m)$ and the single-particle energies ε_i. Moreover, we find it advisable to work with the wave vectors $\boldsymbol{k}_i$ rather than the energies ε_i; so we write

$$E = \frac{\hbar^2 K^2}{2m} = \frac{\hbar^2}{2m}(k_1^2 + k_2^2 + \cdots + k_N^2), \tag{4}$$

where the k_i on the right-hand side are the wave vectors of the individual particles. Imposing periodic boundary conditions, the *normalized* single-particle wave functions are

$$u_k(\boldsymbol{r}) = V^{-1/2} \exp\{i(\boldsymbol{k} \cdot \boldsymbol{r})\}, \tag{5}$$

with

$$\boldsymbol{k} = 2\pi V^{-1/3} \boldsymbol{n}; \tag{6}$$

here, $\boldsymbol{n}$ stands for a three-dimensional vector whose components can have values $0, \pm 1, \pm 2, \ldots$. The wave function ψ of the total system would then be, see eqns (5.4.16) and (5.4.17),

$$\psi_K(1, \ldots, N) = (N!)^{-1/2} \sum_P \delta_P P \left\{ u_{k_1}(1) \ldots u_{k_N}(N) \right\}, \tag{7}$$

where the magnitudes of the individual k_i are such that

$$(k_1^2 + \cdots + k_N^2) = K^2. \tag{8}$$

The number δ_P in the expression for ψ_K is identically equal to $+1$ if the particles are bosons. For fermions, it is $+1$ or -1 according as the permutation P is even or odd. Thus, quite generally, we may write

$$\delta_P = (\pm 1)^{[P]}, \tag{9}$$

where $[P]$ denotes the order of the permutation; note that the upper sign in this expression holds for bosons while the lower sign holds for fermions. The factor $(N!)^{-1/2}$ has been introduced here to secure the normalization of the total wave function.

Now, it makes no difference to the wave function (7) whether the permutations P are carried out on the coordinates $1, \ldots, N$ or on the wave vectors $\boldsymbol{k}_1, \ldots, \boldsymbol{k}_N$, because after all we are going to sum over *all* the $N!$ permutations. Denoting the permuted coordinates by $P1, \ldots, PN$ and the permuted wave vectors by $P\boldsymbol{k}_1, \ldots, P\boldsymbol{k}_N$, eqn. (7) may be written as

$$\psi_K(1, \ldots, N) = (N!)^{-1/2} \sum_P \delta_P \left\{ u_{k_1}(P1) \ldots u_{k_N}(PN) \right\} \tag{10a}$$

$$= (N!)^{-1/2} \sum_P \delta_P \left\{ u_{Pk_1}(1) \ldots u_{Pk_N}(N) \right\}. \tag{10b}$$

Equations (10) may now be substituted into (3), with the result

$$\langle 1, \ldots, N | e^{-\beta \hat{H}} | 1', \ldots, N' \rangle = (N!)^{-1} \sum_K e^{-\beta \hbar^2 K^2 / 2m} \times$$

$$\times\left[\sum_{P}\delta_P\left\{u_{k_1}(P1)\ldots u_{k_N}(PN)\right\}\sum_{\tilde P}\delta_{\tilde P}\left\{u^*_{\tilde P k_1}(1')\ldots u^*_{\tilde P k_N}(N')\right\}\right], \qquad (11)$$

where P and $\tilde P$ are any of the $N!$ possible permutations. Now, since a permutation among the k_i changes the wave function ψ at most by a sign, the quantity $[\psi\psi^*]$ in (11) is insensitive to such a permutation; the same holds for the exponential factor as well. The summation over K is, therefore, equivalent to $(1/N!)$ times a summation over all the vectors $k_1, \ldots, k_N$ *independently of one another*. Next, in view of the N-fold summation over the k_i, all the permutations $\tilde P$ will make equal contributions towards the sum (because they differ from one another only in the ordering of the k_i). Therefore, we may consider only one of these permutations, say the one for which $\tilde P k_1 = k_1, \ldots, \tilde P k_N = k_N$ (and hence $\delta_{\tilde P} = 1$ for both kinds of statistics), and include a factor of $(N!)$ along. The net result is

$$\langle 1, \ldots, N|e^{-\beta\hat H}|1', \ldots, N'\rangle = (N!)^{-1}\sum_{k_1,\ldots,k_N}$$

$$e^{-\beta\hbar^2(k_1^2+\cdots+k_N^2)/2m}\left[\sum_{P}\delta_P\left\{u_{k_1}(P1)u^*_{k_1}(1')\right\}\ldots\left\{u_{k_N}(PN)u^*_{k_N}(N')\right\}\right]. \qquad (12)$$

Substituting from (5) and noting that, in view of the largeness of V, the summations over the k_i may be replaced by integrations, eqn. (12) becomes

$$\langle 1, \ldots, N|e^{-\beta\hat H}|1', \ldots, N'\rangle$$

$$= \frac{1}{N!(2\pi)^{3N}}\sum_{P}\delta_P\left[\int e^{-\beta\hbar^2k_1^2/2m+ik_1\cdot(P1-1')}\,d^3k_1\ldots\right.$$

$$\left.\int e^{-\beta\hbar^2k_N^2/2m+ik_N\cdot(PN-N')}\,d^3k_N\right] \qquad (13)$$

$$= \frac{1}{N!}\left(\frac{m}{2\pi\beta\hbar^2}\right)^{3N/2}\sum_{P}\delta_P[f(P1-1')\ldots f(PN-N')], \qquad (14)$$

where

$$f(\xi) = \exp\left(-\frac{m}{2\beta\hbar^2}\xi^2\right). \qquad (15)$$

Here, use has been made of the mathematical result (5.3.12), which is clearly a special case of the present formula.

Introducing the *mean thermal wavelength*

$$\lambda = \frac{h}{(2\pi mkT)^{1/2}} = \hbar\left(\frac{2\pi\beta}{m}\right)^{1/2} \qquad (16)$$

and rewriting our coordinates as $r_1, \ldots, r_N$, the diagonal elements among (14) take the form

$$\langle r_1, \ldots, r_N|e^{-\beta\hat H}|r_1, \ldots, r_N\rangle = \frac{1}{N!\lambda^{3N}}\sum_{P}\delta_P[f(Pr_1 - r_1)\ldots f(Pr_N - r_N)], \qquad (17)$$

where

$$f(\boldsymbol{r}) = \exp(-\pi r^2/\lambda^2). \qquad (18$$

To obtain the partition function of the system, we have to integrate (17) over all the coordinates involved. However, before we do that, we would like to make some observations on the summation $\sum_P$. First of all, we note that the leading term in this summation, namely the one for which $P\boldsymbol{r}_i = \boldsymbol{r}_i$, is identically equal to unity (because $f(0) = 1$). This is followed by a group of terms in which *only one* pair interchange (among the coordinates) has taken place; a typical term in this group will be $f(\boldsymbol{r}_j - \boldsymbol{r}_i) f(\boldsymbol{r}_i - \boldsymbol{r}_j)$ where $i \neq j$. This group of terms is followed by other groups of terms in which *more than one* pair interchanges have taken place. Thus, we may write

$$\sum_P = 1 \pm \sum_{i<j} f_{ij} f_{ji} + \sum_{i<j<k} f_{ij} f_{jk} f_{ki} \pm \cdots . \qquad (19$$

where $f_{ij} \equiv f(\boldsymbol{r}_i - \boldsymbol{r}_j)$; again note that the upper (lower) signs in this expansion pertain to a system of bosons (fermions). Now, the function f_{ij} vanishes rapidly as the distance r_{ij} becomes much larger than the mean thermal wavelength λ. It follows that if the mean interparticle distance, $(V/N)^{1/3}$, in the system is much larger than the mean thermal wavelength, i.e. if

$$\frac{nh^3}{(2\pi mkT)^{3/2}} \ll 1. \qquad (20)$$

where n is the particle density in the system, then the sum $\sum_P$ in (19) may b approximated by unity. Accordingly, the partition function of the system becomes see eqn. (17).

$$Q_N(V, T) \equiv \mathrm{Tr}\,(e^{-\beta \hat{H}}) \approx \frac{1}{N!\lambda^{3N}} \int 1\,(d^{3N}r) = \frac{1}{N!}\left(\frac{V}{\lambda^3}\right)^N . \qquad (21$$

This is precisely the result obtained earlier for the classical ideal gas; se eqn. (3.5.9). Thus, we have obtained from our quantum-mechanical treatme the precise classical limit for the partition function $Q_N(V, T)$. Incidentally, w have achieved something more. Firstly, we have automatically recovered here the Gibbs correction factor $(1/N!)$ which was introduced into the classical treatmen on an *ad hoc*, semi-empirical basis. We, of course, tried to understand its origi in terms of the inherent indistinguishability of the particles. Here, on the oth hand, we see it coming in a very natural manner and its source indeed lies in the symmetrization of the wave functions of the system (which is ultimately relate to the indistinguishability of the particles); cf. Problem 5.4. Secondly, we fin here a formal justification for computing the number of microstates of a syste corresponding to a given region of its phase space by dividing the volume that region into cells of a "suitable" size and then counting instead the number these cells. This correspondence becomes all the more transparent by noting th formula (21) is exactly equivalent to the classical expression

$$Q_N(V, T) = \frac{1}{N!} \int e^{-\beta(p_1^2 + \cdots + p_N^2)/2m} \left(\frac{d^{3N}q\, d^{3N}p}{\omega_0}\right), \qquad (2$$

with $\omega_0 = h^{3N}$. Thirdly, in deriving the classical limit we have also evolved a riterion which enables us to determine whether a given physical system can be reated classically; mathematically, this criterion is given by condition (20). Now, n statistical mechanical studies, a system which cannot be treated classically is said to be *degenerate*; the quantity $(n\lambda^3)$ is, therefore, referred to as the *degen-acy discriminant*. Accordingly, the condition that classical considerations may e applicable to a given physical system is that "the value of the degeneracy discriminant of the system be much less than unity".

Next, we note that, in the classical limit, the diagonal elements of the density matrix are given by

$$\langle r_1 \dots r_N|\hat{\rho}|r_1, \dots, r_N\rangle \approx \left(\frac{1}{V}\right)^N, \tag{23}$$

which is simply a product of N factors, each equal to $(1/V)$. Recalling that for a single particle in a box of volume V, $\langle \mathbf{r}|\hat{\rho}|\mathbf{r}\rangle = (1/V)$, see eqn. (5.3.16), we infer that in the classical limit there is no spatial correlation among the various particles of the system. In general, however, spatial correlations exist even if the particles are supposedly non-interacting; they arise from the symmetrization of the wave functions and their magnitude is quite significant if the interparticle distances in he system are comparable with the mean thermal wavelength of the particles. To see this more clearly, we consider the simplest relevant case, namely the one with $N = 2$. The sum $\sum_P$ is now exactly equal to $1 \pm [f(r_{12})]^2$. Accordingly,

$$\langle r_1, r_2|e^{-\beta\hat{H}}|r_1, r_2\rangle = \frac{1}{2\lambda^6}[1 \pm \exp(-2\pi r_{12}^2/\lambda^2)] \tag{24}$$

and hence

$$\begin{aligned}
Q_2(V,T) &= \frac{1}{2\lambda^6}\iint [1 \pm \exp(-2\pi r_{12}^2/\lambda^2] \, d^3r_1\, d^3r_2 \\
&= \frac{1}{2}\left(\frac{V}{\lambda^3}\right)^2 \left[1 \pm \frac{1}{V}\int\limits_0^\infty \exp(-2\pi r^2/\lambda^2)\, 4\pi r^2\, dr\right] \\
&= \frac{1}{2}\left(\frac{V}{\lambda^3}\right)^2 \left[1 \pm \frac{1}{2^{3/2}}\left(\frac{\lambda^3}{V}\right)\right] \qquad (25) \\
&\approx \frac{1}{2}\left(\frac{V}{\lambda^3}\right)^2. \qquad (26)
\end{aligned}$$

Combining (24) and (26), we obtain

$$\langle r_1, r_2|\hat{\rho}|r_1, r_2\rangle \approx \frac{1}{V^2}[1 \pm \exp(-2\pi r_{12}^2/\lambda^2)]. \tag{27}$$

Thus, if r_{12} is comparable to λ, the probability density (27) may differ considerably from the classical value $(1/V)^2$. In particular, the probability density for a pair of *osons* to be a distance r apart is larger than the classical, r-independent value by a factor of $[1 + \exp(-2\pi r^2/\lambda^2)]$ which becomes as high as 2 as $r \to 0$. The corresponding result for a pair of *fermions* is smaller than the classical value by a factor of $[1 - \exp(-2\pi r^2/\lambda^2)]$ which becomes low as 0 as $r \to 0$. Thus, we obtain a *positive* spatial correlation among particles obeying Bose statistics and a *negative* spatial correlation among particles obeying Fermi statistics; see also Sec. 6.3.

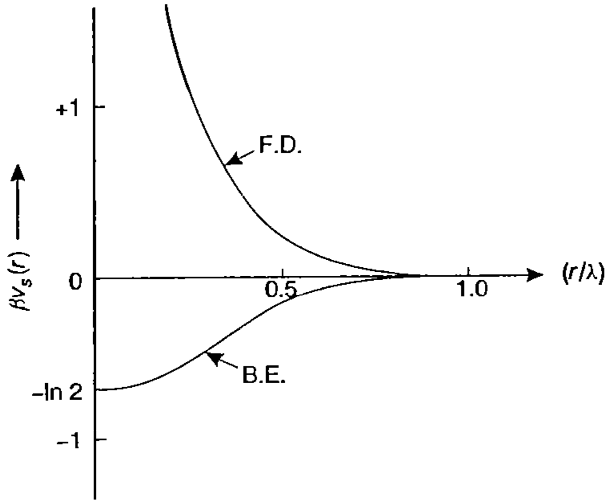


FIG. 5.1. The statistical potential $v_s(r)$ between a pair of particles obeying Bose–Einstein statistics or Fermi–Dirac statistics.

Another way of expressing correlations (among otherwise non-interacting particles) is by introducing a *statistical* interparticle potential $v_s(r)$ and then treating the particles classically (Uhlenbeck and Gropper, 1932). The potential $v_s(r)$ must be such that the Boltzmann factor $\exp(-\beta v_s)$ is precisely equal to the pair distribution function $[\ldots]$ in (27), i.e.,

$$v_s(r) = -kT \ln[1 \pm \exp(-2\pi r^2/\lambda^2)]. \tag{28}$$

Figure 5.1 shows a plot of the statistical potential $v_s(r)$ for a pair of bosons or fermions. In the Bose case, the potential is throughout attractive, thus giving rise to a "statistical attraction" among bosons; in the Fermi case, it is throughout repulsive, giving rise to a "statistical repulsion" among fermions. In either case, the potential vanishes rapidly as r becomes larger than λ; accordingly, its influence becomes less important as the temperature of the system rises.

Problems

5.1. Evaluate the density matrix ρ_{mn} of an electron spin in the representation which makes $\hat{\sigma}_x$ diagonal. Next, show that the value of $\langle\sigma_z\rangle$, resulting from this representation, is precisely the same as the one obtained in Sec. 5.3.

Hint: The representation needed here follows from the one used in Sec. 5.3 by carrying out a transformation with the help of the unitary operator

$$\hat{U} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

5.2. Prove that

$$\langle q|e^{-\beta\hat{H}}|q'\rangle \equiv \exp\left[-\beta\hat{H}\left(-i\hbar\frac{\partial}{\partial q}, q\right)\right]\delta(q - q').$$

where $\hat{H}(-i\hbar\,\partial/\partial q, q)$ is the Hamiltonian of the system, in the q-representation, which formally operates upon the Dirac delta function $\delta(q-q')$. Writing the δ-function in a suitable form, apply this result to (i) a free particle and (ii) a linear harmonic oscillator.

5.3. Derive the density matrix ρ for (i) a free particle and (ii) a linear harmonic oscillator in the *momentum* representation and study its main properties along the lines of Sec. 5.3.

5.4. Study the density matrix and the partition function of a system of free particles, using the *unsymmetrized* wave function (5.4.3) instead of the *symmetrized* wave function (5.5.7). Show that, following this procedure, one obtains neither the Gibbs' correction factor $(1/N!)$ nor a spatial correlation among the particles.

5.5. Show that in the *first* approximation the partition function of a system of N non-interacting, indistinguishable particles is given by

$$Q_N(V,T) = \frac{1}{N!\,\lambda^{3N}} Z_N(V,T),$$

where

$$Z_N(V,T) = \int \exp\left\{-\beta \sum_{i<j} v_s(r_{ij})\right\} d^{3N}r,$$

$v_s(r)$ being the statistical potential (5.5.28). Hence evaluate the *first-order* correction to the equation of state of this system.

5.6. Determine the values of the degeneracy discriminant $(n\lambda^3)$ for hydrogen, helium and oxygen at N.T.P. Make an estimate of the respective temperature ranges where the magnitude of this quantity becomes comparable to unity and hence quantum effects become important.

5.7. Show that the quantum-mechanical partition function of a system of N *interacting* particles approaches the classical form

$$Q_N(V,T) = \frac{1}{N!h^{3N}} \int e^{-\beta E(q,p)}\,(d^{3N}q\,d^{3N}p)$$

as the mean thermal wavelength λ becomes much smaller than (i) the mean interparticle distance $(V/N)^{1/3}$ and (ii) a characteristic length r_0 of the interparticle potential.[10]

5.8. Prove the following theorem due to Peierls.[11]

"If $\hat{H}$ is the hermitian Hamiltonian operator of a given physical system and $\{\phi_n\}$ an arbitrary orthonormal set of wave functions satisfying the symmetry requirements and the boundary conditions of the problem, then the partition function of the system satisfies the following inequality:

$$Q(\beta) \ge \sum_n \exp\left\{-\beta\langle\phi_n|\hat{H}|\phi_n\rangle\right\};$$

the equality holds when $\{\phi_n\}$ is a complete orthonormal set of eigenfunctions of the Hamiltonian itself."

Notes

[1] For simplicity of notation, we suppress the coordinates r_i in the argument of the wave function ψ^k.

[2] It may be noted that in this (so-called energy) representation the density operator $\hat{\rho}$ may be written as

$$\hat{\rho} = \sum_n |\phi_n\rangle \rho_n \langle\phi_n|, \tag{12}$$

for then

$$\rho_{kl} = \sum_n \langle\phi_k|\phi_n\rangle \rho_n \langle\phi_n|\phi_l\rangle = \sum_n \delta_{kn}\rho_n\delta_{nl} = \rho_k\delta_{kl}.$$

[3] The mathematical details of this derivation can be found in Kubo (1965), pp. 175–7.

[4] We are studying here a single-component system composed of "spinless" particles. Generalization to a system composed of particles with spin and to a system composed of two or more components is quite straightforward.

[5] It may be mentioned here that as early as in 1905 Ehrenfest pointed out that to obtain Planck's formula for the black-body radiation one must assign equal *a priori* probabilities to the various states $\{n_i\}$.

[6] An even (odd) permutation is one which can be arrived at from the original order by an even (odd) number of "pair interchanges" among the arguments. For example, of the six permutations

$$(1,2,3),\quad (2,3,1),\quad (3,1,2),\quad (1,3,2),\quad (3,2,1) \quad \text{and} \quad (2,1,3),$$

of the arguments 1, 2 and 3, the first three are *even* permutations while the last three are *odd*. A single interchange, among any two arguments, is clearly an *odd* permutation.

[7] This is directly related to the fact that if we effect an interchange of two particles in the *same* single-particle state, then $P\psi_A$ will obviously be identical with ψ_A. At the same time, if we also have: $P\psi_A = -\psi_A$, then ψ_A must be identically zero.

[8] Note that the condition $\sum_i n_i^2 = N$ necessarily implies that all n_i are either 0 or 1. On the other hand, if any of the n_i are greater than 1, the sum $\sum_i n_i^2$ is necessarily greater than N.

[9] For a general survey of the density matrix and its applications, see ter Haar (1961).

[10] See Huang (1963), Sec. 10.2.

[11] R. E. Peierls (1938), *Phys. Rev.* 54, 918. See also Huang (1963), Sec. 10.3.

CHAPTER 6

THE THEORY OF SIMPLE GASES

WE ARE now fully equipped with the formalism required for determining the macroscopic properties of a large variety of physical systems. In most cases, however, derivations run into serious mathematical difficulties, with the result that one is forced to restrict one's analysis either to simpler kinds of systems or to simplified models of actual systems. In practice, even these restricted studies are carried out in a series of stages, the first stage of the process being highly "idealized". The best example of such an idealization is the familiar *ideal* gas, a study of which is not only helpful in acquiring facility with the mathematical procedures but also throws considerable light on the physical behavior of gases actually met with in nature. In fact, it also serves as a base on which the theory of real gases can be founded; see Chapter 9.

In this chapter we propose to derive, and at some length discuss, the most basic properties of simple gaseous systems obeying quantum statistics; the discussion will include some of the essential features of diatomic and polyatomic gases as well.

6.1. An ideal gas in a quantum-mechanical microcanonical ensemble

We consider a gaseous system of N non-interacting, *indistinguishable* particles, confined to a space of volume V and sharing a given energy E. The statistical quantity of interest in this case is $\Omega(N, V, E)$ which, by definition, denotes the number of *distinct* microstates accessible to the system under the macrostate (N, V, E). While determining this number, we must remember that a failure to take into account the indistinguishability of the particles in a proper manner could lead to results which, except in the classical limit, may not be acceptable. With this in mind, we proceed as follows.

Since, for large V, the single-particle energy levels in the system are very close to one another, we may divide the energy spectrum into a large number of "groups of levels", which may be referred to as the *energy cells*; see Fig. 6.1. Let ε_i denote the average energy of a level, and g_i the (arbitrary) number of levels, in the ith cell; we assume that all $g_i \gg 1$. In a particular situation, we may have n_1 particles in the first cell, n_2 particles in the second cell, and so on. Clearly, the distribution set $\{n_i\}$ must conform to the conditions

$$\sum_i n_i = N \tag{1}$$

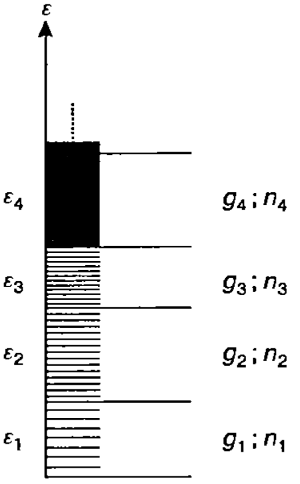


FIG. 6.1. The grouping of the single-particle energy levels into "cells".

and

$$\sum_i n_i \varepsilon_i = E. \tag{2}$$

Then

$$\Omega(N, V, E) = \sum_{\{n_i\}}{}' W\{n_i\}, \tag{3}$$

where $W\{n_i\}$ is the number of *distinct* microstates associated with the distribution set $\{n_i\}$ while the primed summation goes over all distribution sets that conform to the conditions (1) and (2). Next,

$$W\{n_i\} = \prod_i w(i), \tag{4}$$

where $w(i)$ is the number of *distinct* microstates associated with the ith cell of the spectrum (the cell that contains n_i particles, to be accommodated among g_i levels) while the product goes over all the cells in the spectrum. Clearly, $w(i)$ is the number of distinct ways in which the n_i identical, and *indistinguishable*, particles can be distributed among the g_i levels of the ith cell. This number in the Bose–Einstein case is given by, see eqn. (3.8.25),

$$w_{\text{B.E.}}(i) = \frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!}, \tag{5}$$

so that

$$W_{\text{B.E.}}\{n_i\} = \prod_i \frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!}. \tag{6}$$

In the Fermi–Dirac case no single level can accommodate more than one particle; accordingly, the number n_i cannot exceed g_i. The number $w(i)$ is then given by the "number of ways in which the g_i levels can be divided into two sub-groups–one consisting of n_i levels (which will have one particle each) and the other consisting of $(g_i - n_i)$ levels (which will be unoccupied)". This number is given by

$$w_{\text{F.D.}}(i) = \frac{g_i!}{n_i!(g_i - n_i)!}, \tag{7}$$

so that

$$W_{\text{F.D.}}\{n_i\} = \prod_i \frac{g_i!}{n_i!(g_i - n_i)!}. \tag{8}$$

For completeness, we may include the classical–or what is generally known as the Maxwell–Boltzmann–case as well. There the particles are regarded as *distinguishable*, with the result that any of the n_i particles may be put into any of the g_i levels, independently of one another, and the resulting states may all be counted as distinct; the number of these states is clearly $(g_i)^{n_i}$. Moreover, the distribution set $\{n_i\}$ is itself regarded as obtainable in

$$\frac{N!}{n_1!n_2!\dots} \tag{9}$$

different ways which, on the introduction of the Gibbs correction factor, lead to a "weight factor" of

$$\frac{1}{n_1!n_2!\dots} = \prod_i \frac{1}{n_i!}; \tag{10}$$

see also Sec. 1.6, especially eqn. (1.6.2). Combining the two results, we obtain

$$W_{\text{M.B.}}\{n_i\} = \prod_i \frac{(g_i)^{n_i}}{n_i!}. \tag{11}$$

Now, the entropy of the system would be given by

$$S(N, V, E) = k \ln \Omega(N, V, E) = k \ln \left[\sum_{\{n_i\}}{}' W\{n_i\} \right]. \tag{12}$$

It can be shown that, under the conditions of our analysis, the logarithm of the sum on the right-hand side of (12) can be approximated by the logarithm of the largest term in the sum; cf. Problem 3.4. We may, therefore, replace (12) by

$$S(N, V, E) \approx k \ln W\{n_i^*\}, \tag{13}$$

where $\{n_i^*\}$ is the distribution set that maximizes the number $W\{n_i\}$; the numbers n_i^* are clearly the *most probable values* of the distribution numbers n_i. The maximization, however, is to be carried out under the restrictions that the quantities N and E remain constant. This can be done by the method of Lagrange's undetermined multipliers; cf. Sec. 3.2. Our condition for determining the most probable distribution set $\{n_i^*\}$ turns out to be, see eqns (1), (2) and (13),

$$\delta \ln W\{n_i\} - \left[\alpha \sum_i \delta n_i + \beta \sum_i \varepsilon_i \delta n_i \right] = 0. \tag{14}$$

For $\ln W\{n_i\}$, we obtain from eqns (6), (8) and (11), assuming that not only all g_i but also all $n_i \gg 1$ (so that the Stirling approximation $\ln(x!) \approx x \ln x - x$ can be applied to all the factorials that appear),

$$\ln W\{n_i\} = \sum_i \ln w(i)$$

$$\approx \sum_i \left[n_i \ln\left(\frac{g_i}{n_i} - a\right) - \frac{g_i}{a} \ln\left(1 - a\frac{n_i}{g_i}\right)\right], \tag{15}$$

where $a = -1$ for the B.E. case, $+1$ for the F.D. case, and 0 for the M.B. case. Equation (14) then becomes

$$\sum_i \left[\ln\left(\frac{g_i}{n_i} - a\right) - \alpha - \beta\varepsilon_i\right]_{n_i = n_i^*} \delta n_i = 0. \tag{16}$$

In view of the arbitrariness of the increments δn_i in (16), we must have (for all i)

$$\ln\left(\frac{g_i}{n_i^*} - a\right) - \alpha - \beta\varepsilon_i = 0, \tag{17}$$

so that[1]

$$n_i^* = \frac{g_i}{e^{\alpha + \beta\varepsilon_i} + a}. \tag{18}$$

The fact that n_i^* turns out to be *directly* proportional to g_i prompts us to interpret the quantity

$$\frac{n_i^*}{g_i} = \frac{1}{e^{\alpha + \beta\varepsilon_i} + a}, \tag{18a}$$

which is actually the most probable number of particles *per* energy level in the ith cell, as the most probable number of particles in a *single* level of energy ε_i. Incidentally, our final result (18a) is totally independent of the manner in which the energy levels of the particles are grouped into cells, so long as the number of levels in each cell is sufficiently large. As shown in Sec. 6.2, formula (18a) can also be derived without grouping energy levels into cells at all; in fact, it is only then that this result becomes truly acceptable.

Substituting (18) into (15), we obtain for the entropy of the gas

$$\frac{S}{k} \approx \ln W\{n_i^*\} = \sum_i \left[n_i^* \ln\left(\frac{g_i}{n_i^*} - a\right) - \frac{g_i}{a} \ln\left(1 - a\frac{n_i^*}{g_i}\right)\right]$$

$$= \sum_i \left[n_i^*(\alpha + \beta\varepsilon_i) + \frac{g_i}{a} \ln\left\{1 + a e^{-\alpha - \beta\varepsilon_i}\right\}\right]. \tag{19}$$

The first sum on the right-hand side of (19) is identically equal to αN while the second sum is identically equal to βE. For the third sum, therefore, we have

$$\frac{1}{a}\sum_i g_i \ln\left\{1 + a e^{-\alpha - \beta\varepsilon_i}\right\} = \frac{S}{k} - \alpha N - \beta E. \tag{20}$$

Now, the physical interpretation of the parameters α and β here is going to be precisely the same as in Sec. 4.3, namely

$$\alpha = -\frac{\mu}{kT}, \quad \beta = \frac{1}{kT}; \tag{21}$$

see Sec. 6.2. The right-hand side of eqn. (20) is, therefore, equal to

$$\frac{S}{k}+\frac{\mu N}{kT}-\frac{E}{kT}=\frac{G-(E-TS)}{kT}=\frac{PV}{kT}. \tag{22}$$

The thermodynamic pressure of the system is, therefore, given by

$$PV=\frac{kT}{a}\sum_i\left[g_i\ln\left\{1+ae^{-\alpha-\beta\varepsilon_i}\right\}\right]. \tag{23}$$

In the Maxwell–Boltzmann case ($a\to 0$), eqn. (23) takes the form

$$PV=kT\sum_i g_ie^{-\alpha-\beta\varepsilon_i}=kT\sum_i n_i^*=NkT, \tag{24}$$

which is the familiar equation of state of the classical ideal gas. Note that eqn. (24) for the Maxwell–Boltzmann case holds irrespective of the details of the energy spectrum ε_i.

It will be recognized that the expression $a^{-1}\sum_i[\ ]$ in eqn. (23), being equal to the thermodynamic quantity (PV/kT), ought to be identical with the q-potential of the ideal gas. One may, therefore, expect to obtain from this expression all the macroscopic properties of this system. However, we prefer to first develop the formal theory of an ideal gas in the canonical and grand canonical ensembles.

6.2. An ideal gas in other quantum-mechanical ensembles

In the canonical ensemble the thermodynamics of a given system is derived from its partition function:

$$Q_N(V,T)=\sum_E e^{-\beta E} \tag{1}$$

where E denotes the energy eigenvalues of the system while $\beta=1/kT$. Now, an energy value E can be expressed in terms of the single-particle energies ε; for instance,

$$E=\sum_\varepsilon n_\varepsilon\varepsilon, \tag{2}$$

where n_ε is the number of particles in the single-particle energy state ε. The values of the numbers n_ε must satisfy the condition

$$\sum_\varepsilon n_\varepsilon=N. \tag{3}$$

Equation (1) may then be written as

$$Q_N(V.T)={\sum_{\{n_\varepsilon\}}}' g\{n_\varepsilon\}e^{-\beta\sum\limits_\varepsilon n_\varepsilon\varepsilon} \tag{4}$$

where $g\{n_\varepsilon\}$ is the *statistical weight factor* appropriate to the distribution set $\{n_\varepsilon\}$ and the summation $\sum'$ goes over all distribution sets that conform to the restrictive condition (3). The statistical weight factor in diffe cases is given by

$$g_{\text{B.E.}}\{n_\varepsilon\} = 1, \tag{5}$$

$$g_{\text{F.D.}}\{n_\varepsilon\} = \begin{cases} 1 & \text{if all } n_\varepsilon = 0 \text{ or } 1 \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

and

$$g_{\text{M.B.}}\{n_\varepsilon\} = \prod_\varepsilon \frac{1}{n_\varepsilon!}. \tag{7}$$

Note that in the present treatment we are dealing with single-particle states as *individual* states, without requiring them to be grouped into cells; indeed, the weight factors (5), (6) and (7) follow straightforwardly from their respective predecessors (6.1.6), (6.1.8) and (6.1.11) by putting all $g_i = 1$.

First of all, we work out the Maxwell–Boltzmann case. Substituting (7) into (4), we get

$$\begin{aligned} Q_N(V,T) &= \sum_{\{n_\varepsilon\}}{}' \left[\left(\prod_\varepsilon \frac{1}{n_\varepsilon!}\right) \prod_\varepsilon \left(e^{-\beta\varepsilon}\right)^{n_\varepsilon}\right] \\ &= \frac{1}{N!} \sum_{\{n_\varepsilon\}}{}' \left[\frac{N!}{\prod\limits_\varepsilon n_\varepsilon!} \prod_\varepsilon \left(e^{-\beta\varepsilon}\right)^{n_\varepsilon}\right] \end{aligned} \tag{8}$$

Since the summation here is governed by condition (3), it can be evaluated with the help of the multinomial theorem, with the result

$$\begin{aligned} Q_N(V,T) &= \frac{1}{N!}\left[\sum_\varepsilon e^{-\beta\varepsilon}\right]^N \\ &= \frac{1}{N!}\left[Q_1(V,T)\right]^N, \end{aligned} \tag{9}$$

in agreement with eqn. (3.5.15). The evaluation of Q_1 is, of course, straightforward: one obtains, using the asymptotic formula (2.4.7) for the number of single-particle states with energies lying between ε and $\varepsilon + d\varepsilon$,

$$\begin{aligned} Q_1(V,T) &\equiv \sum_\varepsilon e^{-\beta\varepsilon} \approx \frac{2\pi V}{h^3}(2m)^{3/2} \int_0^\infty e^{-\beta\varepsilon} \varepsilon^{1/2}\, d\varepsilon \\ &= V/\lambda^3, \end{aligned} \tag{10}$$

where $\lambda\,[= h/(2\pi mkT)^{1/2}]$ is the mean thermal wavelength of the particles. Hence

$$Q_N(V,T) = \frac{V^N}{N!\lambda^{3N}}, \tag{11}$$

from which complete thermodynamics of this system can be derived; see, for example, Sec. 3.5. Further, we obtain for the grand partition function of this system

$$\mathcal{Q}(z,V,T) = \sum_{N=0}^{\infty} z^N Q_N(V,T) = \exp\left(zV/\lambda^3\right); \tag{12}$$

cf. eqn. (4.4: We know that the thermodynamics of the system follows equally well from the expression for $\mathscr{Q}$.

In the Bose–Einstein and Fermi–Dirac cases we obtain, by substituting (5) and (6) into (4),

$$Q_N(V,T) = {\sum_{\{n_\varepsilon\}}}' \left(e^{-\beta \sum_\varepsilon n_\varepsilon \varepsilon} \right); \tag{13}$$

the difference between the two cases, B.E. and F.D., arises from the values that the number n_ε can take. Now, in view of restriction (3) on the summation $\sum'$, an explicit evaluation of the partition function Q_N is rather cumbersome. The grand partition function $\mathscr{Q}$, on the other hand, turns out to be more easily tractable; we have

$$\mathscr{Q}(z,V,T) = \sum_{N=0}^{\infty} \left[z^N {\sum_{\{n_\varepsilon\}}}' e^{-\beta \sum_\varepsilon n_\varepsilon \varepsilon} \right] \tag{14a}$$

$$= \sum_{N=0}^{\infty} \left[{\sum_{\{n_\varepsilon\}}}' \prod_\varepsilon \left(z e^{-\beta\varepsilon} \right)^{n_\varepsilon} \right] \tag{14b}$$

Now, the double summation in (14b) (first over the numbers n_ε constrained by a *fixed* value of the total number N, and then over all possible values of N) is equivalent to a summation over all possible values of the numbers n_ε, *independently of one another*. Hence, we may write

$$\mathscr{Q}(z,V,T) = \sum_{n_0, n_1, \ldots} \left[\left(z e^{-\beta\varepsilon_0} \right)^{n_0} \left(z e^{-\beta\varepsilon_1} \right)^{n_1} \cdots \right]$$

$$= \left[\sum_{n_0} \left(z e^{-\beta\varepsilon_0} \right)^{n_0} \right] \left[\sum_{n_1} \left(z e^{-\beta\varepsilon_1} \right)^{n_1} \right] \cdots \tag{15}$$

Now, in the Bose–Einstein case the n_ε can be either 0 or 1 or 2 or …, while in the Fermi–Dirac case they can be only 0 or 1. Therefore,

$$\mathscr{Q}(z,V,T) = \begin{cases} \displaystyle\prod_\varepsilon \frac{1}{(1 - z e^{-\beta\varepsilon})} & \text{in the B.E. case, with } z e^{-\beta\varepsilon} < 1 \quad (16a) \\ \displaystyle\prod_\varepsilon (1 + z e^{-\beta\varepsilon}) & \text{in the F.D. case.} \quad (16b) \end{cases}$$

The q-potential of the system is thus given by

$$q(z,V,T) \equiv \frac{PV}{kT} \equiv \ln \mathscr{Q}(z,V,T)$$

$$= \mp \sum_\varepsilon \ln \left(1 \mp z e^{-\beta\varepsilon} \right); \tag{17}$$

cf. eqn. (6.1.23), with all $g_i = 1$. The identification of the fugacity z with the quantity $e^{-\alpha}$ of eqn. (6.1.23) is quite natural; accordingly, $\alpha = -\mu/kT$. As usual, the upper (lower) sign in eqn. (17) corresponds to the Bose (Fermi) case.

In the end, we may write our results for q in a form applicable to all three cases:

$$q(z, V, T) \equiv \frac{PV}{kT} = \frac{1}{a}\sum_{\varepsilon} \ln\left(1 + aze^{-\beta\varepsilon}\right), \tag{18}$$

where $a = -1$, $+1$ or 0, depending upon the statistics governing the system. In particular, the classical case ($a \to 0$) gives

$$q_{\text{M.B.}} = z\sum_{\varepsilon} e^{-\beta\varepsilon} = zQ_1, \tag{19}$$

in agreement with eqn. (4.4.4). From (18) it follows that

$$\overline{N} \equiv z\left(\frac{\partial q}{\partial z}\right)_{V,T} = \sum_{\varepsilon} \frac{1}{z^{-1}e^{\beta\varepsilon} + a} \tag{20}$$

and

$$\overline{E} \equiv -\left(\frac{\partial q}{\partial \beta}\right)_{z,V} = \sum_{\varepsilon} \frac{\varepsilon}{z^{-1}e^{\beta\varepsilon} + a}. \tag{21}$$

At the same time, the *mean occupation number* $\langle n_\varepsilon \rangle$ of level ε turns out to be, see eqns (14a) and (17),

$$\begin{aligned}\langle n_\varepsilon \rangle &= \frac{1}{\mathcal{Q}}\left[-\frac{1}{\beta}\left(\frac{\partial \mathcal{Q}}{\partial \varepsilon}\right)_{z,T,\ \text{all other } \varepsilon}\right] \\ &\equiv -\frac{1}{\beta}\left(\frac{\partial q}{\partial \varepsilon}\right)_{z,T,\ \text{all other } \varepsilon} \\ &= \frac{1}{z^{-1}e^{\beta\varepsilon} + a},\end{aligned} \tag{22}$$

in keeping with eqns (20) and (21). Comparing our final result (22) with its counterpart (6.1.18a), we find that the *mean* value $\langle n \rangle$ and the *most probable* value n^* of the occupation number n of a single-particle state are identical.

6.3. Statistics of the occupation numbers

Equation (6.2.22) gives the *mean occupation number* of a single-particle state with energy ε as an explicit function of the quantity $(\varepsilon - \mu)/kT$:

$$\langle n_\varepsilon \rangle = \frac{1}{e^{(\varepsilon-\mu)/kT} + a}. \tag{1}$$

The functional behavior of this number is shown in Fig. 6.2. In the Fermi–Dirac case ($a = +1$) the mean occupation number never exceeds unity, for the variable n_ε itself cannot have a value other than 0 or 1. Moreover, for $\varepsilon < \mu$ and $|\varepsilon - \mu| \gg kT$, the mean occupation number tends to its maximum possible value 1. In the Bose–Einstein case ($a = -1$), we must have $\mu <$ all ε; see eqn. (6.2.16a). In fact, when μ becomes equal to the lowest value of ε, say ε_0, the occupancy of that particular level becomes infinitely high, which leads to the phenomenon of *Bose–Einstein condensation* (see Sec. 7.1). For $\mu < \varepsilon_0$, all values

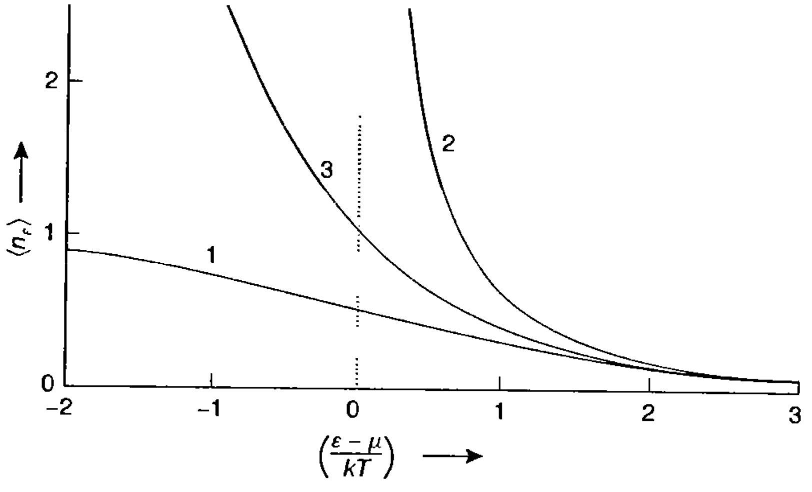


FIG. 6.2. The mean occupation number $\langle n_\varepsilon \rangle$ of a single-particle energy state ε in a system of non-interacting particles: curve 1 is for fermions, curve 2 for bosons and curve 3 for the Maxwell–Boltzmann particles.

of $(\varepsilon-\mu)$ are positive and the behavior of all $\langle n_\varepsilon \rangle$ is nonsingular. Finally, in the Maxwell–Boltzmann case $(a=0)$, the mean occupation number takes the familiar form

$$\langle n_\varepsilon \rangle_{\text{M.B.}} = \exp\{(\mu-\varepsilon)/kT\} \propto \exp(-\varepsilon/kT). \tag{2}$$

The important thing to note here is that the distinction between the quantum statistics $(a = \mp 1)$ and the classical statistics $(a = 0)$ becomes imperceptible when, for all values of ε that are of practical interest,

$$\exp\{(\varepsilon-\mu)/kT\} \gg 1. \tag{3}$$

In that event, eqn. (1) essentially reduces to (2) and we may write, instead of (3),

$$\langle n_\varepsilon \rangle \ll 1. \tag{4}$$

Condition (4) is quite understandable, for it implies that the probability of any of the n_ε being greater than unity is quite negligible, with the result that the classical weight factors $g\{n_\varepsilon\}$, as given by eqn. (6.2.7), become essentially equal to 1. The distinction between the classical treatment and the quantum-mechanical treatment then becomes physically insignificant. Correspondingly, we find, see Fig. 6.2, that for large values of $(\varepsilon-\mu)/kT$ the quantum curves 1 and 2 essentially merge into the classical curve 3. Since we already know that the higher the temperature of the system the better the validity of the classical treatment, condition (3) also implies that μ, the chemical potential of the system, must be negative and large in magnitude. This means that the fugacity $z[\equiv \exp(\mu/kT)]$ of the system must be much smaller than unity; see also eqn. (6.2.22). One can see, from eqns (4.4.6) and (4.4.29), that this is further equivalent to the requirement

$$\frac{N\lambda^3}{V} \ll 1, \tag{5}$$

which agrees with condition (5.5.20).

We shall now examine statistical fluctuations in variable n_ε. Going a step further from the calculation that led to eqn. (6.2.22), we have

$$\langle n_\varepsilon^2 \rangle = \frac{1}{\mathcal{Q}} \left[\left(-\frac{1}{\beta} \frac{\partial}{\partial \varepsilon} \right)^2 \mathcal{Q} \right]_{z,T,\text{ all other } \varepsilon}; \tag{6}$$

it follows that

$$\begin{aligned} \langle n_\varepsilon^2 \rangle - \langle n_\varepsilon \rangle^2 &= \left[\left(-\frac{1}{\beta} \frac{\partial}{\partial \varepsilon} \right)^2 \ln \mathcal{Q} \right]_{z,T,\text{ all other } \varepsilon} \\ &= \left[\left(-\frac{1}{\beta} \frac{\partial}{\partial \varepsilon} \right) \langle n_\varepsilon \rangle \right]_{z,T} \end{aligned} \tag{7}$$

For the relative mean-square fluctuation, we obtain (irrespective of the statistics obeyed by the particles)

$$\frac{\langle n_\varepsilon^2 \rangle - \langle n_\varepsilon \rangle^2}{\langle n_\varepsilon \rangle^2} = \left(\frac{1}{\beta} \frac{\partial}{\partial \varepsilon} \right) \left\{ \frac{1}{\langle n_\varepsilon \rangle} \right\} = z^{-1} e^{\beta\varepsilon}; \tag{8}$$

of course, the actual value of this quantity will depend upon the statistics of the particles because, for a given particle density (N/V) and a given temperature T, the value of z will be different for different statistics.

It seems more instructive to write (8) in the form

$$\frac{\langle n_\varepsilon^2 \rangle - \langle n_\varepsilon \rangle^2}{\langle n_\varepsilon \rangle^2} = \frac{1}{\langle n_\varepsilon \rangle} - a. \tag{9}$$

In the classical case ($a = 0$), the relative fluctuation is *normal*, so to say. In the Fermi–Dirac case, it is given by $1/\langle n_\varepsilon \rangle - 1$, which is below normal and tends to vanish as $\langle n_\varepsilon \rangle \to 1$. In the Bose–Einstein case, the fluctuation is clearly *above normal*.[2] Obviously, this result would apply to a gas of photons and, hence, to the oscillator states in the black-body radiation. In the latter context, Einstein derived this result as early as 1909 following Planck's approach and even pointed out that the term 1 in the expression for the fluctuation may be attributed to the wave character of the radiation and the term $1/\langle n_\varepsilon \rangle$ to the particle character of the photons; for details, see Kittel (1958), ter Haar (1968). Closely related to the subject of fluctuations is the problem of "statistical correlations in photon beams" which have been observed experimentally (see Hanbury Brown and Twiss, 1956–8) and have been explained theoretically in terms of the quantum-statistical nature of these fluctuations (see Purcell, 1956; Kothari and Auluck, 1957). For further details, refer to Mandel, Sudarshan and Wolf (1964); Holliday and Sage (1964).

For greater understanding of the statistics of the occupation numbers, we evaluate the quantity $p_\varepsilon(n)$, which is the probability that there are exactly n particles in the state of energy ε. Referring to eqn. (6.2.14b), we infer that $p_\varepsilon(n) \propto (ze^{-\beta\varepsilon})^n$. On normalization, it becomes in the Bose–Einstein case

$$\begin{aligned} p_\varepsilon(n)|_{\text{B.E.}} &= \left(ze^{-\beta\varepsilon}\right)^n \left[1 - ze^{-\beta\varepsilon}\right] \\ &= \left(\frac{\langle n_\varepsilon \rangle}{\langle n_\varepsilon \rangle + 1} \right)^n \frac{1}{\langle n_\varepsilon \rangle + 1} = \frac{(\langle n_\varepsilon \rangle)^n}{(\langle n_\varepsilon \rangle + 1)^{n+1}}. \end{aligned} \tag{10}$$

In the Fermi–Dirac case, we get

$$\begin{aligned} p_\varepsilon(n)|_{\text{F.D.}} &= \left(ze^{-\beta\varepsilon}\right)^n \left[1+ze^{-\beta\varepsilon}\right]^{-1} \\ &= \begin{cases} 1-\langle n_\varepsilon\rangle & \text{for} \quad n=0 \\ \langle n_\varepsilon\rangle & \text{for} \quad n=1. \end{cases} \end{aligned} \tag{11}$$

In the Maxwell–Boltzmann case, $p_\varepsilon(n) \propto (ze^{-\beta\varepsilon})^n/n!$ instead; see eqn. (6.2.8). On normalization, we get

$$p_\varepsilon(n)|_{\text{M.B.}} = \frac{\left(ze^{-\beta\varepsilon}\right)^n/n!}{\exp\left(ze^{-\beta\varepsilon}\right)} = \frac{(\langle n_\varepsilon\rangle)^n}{n!}e^{-\langle n_\varepsilon\rangle} \tag{12}$$

Distribution (12) is clearly a *Poisson distribution*, for which the mean square deviation of the variable is equal to the mean value itself; cf. eqn. (9), with $a=0$. It also resembles the distribution of the total particle number N in a grand canonical ensemble consisting of ideal, classical systems; see Problem 4.4. We also note that the ratio $p_\varepsilon(n)/p_\varepsilon(n-1)$ in this case varies inversely with n, which is "normal" statistical behavior of uncorrelated events. On the other hand, the distribution in the Bose–Einstein case is *geometric*, with a constant common ratio $\langle n_\varepsilon\rangle/(\langle n_\varepsilon\rangle + 1)$. This means that the probability of a state ε acquiring one more particle for itself is independent of the number of particles already occupying that state; thus, in comparison with the "normal" statistical behavior, bosons exhibit a special tendency of "bunching" together, which means a *positive* statistical correlation. In contrast, fermions exhibit a *negative* statistical correlation.

6.4. Kinetic considerations

The thermodynamic pressure of an ideal gas is given by eqn. (6.1.23) or (6.2.18). In view of the largeness of volume V, the single-particle energy states ε would be so close to one another that a summation over them may be replaced by an integration. One thereby gets

$$\begin{aligned} P &= \frac{kT}{a}\int_0^\infty \ln\left[1+aze^{-\beta\varepsilon(p)}\right]\frac{4\pi p^2\,dp}{h^3} \\ &= \frac{4\pi kT}{ah^3}\left[\frac{p^3}{3}\ln\left[1+aze^{-\beta\varepsilon(p)}\right]\Big|_0^\infty + \int_0^\infty \frac{p^3}{3}\frac{aze^{-\beta\varepsilon(p)}}{1+aze^{-\beta\varepsilon(p)}}\beta\frac{d\varepsilon}{dp}\,dp\right] \end{aligned}$$

The integrated part vanishes at both limits while the rest of the expression reduces to

$$P = \frac{4\pi}{3h^3}\int_0^\infty \frac{1}{z^{-1}e^{\beta\varepsilon(p)}+a}\left(p\frac{d\varepsilon}{dp}\right)p^2\,dp. \tag{1}$$

Now, the total number of particles in the system is given by

$$N = \int \langle n_p\rangle \frac{V\,d^3p}{h^3} = \frac{4\pi V}{h^3}\int_0^\infty \frac{1}{z^{-1}e^{\beta\varepsilon(p)}+a}p^2\,dp. \tag{2}$$

Comparing (1) and (2), we can write

$$P = \frac{1}{3}\frac{N}{V}\left\langle p\frac{d\varepsilon}{dp}\right\rangle = \frac{1}{3}n\langle pu\rangle, \tag{3}$$

where n is the particle density in the gas and u the speed of an individual particle. If the relationship between the energy ε and the momentum p is of the form $\varepsilon \propto p^s$, then

$$P = \frac{s}{3}n\langle\varepsilon\rangle = \frac{s}{3}\frac{E}{V}; \tag{4}$$

The particular cases $s = 1$ and $s = 2$ are pretty easy to recognize. It should be noted here that the results (3) and (4) hold independently of the statistics obeyed by the particles.

The structure of formula (3) suggests that the pressure of the gas arises essentially from the physical motion of the particles; it should, therefore, be derivable from kinetic considerations alone. To do this, we consider the bombardment, by the particles of the gas, on the walls of the container. Let us take, for example, an element of area dA on one of the walls normal to the z-axis, see Fig. 6.3, and focus our attention on those particles whose velocity lies between $\mathbf{u}$ and $\mathbf{u} + d\mathbf{u}$; the number of such particles per unit volume may be denoted by $nf(\mathbf{u})d\mathbf{u}$, where

$$\int_{\text{all } \mathbf{u}} f(\mathbf{u})d\mathbf{u} = 1. \tag{5}$$

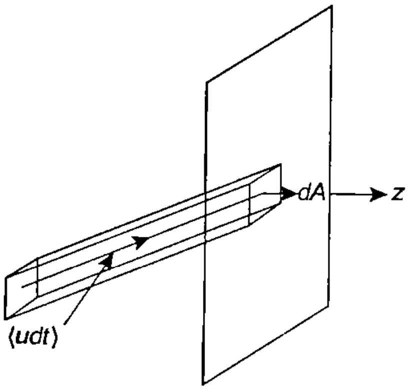


FIG. 6.3. The molecular bombardment on one of the walls of the container.

Now, the question is: how many of these particles will strike the area dA in time dt? The answer is: all those that happen to lie in a cylindrical region of base dA and height $u\,dt$, as shown in Fig. 6.3. Since the volume of this region is $(d\mathbf{A} \cdot \mathbf{u})dt$, the relevant number of particles would be $\{(d\mathbf{A} \cdot \mathbf{u})dt \times nf(\mathbf{u})d\mathbf{u}\}$. On reflection from the wall, the normal component of the momentum of a particle would undergo a change from p_z to $-p_z$; as a result, the normal momentum imparted by these particles per unit time to a unit area of the wall would be $2p_z\{u_z nf(\mathbf{u})d\mathbf{u}\}$. Integrating this expression over all relevant $\mathbf{u}$, we obtain the total normal momentum imparted per unit time to a unit area of the wall by all the particles of the gas, which is, by definition, the *kinetic* pressure of the gas:

$$P = 2n \int_{u_x=-\infty}^{\infty} \int_{u_y=-\infty}^{\infty} \int_{u_z=0}^{\infty} p_z u_z f(u)\, du_x\, du_y\, du_z. \qquad (6)^3$$

Since (i) $f(u)$ is a function of u alone and (ii) the product $(p_z u_z)$ is an even function of u_z, the foregoing result may be written as

$$P = n \int_{\text{all } u} (p_z u_z) f(u)\, d\mathbf{u}. \qquad (7)$$

Comparing (7) with (5), we obtain

$$P = n\langle p_z u_z\rangle = n\langle pu\cos^2\theta\rangle \qquad (8)$$

$$= \tfrac{1}{3}n\langle pu\rangle, \qquad (9)$$

which is identical with (3).

In a similar manner, we can determine the rate of *effusion* of the gas particles through a hole (of unit area) in the wall. This is given by, cf. (6),

$$R = n \int_{u_x=-\infty}^{\infty} \int_{u_y=-\infty}^{\infty} \int_{u_z=0}^{\infty} u_z f(u)\, du_x\, du_y\, du_z \qquad (10)$$

$$= n \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi/2} \int_{u=0}^{\infty} \{u\cos\theta f(u)\}(u^2 \sin\theta\, du\, d\theta\, d\phi); \qquad (11)$$

note that the condition $u_z > 0$ restricts the range of the angle θ between the values 0 and $\pi/2$. Carrying out integrations over θ and ϕ, we obtain

$$R = n\pi \int_0^{\infty} f(u)u^3\, du. \qquad (12)$$

In view of the fact that

$$\int_0^{\infty} f(u)(4\pi u^2\, du) = 1, \qquad (5a)$$

eqn. (12) may be written as

$$R = \tfrac{1}{4}n\langle u\rangle. \qquad (13)$$

Again, this result holds independently of the statistics obeyed by the particles.

It is obvious that the velocity distribution among the effused particles is considerably different from the one among the particles inside the container. This is due to the fact that, firstly, the velocity component u_z of the effused particles must be positive (which introduces an element of anisotropy into the distribution) and, secondly, the particles with larger values of u_z appear with an extra weightage, the weightage being directly proportional to the value of u_z; see eqn. (10). As a result of this, (i) the effused particles carry with them a net forward momentum, thus causing the container to experience a recoil force, and (ii) they carry away a relatively larger amount of energy per particle, thus leaving the gas in the container at not only a progressively decreasing pressure and density but also a progressively decreasing temperature; see Problem 6.13.

6.5. Gaseous systems composed of molecules with internal motion

In most of our studies so far we have considered only the translational part of the molecular motion. Though this aspect of motion is invariably present in a gaseous system, other aspects, which are essentially concerned with the *internal* motion of the molecules, also exist. It is only natural that in the calculation of the physical properties of such a system, contributions arising from these motions are also taken into account. In doing so, we shall assume here that (i) the effects of the intermolecular interactions are negligible and (ii) the nondegeneracy criterion

$$\frac{nh^3}{(2\pi mkT)^{3/2}} \ll 1 \tag{5.5.20}$$

is fulfilled; this makes our system an *ideal, Boltzmannian* gas. Under these assumptions, which hold sufficiently well in a large number of applications, the partition function of the system is given by

$$Q_N(V,T) = \frac{1}{N!}[Q_1(V,T)]^N, \tag{1}$$

where

$$Q_1(V,T) = \left\{\frac{V}{h^3}(2\pi mkT)^{3/2}\right\} j(T); \tag{2}$$

the factor within the curly brackets is the familiar translational partition function of a molecule, while $j(T)$ is the partition function corresponding to internal motions. The latter may be written as

$$j(T) = \sum_i g_i e^{-\varepsilon_i/kT}, \tag{3}$$

where ε_i is the energy associated with a state of internal motion (characterized by the quantum numbers i), while g_i is the multiplicity of that state.

The contributions made by the internal motions of the molecules to the various thermodynamic properties of the system follow straightforwardly from the function $j(T)$. We obtain

$$A_{\text{int}} = -NkT \ln j, \tag{4}$$

$$\mu_{\text{int}} = -kT \ln j, \tag{5}$$

$$S_{\text{int}} = NK\left(\ln j + T\frac{\partial}{\partial T}\ln j\right) \tag{6}$$

$$U_{\text{int}} = NkT^2\frac{\partial}{\partial T}\ln j \tag{7}$$

and

$$(C_V)_{\text{int}} = Nk\frac{\partial}{\partial T}\left\{T^2\frac{\partial}{\partial T}\ln j\right\}. \tag{8}$$

Thus, the central problem in this study is to derive an explicit expression for the function $j(T)$ from a knowledge of the internal states of the molecules. For this purpose, we note that the internal state of a molecule is determined by (i) the electronic sta. , (ii) the state of the nuclei, (iii) the vibrational state and (iv) the rotational state. Rigorously speaking, these four modes of excitation mutually interact; in many cases, however, they can be treated independently of one another. We can then write

$$j(T) = j_{\text{elec}}(T) j_{\text{nuc}}(T) j_{\text{vib}}(T) j_{\text{rot}}(T), \tag{3a}$$

with the result that the net contribution made by the internal motions to the various thermodynamic quantities of the system is given by a simple sum of the four respective contributions. There is one interaction, however, which plays a special role in the case of homonuclear molecules, such as AA, and that is between the states of the nuclei and the rotational states. In such a case, we may write

$$j(T) = j_{\text{elec}}(T) j_{\text{nuc-rot}}(T) j_{\text{vib}}(T). \tag{3b}$$

We now examine this problem for various systems in the order of increasing complexity.

A. Monatomic molecules

For simplicity, we consider a monatomic gas at temperatures such that the thermal energy kT is small in comparison with the ionization energy E_{ion}; for different atoms, this amounts to the condition: $T \ll E_{\text{ion}}/k \sim 10^4\text{–}10^5$ K. At these temperatures the number of ionized atoms in the gas would be insignificant. The same would be true of atoms in the excited states, for the separation of any of the excited states from the ground state of the atom is generally of the same order of magnitude as the ionization energy itself. Thus, we may regard all atoms in the gas to be in their (electronic) ground state.

Now, there is a special class of atoms, namely He, Ne, A, ..., which, in their ground state, possess neither orbital angular momentum nor spin ($L = S = 0$). Their (electronic) ground state is clearly a singlet: $g_e = 1$. The nucleus, however, possesses a degeneracy that arises from the possibility of different orientations of the nuclear spin.[4] If the value of this spin is S_n, the corresponding degeneracy factor $g_n = 2S_n + 1$. Moreover, a monatomic molecule is incapable of having any vibrational or rotational states. The internal partition function (3a) of such a molecule is, therefore, given by

$$j(T) = (g)_{\text{gr. st.}} = g_e \cdot g_n = 2S_n + 1. \tag{9}$$

Equations (4)–(8) then tell us that the internal motions in this case contribute only towards properties such as the chemical potential and the entropy of the gas; they do not contribute towards the internal energy and the specific heat.

If, on the other hand, the ground state does not possess orbital angular momentum but possesses spin ($L = 0$, $S \neq 0$–as, for example, in the case of alkali atoms), then the ground state will still have no fine structure; it will, however, have a degeneracy $g_e = 2S + 1$. As a result, the internal partition function $j(T)$ will get multiplied by a factor of $(2S + 1)$ and the properties such as the chemical potential and the entropy of the gas will get modified accordingly.

In other cases, the ground state of the atom may possess both orbital angular momentum and spin ($L \neq 0$, $S \neq 0$); the ground state would then possess a definite fine structure. The intervals of this structure are, in general, comparable with kT; hence, in the evaluation of the partition function, the energies of the various components of the fine structure will have to be taken into account Since these components differ from one another in the value of the total angular momentum J, the relevant partition function may be written as

$$j_{\text{elec}}(T) = \sum_J (2J+1)e^{-\varepsilon_J/kT}. \tag{10}$$

The foregoing expression simplifies considerably in the following limiting cases:

(a) $kT \gg$ all ε_J; then

$$j_{\text{elec}}(T) \simeq \sum_J (2J+1) = (2L+1)(2S+1). \tag{10a}$$

(b) $kT \ll$ all ε_J; then

$$j_{\text{elec}}(T) \simeq (2J_0+1)e^{-\varepsilon_0/kT}, \tag{10b}$$

where J_0 is the total angular momentum, and ε_0 the energy, of the atom in the lowest state. In either case, the electronic motion makes no contribution towards the specific heat of the gas. Of course, at intermediate temperatures we do obtain a contribution towards this property. And, in view of the fact that both at high and low temperatures the specific heat tends to be equal to the translational value $\frac{3}{2}Nk$, it must pass through a maximum at a temperature comparable to the separation of the fine structure levels.[5] Needless to say, the multiplicity $(2S_n+1)$ introduced by the nuclear spin must be taken into account in each case.

B. Diatomic molecules

Now we consider a diatomic gas at temperatures such that kT is small compared with the energy of dissociation; for different molecules this amounts once again to the condition: $T \ll E_{\text{diss}}/k \sim 10^4$–$10^5$ K. At these temperatures the number of dissociated molecules in the gas would be insignificant. At the same time, in most cases, there would be practically no molecules in the excited states as well, for the separation of any of these states from the ground state of the molecule is in general comparable to the dissociation energy itself.[6] Accordingly, in the evaluation of $j(T)$, we have to take into account only the lowest electronic state of the molecule.

The lowest electronic state, in most cases, is non-degenerate: $g_e = 1$. We then need not consider any further the question of the electronic state making a contribution towards the thermodynamic properties of the gas. However, certain molecules (though not very many) have, in their lowest electronic state, either (i) a nonzero orbital angular momentum ($\Lambda \neq 0$) or (ii) a nonzero spin ($S \neq 0$) or (iii) both. In case (i), the electronic state acquires a twofold degeneracy corresponding to the two possible orientations of the oribital angular momentum relative to the molecular axis;[7] as a result, $g_e = 2$. In case (ii), the state acquires a degeneracy $2S+1$ corresponding to the space quantization of the spin.[8] In both these cases the chemical potential and the entropy of the gas are modified by the multiplicity of the electronic state, while the energy and the specific heat remain unaffected. In case (iii), we encounter a fine structure which necessitates a rather detailed study because the intervals of this structure are generally of the same order of magnitude as kT. In particular, for a doublet fine-structure term, such as the one that arises in the molecule NO ($\Pi_{1/2,3/2}$ with a separation of 178 K, the components themselves being Λ-doublets), we have for the electronic partition function

$$j_{\text{elec}}(T) = g_0 + g_1 e^{-\Delta/kT}, \tag{11}$$

where g_0 and g_1 are the degeneracy factors of the two components while Δ is their separation energy. The contribution made by (11) towards the various thermodynamic properties of the gas can be calculated with the help of formulae (4)–(8). In particular, we obtain for the contribution towards the specific heat

$$(C_V)_{\text{elec}} = Nk \frac{(\Delta/kT)^2}{\left[1 + (g_0/g_1)e^{\Delta/kT}\right]\left[1 + (g_1/g_0)e^{-\Delta/kT}\right]} \tag{12}$$

We note that this contribution vanishes both for $T \ll \Delta/k$ and for $T \gg \Delta/k$ and is maximum for a certain temperature $\sim \Delta/k$; cf. the corresponding situation in the case of monatomic molecules.

We now consider the effect of the vibrational states of the molecules on the thermodynamic properties of the gas. To have an idea of the temperature range over which this effect would be significant, we note that the magnitude of the corresponding quantum of energy, namely $\hbar\omega$, for different diatomic gases is of order 10^3 K. Thus, we would obtain full contributions (consistent with the dictates of the equipartition theorem) at temperatures of the order of 10^4 K or more, and practically no contribution at temperatures of the order of 10^2 K or less. Let us assume that the temperature is not high enough to excite vibrational states of large energy; the oscillations of the nuclei then remain small in amplitude and hence harmonic. The energy levels for a mode of frequency ω are then given by the well-known expression $(n + \frac{1}{2})\hbar\omega$.[9] The evaluation of the vibrational partition function $j_{\text{vib}}(T)$ is quite elementary; see Sec. 3.8. In view of the rapid convergence of the series involved, the summation may formally be extended to $n = \infty$. The corresponding contributions towards the various thermodynamic properties of the system are then given by eqns (3.8.16) to (3.8.21). In particular,

$$(C_V)_{\text{vib}} = Nk\left(\frac{\Theta_v}{T}\right)^2 \frac{e^{\Theta_v/T}}{\left(e^{\Theta_v/T} - 1\right)^2}; \quad \Theta_v = \frac{\hbar\omega}{k}. \tag{13}$$

We note that for $T \gg \Theta_v$, the vibrational specific heat is very nearly equal to the equipartition value Nk; otherwise, it is always less than Nk. In particular, for $T \ll \Theta_v$, the specific heat tends to zero (see Fig. 6.4); the vibrational degrees of freedom are then said to be "frozen".

At sufficiently high temperatures, when vibrations with large n are also excited, the effects of anharmonicity and of interaction between the vibrational and the rotational modes of the molecule can become important.[10] However, since this happens only at large n, the relevant corrections to the various thermodynamic quantities can b[illegible]determined even classically; see Problems 3.29 and 3.30. One finds that the fir[illegible]der correction to C_{vib} is directly proportional to the temperature of the gas.

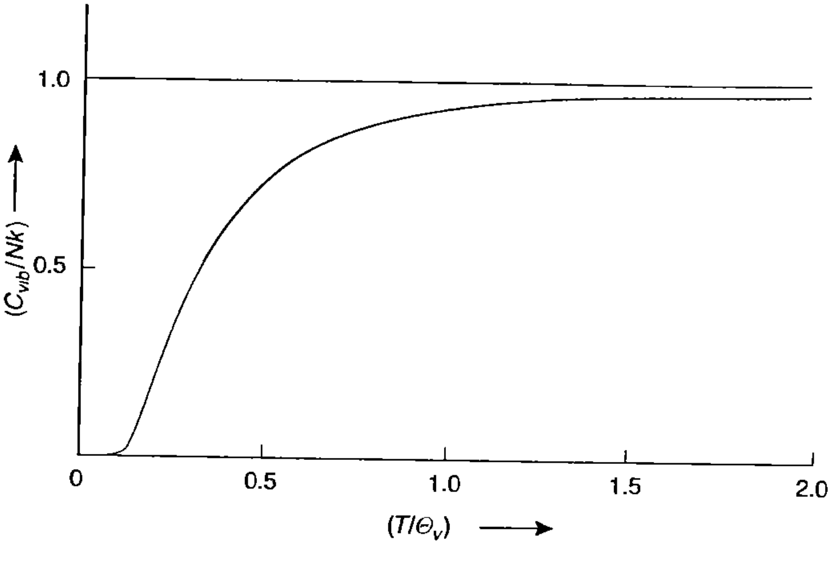


FIG. 6.4. The vibrational specific heat of a gas of diatomic molecules. At $T = \Theta_v$, the specific heat is already about 93 per cent of the equipartition value.

Finally, we consider the effect of (i) the states of the nuclei and (ii) the rotational states of the molecule; wherever necessary, we shall take into account the mutual interaction of these modes. This interaction is of no relevance in the case of *heteronuclear* molecules, such as AB; it is, however, important in the case of *homonuclear* molecules, such as AA. We may, therefore, consider the two cases separately.

The states of the nuclei in the heteronuclear case may be treated separately from the rotational states of the molecule. Proceeding in the same manner as for monatomic molecules, we conclude that the effect of the nuclear states is adequately taken care of through a degeneracy factor g_n. Denoting the spins of the two nuclei by S_A and S_B,

$$g_n = (2S_A + 1)(2S_B + 1). \tag{14}$$

As before, we obtain a finite contribution towards the chemical potential and the entropy of the gas but none towards the internal energy and the specific heat.

Now, the rotational levels of a linear "rigid" rotator, with two degrees of freedom (for the axis of rotation) and the principal moments of inertia $(I, I, 0)$, are given by

$$\varepsilon_{\text{rot}} = l(l+1)\hbar^2/2I, \quad l = 0, 1, 2, \ldots; \tag{15}$$

here, $I = \mu r_0^2$ where $\mu[= m_1 m_2/(m_1 + m_2)]$ is the *reduced mass* of the nuclei and r_0 the *equilibrium distance* between them. The rotational partition function of the molecule is then given by

$$\begin{aligned} j_{\text{rot}}(T) &= \sum_{l=0}^{\infty} (2l+1) \exp\left\{-l(l+1)\frac{\hbar^2}{2IkT}\right\} \\ &= \sum_{l=0}^{\infty} (2l+1) \exp\left\{-l(l+1)\frac{\Theta_r}{T}\right\}; \quad \Theta_r = \frac{\hbar^2}{2Ik}. \end{aligned} \tag{16}$$

The values of ,, for all gases except the ones involving the isotopes H and D, are much smaller than the room temperature. For example, the value of Θ_r for HCl is about 15 K, for N_2, O_2 and NO it lies between 2 and 3 K, while for Cl_2 it is about one-third of a degree. On the other hand, the values of Θ_r for H_2, D_2 and HD are, respectively, 85 K, 43 K and 64 K. These numbers give us an idea of the respective temperature ranges in which the effects arising from the *discreteness* of the rotational states are expected to be important.

For $T \gg \Theta_r$, the spectrum of the rotational states may be approximated by a continuum. The summation in (16) is then replaced by an integration:

$$j_{\text{rot}}(T) \approx \int_0^\infty (2l+1)\exp\left\{-l(l+1)\frac{\Theta_r}{T}\right\} dl = \frac{T}{\Theta_r}. \tag{17}$$

The rotational specific heat is then given by

$$(C_V)_{\text{rot}} = Nk, \tag{18}$$

consistent with the equipartition theorem.

A better evaluation of the sum in (16) can be made with the help of the Euler–Maclaurin formula, viz.

$$\sum_{n=0}^{\infty} f(n) = \int_0^{\infty} f(x)\,dx + \tfrac{1}{2}f(0) - \tfrac{1}{12}f'(0) + \tfrac{1}{720}f'''(0) - \tfrac{1}{30240}f^{\text{v}}(0) + \cdots \tag{19}$$

Writing

$$f(x) = (2x+1)\exp\{-x(x+1)\Theta_r/T\},$$

one obtains

$$j_{\text{rot}}(T) = \frac{T}{\Theta_r} + \frac{1}{3} + \frac{1}{15}\frac{\Theta_r}{T} + \frac{4}{315}\left(\frac{\Theta_r}{T}\right)^2 + \cdots \tag{20}$$

which is the so-called *Mulholland's formula*; as expected, the main term of this formula is identical with the classical partition function (17). The corresponding result for the specific heat is

$$(C_V)_{\text{rot}} = Nk\left\{1 + \frac{1}{45}\left(\frac{\Theta_r}{T}\right)^2 + \frac{16}{945}\left(\frac{\Theta_r}{T}\right)^3 + \cdots\right\} \tag{21}$$

which shows that at high temperatures the rotational specific heat *decreases* with temperature and ultimately tends to the classical value Nk. Thus, at high (but finite) temperatures the rotational specific heat of a diatomic gas is greater than the classical value. On the other hand, it must go to zero as $T \to 0$. We therefore conclude that it passes through at least one maximum. Numerical studies show that there is only one maximum which appears at a temperature of about 0.8 Θ_r and has a value of about $1.1Nk$; see Fig. 6.5.

In the other limiting case, when $T \ll \Theta_r$, one may retain only the first few terms of the sum in (16); then

$$j_{\text{rot}}(T) = 1 + 3e^{-2\Theta_r/T} + 5e^{-6\Theta_r/T} + \cdots, \tag{22}$$

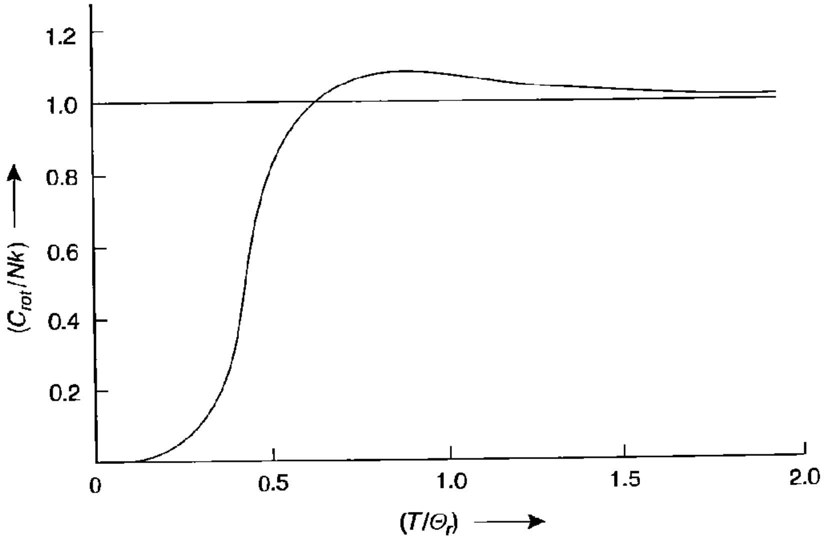


FIG. 6.5. The rotational specific heat of a gas of heteronuclear diatomic molecules.

whence one obtains, in the *lowest* approximation,

$$(C_V)_{\text{rot}} \simeq 12Nk\left(\frac{\Theta_r}{T}\right)^2 e^{-2\Theta_r/T}. \tag{23}$$

Thus, as $T \to 0$, the specific heat drops exponentially to zero; see again Fig. 6.5. We therefore conclude that at low enough temperatures the rotational degrees of freedom of the molecules are also "frozen".

At this stage it appears worthwhile to remark that, since the internal motions of the molecules do not make any contribution towards the pressure of the gas (A_{int} being independent of V), the quantity $(C_P - C_V)$ is the same for a diatomic gas as for a monatomic one. Moreover, under the assumptions made in the very beginning of this section, the value of this quantity at all temperatures of interest would be equal to the classical value Nk. Thus, at sufficiently low temperatures (when rotational as well as vibrational degrees of freedom of the molecules are "frozen"), we have, by virtue of the translational motion alone,

$$C_V = \tfrac{3}{2}Nk, \quad C_P = \tfrac{5}{2}NK; \quad \gamma = \tfrac{5}{3}. \qquad (24$$

As temperature rises, the rotational degrees of freedom begin to "loosen up" till we reach temperatures that are much larger than Θ_r but much smaller than Θ_V; the rotational degrees of freedom are then fully excited while the vibrational ones are still "frozen". Accordingly, for $\Theta_r \ll T \ll \Theta_V$,

$$C_V = \tfrac{5}{2}Nk, \quad C_P = \tfrac{7}{2}Nk; \quad \gamma = \tfrac{7}{5}. \qquad (25$$

As temperature rises further, the vibrational degrees of freedom as well start loosening up, till we reach temperatures that are much larger than Θ_V. Then, the vibrational degrees of freedom are also fully excited and we have

$$C_V = \tfrac{7}{2}Nk, \quad C_P = \tfrac{9}{2}Nk; \quad \gamma = \tfrac{9}{7}. \qquad (26$$

These features are displayed in Fig. 6.6 where the experimental results for C_P are plotted for three gases HD, HT and DT. We note that, in view of the considerable difference between the values of Θ_r and Θ_v, the situation depicted by (25) obtains over a considerably large range of temperatures. In passing, it may be pointed out that, for most diatomic gases, the situation at room temperatures corresponds to the one depicted by (25).

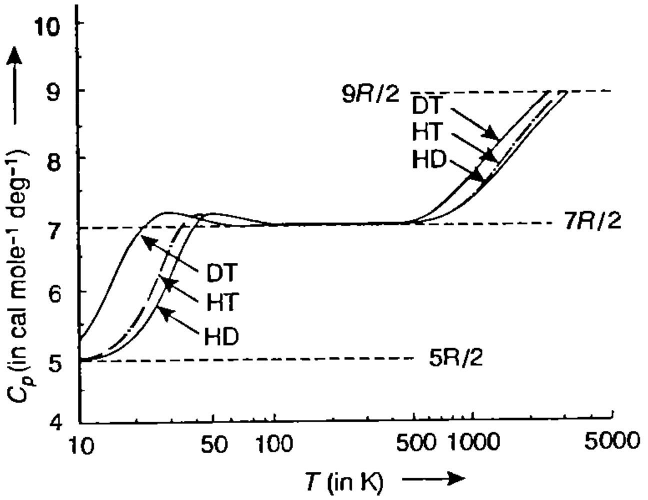


FIG. 6.6. The rotational-vibrational specific heat, C_P, of the diatomic gases HD, HT and DT.

We now study the case of *homonuclear* molecules, such as AA. To start with, we consider the limiting case of high temperatures where classical approximation is admissible. The rotational motion of the molecule may then be visualized as a rotation of the molecular axis, i.e. the line joining the two nuclei, about an "axis of rotation" which is perpendicular to the molecular axis and passes through the center of mass of the molecule. Then, the two opposing positions of the molecular axis, viz. the ones corresponding to the azimuthal angles ϕ and $\phi + \pi$, differ simply by an interchange of the two identical nuclei and, hence, correspond to only *one* distinct state of the molecule. Therefore, in the evaluation of the partition function, the range of the angle ϕ should be taken as $(0, \pi)$ instead of the customary $(0, 2\pi)$. Moreover, since the energy of rotational motion does not depend upon the angle ϕ, the only effect of this on the partition function of the molecule would be to reduce it by a factor of 2. We thus obtain, *in the classical approximation*,[11]

$$j_{\text{nuc-rot}}(T) = (2S_A + 1)^2 \frac{T}{2\Theta_r}. \tag{27}$$

Obviously, the factor 2 here will not affect the specific heat of the gas; in the classical approximation, therefore, the specific heat of a gas of homonuclear molecules is the same as that of a corresponding gas of heteronuclear molecules.

In contrast, significant changes result at relatively lower temperatures where the states of rotational motion have to be treated as *discrete*. These changes arise from the coupling between the nuclear and the rotational states which in turn arises from the symmetry character of the nuclear-rotational wave function. As discussed in Sec. 5.4, the [illegible]l wave function of a physical state must be either symmetric or antisymmetri[illegible] depending upon the statistics obeyed by the particles involved) with respect to an interchange of two identical particles. Now, the rotational wave function of a diatomic molecule is symmetric or antisymmetric according as the quantum number l is even or odd. The nuclear wave function, on the other hand, consists of a linear combination of the spin functions of the two nuclei and its symmetry character depends upon the manner in which the combination is formed. It is not difficult to see that, of the $(2S_A + 1)^2$ different combinations that one constructs, exactly $(S_A + 1)(2S_A + 1)$ are symmetric with respect to an interchange of the nuclei and the remaining $S_A(2S_A + 1)$ antisymmetric.[12] In constructing the total wave function, as a product of the nuclear and the rotational wave functions, we then proceed as follows:

(i) If the nuclei are fermions ($S_A = \frac{1}{2}, \frac{3}{2}, \ldots$), as in the molecule H_2, the total wave function must be antisymmetric. To secure this, we may associate any one of the $S_A(2S_A + 1)$ antisymmetric nuclear wave functions with any one of the even-l rotational wave functions *or* any one of the $(S_A + 1)(2S_A + 1)$ symmetric nuclear wave functions with any one of the odd-l rotational wave functions. Accordingly, the nuclear-rotational partition function of such a molecule would be

$$j^{(\text{F.D.})}_{\text{nuc-rot}}(T) = S_A(2S_A + 1)r_{\text{even}} + (S_A + 1)(2S_A + 1)r_{\text{odd}}, \tag{28}$$

where

$$r_{\text{even}} = \sum_{l=0,2,\ldots}^{\infty} (2l + 1)\exp\{-l(l + 1)\Theta_r/T\} \tag{29}$$

and

$$r_{\text{odd}} = \sum_{l=1,3,\ldots}^{\infty} (2l + 1)\exp\{-l(l + 1)\Theta_r/T\}. \tag{30}$$

(ii) If the nuclei are bosons ($S_A = 0, 1, 2, \ldots$), as in the molecule D_2, the total wave function must be symmetric. To secure this, we may associate any one of the $(S_A + 1)(2S_A + 1)$ symmetric nuclear wave functions with any one of the even-l rotational wave functions *or* any one of the $S_A(2S_A + 1)$ antisymmetric nuclear wave functions with any one of the odd-l rotational wave functions. We then have

$$j^{(\text{B.E.})}_{\text{nuc-rot}}(T) = (S_A + 1)(2S_A + 1)r_{\text{even}} + S_A(2S_A + 1)r_{\text{odd}}. \tag{31}$$

At high temperatures, it is the large values of l that contribute most to the sums (29) and (30). The difference between the two sums is then negligibly small, and we have

$$r_{\text{even}} \simeq r_{\text{odd}} \simeq \tfrac{1}{2} j_{\text{rot}}(T) = T/2\Theta_r; \tag{32}$$

see eqns (16) and (17). Consequently,

$$j^{(\text{B.E.})}_{\text{nuc-rot}} \simeq j^{(\text{F.D.})}_{\text{nuc-rot}} = (2S_A + 1)^2 T/2\Theta_r, \tag{33}$$

in agreement with our previous result (27). Under these circumstances, the statistics governing the nuclei does not make a significant difference to the thermodynamic behavior of the gas.

Things change when the temperature of the gas is in a range comparable to the value of Θ_r. It seems most reasonable then to regard the gas as a mixture of two components, generally referred to as *ortho-* and *para-*, whose relative concentrations in equilibrium are determined by the relative magnitudes of the two parts of the partition function (28) or (31), as the case may be. Customarily, the name ortho- is given to that component which carries the larger statistical weight. Thus, in the case of fermions (as in H_2), the ortho- to para- ratio is given by

$$n^{(\text{F.D.})} = \frac{(S_A + 1)r_{\text{odd}}}{S_A r_{\text{even}}}, \tag{34}$$

while in the case of bosons (as in D_2), the corresponding ratio is given by

$$n^{(\text{B.E.})} = \frac{(S_A + 1)r_{\text{even}}}{S_A r_{\text{odd}}}. \tag{35}$$

As temperature rises, the factor $r_{\text{odd}}/r_{\text{even}}$ tends to unity and the ratio n, in each case, approaches the temperature-independent value $(S_A + 1)/S_A$. In the case of H_2, this limiting value is 3 (since $S_A = \frac{1}{2}$) while in the case of D_2 it is 2 (since $S_A = 1$). At sufficiently low temperatures, one may retain only the main terms of the sums (29) and (30), with the result that

$$\frac{r_{\text{odd}}}{r_{\text{even}}} \simeq 3 \exp\left(-\frac{2\Theta_r}{T}\right) \quad (T \ll \Theta_r), \tag{36}$$

which tends to zero as $T \to 0$. The ratio n, then, tends to zero in the case of fermions and to infinity in the case of bosons. Hence, as $T \to 0$, the hydrogen gas is wholly para-, while deuterium is wholly ortho-; of course, in each case, the molecules settle down in the rotational state $l = 0$.

At intermediate temperatures, one has to work with the equilibrium ratio (34), or (35), and with the composite partition function (28), or (31), in order to compute the thermodynamic properties of the gas. One finds, however, that the theoretical results so derived do not *generally* agree with the ones obtained experimentally. The discrepancy was resolved by Dennison (1927) who pointed out that the samples of hydrogen, or deuterium, ordinarily subjected to experiment are not in thermal equilibrium as regards the relative magnitudes of the ortho- and para-components. These samples are ordinarily prepared and kept at room temperatures which are well above Θ_r, with the result that the ortho- to para- ratio in them is very nearly equal to the limiting value $(S_A + 1)S_A$. If now the temperature is lowered, one would expect the ratio to change in accordance with eqn. (34), or (35). However, it does not do so for the following reason. Since the transition of a molecule from one form of existence to another involves the flipping of the spin of one of its nuclei, the transition probability of the process is quite small. Actually, the periods involved are of the order of a year. Obviously, one cannot expect to attain the equilibrium ratio n during the short times available. Consequently, even at lower temperatures, what one generally has is a *non-equilibrium* mixture of two independent substances, the relative concentration of which is preassigned. The partition functions (28) and (31) as such are, therefore, inapplicable; we rather have directly for the specific heat

$$C^{(\text{F.D.})} = \frac{S_A}{2S_A + 1} C_{\text{even}} + \frac{S_A + 1}{2S_A + 1} C_{\text{odd}} \tag{37}$$

and

$$C^{(\mathrm{B.E.})} = \frac{S_A + 1}{2S_A + 1} C_{\mathrm{even}} + \frac{S_A}{2S_A + 1} C_{\mathrm{odd}}, \tag{38}$$

where

$$C_{\mathrm{even/odd}} = Nk \frac{\partial}{\partial T} \left\{ T^2 (\partial/\partial T) \ln r_{\mathrm{even/odd}} \right\}. \tag{39}$$

We have, therefore, to compute C_{even} and C_{odd} separately and then derive the net value of the rotational specific heat with the help of the formula (37) or (38), as the case may be. Figure 6.7 shows the relevant results for hydrogen. Curves 1 and 2 correspond to the para-hydrogen (C_{even}) and the ortho-hydrogen (C_{odd}), respectively, while curve 3 represents the weighted mean, as given by eqn. (37). The experimental results are also shown in the figure; the agreement between theory and experiment is clearly good.

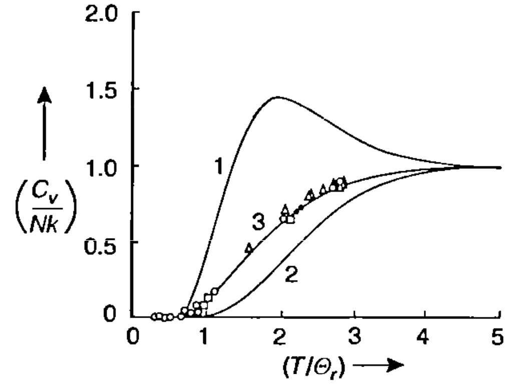


FIG. 6.7. The theoretical specific heat of a 1:3 mixture of para-hydrogen and ortho-hydrogen. The experimental points originate from various sources listed in Wannier (1966).

Further evidence in favor of Dennison's explanation is obtained by performing experiments with ortho–para mixtures of different relative concentration. This can be done by speeding up the ortho–para conversion by passing hydrogen over activated charcoal. By doing this at various temperatures, and afterwards removing the catalyst, one can fix the ratio n at any desired value. The specific heat then follows a curve obtained by mixing C_{even} and C_{odd} with appropriate weight factors. Further, if one measures the specific heat of the gas in such a way that the ratio n, at every temperature T, has the value given by the formula (34), it indeed follows the curve obtained from expression (28) for the partition function.

C. Polyatomic molecules

Once again, the translational degrees of freedom of the molecules contribute their usual share, $\frac{3}{2}k$ per molecule, towards the specific heat of the gas. As regards the lowest electronic state, it is, in most cases, far below any of the excited states; nevertheless, it generally possesses a multiplicity (d[illegible]nding upon the orbital and spin angular momenta of the state) which can be t[illegible]n care of by a degeneracy factor g_e. As regards the rotational states, they can be treated classically because the large values of the moments of inertia obtaining in polyatomic molecules make the quantum of the rotational energy, $\hbar^2/2I_i$, much smaller than the thermal energy, kT, at practically all temperatures of interest. Consequently, the interaction between the rotational states and the states of the nuclei can also be treated classically. As a result, the nuclear-rotational partition function is given by the product of the respective partition functions, divided by a symmetry number γ which denotes the number of physically indistinguishable configurations realized during one complete rotation of the molecule:[13]

$$j_{\text{nuc-rot}}(T) = \frac{g_{\text{nuc}} j^{C}_{\text{rot}}(T)}{\gamma}; \tag{40}$$

cf. eqn. (27). Here, $j^{C}_{\text{rot}}(T)$ is the rotational partition function of the molecule evaluated in the classical approximation (without paying regard to the presence of *identical* nuclei, if any); it is given by

$$j^{C}_{\text{rot}}(T) = \pi^{1/2}\left(\frac{2I_1kT}{\hbar^2}\right)^{1/2}\left(\frac{2I_2kT}{\hbar^2}\right)^{1/2}\left(\frac{2I_3kT}{\hbar^2}\right)^{1/2}, \tag{41}$$

where I_1, I_2 and I_3 are the principal moments of inertia of the molecule; see Problem 6.26.[14] The rotational specific heat is then given by

$$C_{\text{rot}} = Nk\frac{\partial}{\partial T}\left\{T^2\frac{\partial}{\partial T}\ln j^{C}_{\text{rot}}(T)\right\} = \frac{3}{2}Nk, \tag{42}$$

consistent with the equipartition theorem.

As regards the vibrational states, we first note that, unlike a diatomic molecule, a polyatomic molecule has not one but several vibrational degrees of freedom. In particular, a non-collinear molecule consisting of n atoms has $3n - 6$ vibrational degrees of freedom, six degrees of freedom out of the total $3n$ having gone into the translational and rotational motions. On the other hand, a collinear molecule consisting of n atoms would have $3n - 5$ vibrational degrees of freedom, for the rotational motion in this case has only two, not three, degrees of freedom. The vibrational degrees of freedom correspond to a set of normal modes characterized by a set of frequencies ω_i. It might happen that some of these frequencies have identical values; we then speak of *degenerate frequencies*.[15]

In the harmonic approximation, these normal modes may be treated independently of one another. The vibrational partition function of the molecule is then given by the product of the partition functions corresponding to individual normal modes, i.e.

$$j_{\text{vib}}(T) = \prod_i \frac{e^{-\Theta_i/2T}}{1 - e^{-\Theta_i/T}}; \qquad \Theta_i = \frac{\hbar\omega_i}{k}, \tag{43}$$

and the vibrational specific heat is given by the sum of the contributions arising from the individual modes,

$$C_{\text{vib}} = Nk\sum_i\left\{\left(\frac{\Theta_i}{T}\right)^2\frac{e^{\Theta_i/T}}{\left(e^{\Theta_i/T} - 1\right)^2}\right\}. \tag{44}$$

In general, the various Θ_i are of order 10^3 K; for instance, in the case of CO_2, which was cited in Note 15, $\Theta_1 = \Theta_2 = 960$ K, $\Theta_3 = 1990$ K and $\Theta_4 = 3510$ K.

For temperatures large in comparison with all Θ_i, specific heat would be given by the equipartition value, namely Nk for each of the normal modes. In practice, however, this limit can hardly be realized because the polyatomic molecules generally break up well before such high temperatures are reached. Secondly, the different frequencies ω_i of a polyatomic molecule are generally spread over a rather wide range of values. Consequently, as temperature rises, different modes of vibration get gradually "included" into the process; in between these "inclusions", the specific heat of the gas may stay constant over considerably large stretches of temperature.

Problems

6.1. Show that the entropy of an ideal gas in thermal equilibrium is given by the formula

$$S = k\sum_{\varepsilon}[\langle n_\varepsilon + 1\rangle \ln\langle n_\varepsilon + 1\rangle - \langle n_\varepsilon\rangle \ln\langle n_\varepsilon\rangle]$$

in the case of *bosons* and by the formula

$$S = k\sum_{\varepsilon}[-\langle 1 - n_\varepsilon\rangle \ln\langle 1 - n_\varepsilon\rangle - \langle n_\varepsilon\rangle \ln\langle n_\varepsilon\rangle]$$

in the case of *fermions*. Verify that these results are consistent with the general formula

$$S = -k\sum_{\varepsilon}\left\{\sum_{n} p_\varepsilon(n)\ln p_\varepsilon(n)\right\},$$

where $p_\varepsilon(n)$ is the probability that there are exactly n particles in the energy state ε.

6.2. Derive, for all three statistics, the relevant expressions for the quantity $\langle n_\varepsilon^2\rangle - \langle n_\varepsilon\rangle^2$ from the respective probabilities $p_\varepsilon(n)$. Show that, quite generally,

$$\langle n_\varepsilon^2\rangle - \langle n_\varepsilon\rangle^2 = kT\left(\frac{\partial\langle n_\varepsilon\rangle}{\partial\mu}\right)_T;$$

compare with the corresponding result, (4.5.3), for a system embedded in a grand canonical ensemble.

6.3 Refer to Sec. 6.2 and show that, if the occupation number n_ε of an energy level ε is restricted to the values $0, 1, \ldots, l$, then the mean occupation number of that level is given by

$$\langle n_\varepsilon\rangle = \frac{1}{z^{-1}e^{\beta\varepsilon} - 1} - \frac{l+1}{(z^{-1}e^{\beta\varepsilon})^{l+1} - 1}.$$

Check that while $l = 1$ leads to $\langle n_\varepsilon\rangle_{\text{F.D}}$, $l \to \infty$ leads to $\langle n_\varepsilon\rangle_{\text{B.E.}}$

6.4. The potential energy of a system of charged particles, characterized by particle charge e and number density $n(\boldsymbol{r})$, is given by

$$U = \frac{e^2}{2}\int\frac{n(\boldsymbol{r})n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}\,d\boldsymbol{r}\,d\boldsymbol{r}' + e\int n(\boldsymbol{r})\phi_{\text{ext}}(\boldsymbol{r})\,d\boldsymbol{r},$$

where $\phi_{\text{ext}}(\boldsymbol{r})$ is the potential of an external electric field. Assume that the entropy of the system, apart from an additive constant, is given by the formula

$$S = -k\int n(\boldsymbol{r})\ln n(\boldsymbol{r})\,d\boldsymbol{r};$$

cf. formula (3.3.13). Using these expressions, derive the equilibrium equations satisfied by the number density $n(\boldsymbol{r})$ and the total potential $\phi(\boldsymbol{r})$, the latter being

$$\phi_{\text{ext}}(\boldsymbol{r}) + e\int\frac{n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}\,d\boldsymbol{r}'.$$

6.5. Show that root-mean-square deviation in the molecular energy ε, in a system obeying Maxwell–Boltzmann distribution, is $\sqrt{(2/3)}$ times the mean molecular energy $\bar{\varepsilon}$. Compare this result with that of Problem 3.18.

6.6. Show that, for any law of distribution of molecular speeds,

$$\left\{\langle u\rangle\left\langle\frac{1}{u}\right\rangle\right\} \geq 1.$$

Check that the value of this quantity for the Maxwellian distribution is $4/\pi$.

6.7. Through a small window in a furnace, which contains a gas at a high temperature T, the spectral lines emitted by the gas molecules are observed. Because of molecular motions, each spectral line exhibits *Doppler broadening*. Show that the variation of the relative intensity $I(\lambda)$ with wavelength λ in a line is given by

$$I(\lambda) \propto \exp\left\{-\frac{mc^2(\lambda-\lambda_0)^2}{2\lambda_0^2 kT}\right\},$$

where m is the molecular mass, c the speed of light and λ_0 the mean wavelength of the line.

6.8. An ideal classical gas composed of N particles, each of mass m, is enclosed in a vertical cylinder of height L placed in a uniform gravitational field (of acceleration g) and is in thermal equilibrium; ultimately, both N and $L \to \infty$. Evaluate the partition function of the gas and derive expressions for the major thermodynamic quantities. Explain why the specific heat of this system is larger than that of a corresponding system in free space.

6.9. (a) Show that, if the temperature is uniform, the pressure of a classical gas in a uniform gravitational field decreases with height according to the *barometric formula*

$$P(z) = P(0)\exp\{-mgz/kT\},$$

where the various symbols have their usual meanings.[16]

(b) Derive the corresponding formula for an *adiabatic* atmosphere, i.e. the one in which (PV^{γ}), rather than (PV), stays constant. Also study the variation, with height, of the temperature T and the density n of the atmosphere.

6.10. (a) Show that the momentum distribution of particles in a relativistic Boltzmannian gas, with $\varepsilon = c(p^2 + m_0^2c^2)^{1/2}$, is given by

$$f(p)\,dp = Ce^{-\beta c(p^2+m_0^2c^2)^{1/2}}\,p^2\,dp,$$

with the normalization constant

$$C = \frac{\beta}{m_0^2 cK_2(\beta m_0 c^2)},$$

$K_\nu(z)$ being a modified Bessel function.

(b) Check that in the nonrelativistic limit $(kT \ll m_0c^2)$ we recover the Maxwellian distribution,

$$f(p)\,dp = \left(\frac{\beta}{2\pi m_0}\right)^{3/2} e^{-\beta p^2/2m_0}(4\pi p^2\,dp),$$

while in the extreme relativistic limit $(kT \gg m_0c^2)$ we obtain

$$f(p)\,dp = \frac{(\beta c)^3}{8\pi}e^{-\beta pc}(4\pi p^2\,dp).$$

(c) Verify that, quite generally,

$$\langle pu\rangle = 3kT.$$

6.11. (a) Considering the loss of translational energy suffered by the molecules of a gas on reflection from a *receding* wall, derive, for a quasi-static adiabatic expansion of an ideal nonrelativistic gas, the well-known relation

$$PV^{\gamma} = \text{const.},$$

where $\gamma = (3a+2)/3a$, a being the ratio of the total energy to the translational energy of the gas.

(b) Show that, in the case of an ideal extreme relativistic gas, $\gamma = (3a+1)/3a$.

6.12. (a) Determine the number of impacts made by gas molecules on a unit area of the wall in a unit time for which the angle of incidence lies between θ and $\theta + d\theta$.

(b) Determine the number of impacts made by gas molecules on a unit area of the wall in a unit time for which the speed of the molecules lies between u and $u + du$.

(c) A molecule AB dissociates if it hits the surface of a solid catalyst with a translational energy greater than 10^{-19} J. Show that the rate of the dissociative reaction $AB \rightarrow A + B$ is more than doubled by raising the temperature of the gas from 300 K to 310 K.

6.13. Consider the effusion of molecules of a Maxwellian gas through an opening of area a in the walls of a vessel of volume V.

(a) Show that, while the molecules inside the vessel have a mean kinetic energy $\frac{3}{2}kT$, the effused ones have a mean kinetic energy $2kT$, T being the *quasi-static* equilibrium temperature of the gas.

(b) Assuming that the effusion is so slow that the gas inside is always in a state of quasi-static equilibrium, determine the manner in which the density, the temperature and the pressure of the gas vary with time.

6.14. A polyethylene balloon at an altitude of 30,000 m is filled with helium gas at a pressure of 10^{-2} atm and a temperature of 300 K. The balloon has a diameter of 10 m, and has numerous pinholes of diameter 10^{-5} m each. How many pinholes per square meter of the surface must there be if 1 per cent of the gas were to leak out in 1 hour?

6.15. Consider two Boltzmannian gases A and B, at pressures P_A and P_B and temperatures T_A and T_B respectively, contained in two regions of space that communicate through a very narrow opening in the partitioning wall; see Fig. 6.8. Show that the dynamic equilibrium resulting from the mutual effusion of the two kinds of molecules satisfies the condition

$$P_A/P_B = (T_A/T_B)^{1/2},$$

rather than $P_A = P_B$ (which would be the case if the equilibrium had resulted from a hydrodynamic flow).

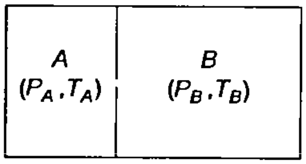


FIG. 6.8. The molecules of the gases A and B undergoing a two-way effusion.

6.16. A *small* sphere, with initial temperature T, is immersed in an ideal Boltzmannian gas at temperature T_0. Assuming that the molecules incident on the sphere are first absorbed and then re-emitted with the temperature of the sphere, determine the variation of the temperature of the sphere with time.

[*Note*: The radius of the sphere may be assumed to be much smaller than the mean free path of the molecules.]

6.17. Show that the mean value of the *relative* speed of two molecules in a Maxwellian gas is $\sqrt{2}$ times the mean speed of a molecule with respect to the walls of the container.

[Note that a similar result for the root-mean-square speeds (instead of the mean speeds) holds under much more general conditions.]

6.18. What is the probability that two molecules picked up at random from a Maxwellian gas will have a total energy between E and $E + dE$? Verify that $\langle E \rangle = 3kT$.

6.19. The energy difference between the lowest electronic state $^1S_0$ and the first excited state $^3S_1$ of the helium atom is 159,843 cm^{-1}. Evaluate the relative fraction of the excited atoms in a sample of helium gas at a temperature of 6000 K.

6.20. Derive an expression for the equilibrium constant $K(T)$ for the reaction $H_2 + D_2 \leftrightarrow 2HD$ at temperatures high enough to allow classical approximation for the rotational motion of the molecules. Show that $K(\infty) = 4$.

[*Note*: For the definition of the equilibrium constant, see Problem 3.14.]

6.21. With the help of the Euler–Maclaurin formula (6.5.19), derive high-temperature expansions for r_{even} and r_{odd}, as defined by eqns (6.5.29) and (6.5.30), and obtain corresponding expansions for C_{even} and C_{odd}, as defined by eqn. (6.5.39). Compare the mathem[illegible]l trend of these results with the nature of the corresponding curves in Fig. 6.7. Also study the lo[illegible]mperature behavior of the two specific heats and again compare your results with the relevant parts of the aforementioned curves.

6.22. The potential energy between the atoms of a hydrogen molecule is given by the (semi-empirical) *Morse potential*

$$V(r) = V_0 \left\{ e^{-2(r-r_0)/a} - 2e^{-(r-r_0)/a} \right\},$$

where $V_0 = 7 \times 10^{-12}$ erg, $r_0 = 8 \times 10^{-9}$ cm and $a = 5 \times 10^{-9}$ cm. Calculate the rotational and vibrational quanta of energy, and estimate the temperatures at which the rotational and vibrational modes of the molecules would begin to contribute towards the specific heat of the hydrogen gas.

6.23. Show that the fractional change in the equilibrium value of the internuclear distance of a diatomic molecule, as a result of rotation, is given by

$$\frac{\Delta r_0}{r_0} \simeq \left(\frac{\hbar}{\mu r_0^2 \omega} \right)^2 J(J+1) = 4 \left(\frac{\Theta_r}{\Theta_v} \right)^2 J(J+1);$$

here, ω is the angular frequency of the vibrational state in which the molecule happens to be. Estimate the numerical value of this fraction in a typical case.

6.24. The ground state of an oxygen atom is a triplet, with the following *fine structure*:

$$\varepsilon_{J=2} = \varepsilon_{J=1} - 158.5 \text{ cm}^{-1} = \varepsilon_{J=0} - 226.5 \text{ cm}^{-1}.$$

Calculate the relative fractions of the atoms occupying different J-levels in a sample of atomic oxygen at 300 K.

6.25. Calculate the contribution of the first excited electronic state, viz. $^1\Delta$ with $g_e = 2$, of the O_2 molecule towards the Helmholtz free energy and the specific heat of oxygen gas at a temperature of 5000 K; the separation of this state from the ground state, viz. $^3\Sigma$ with $g_e = 3$, is 7824 cm^{-1}. How would these results be affected if the parameters Θ_r and Θ_v of the O_2 molecule had different values in the two electronic states?

6.26. The rotational kinetic energy of a rotator with three degrees of freedom can be written as

$$\varepsilon_{\text{rot}} = \frac{M_\xi^2}{2I_1} + \frac{M_\eta^2}{2I_2} + \frac{M_\zeta^2}{2I_3},$$

where (ξ, η, ζ) are coordinates in a rotating frame of reference whose axes coincide with the principal axes of the rotator, while (M_ξ, M_η, M_ζ) are the corresponding angular momenta. Carrying out integrations in the phase space of the rotator, derive expression (6.5.41) for the partition function $j_{\text{rot}}(T)$ in the classical approximation.

6.27. Determine the translational, rotational and vibrational contributions towards the molar entropy and the molar specific heat of carbon dioxide at N.T.P. Assume the ideal-gas formulae and use the following data: molecular weight $M = 44.01$; moment of inertia I of a CO_2 molecule $= 71.67 \times 10^{-40}$ g cm^2; wave numbers of the various modes of vibration: $\bar{\nu}_1 = \bar{\nu}_2 = 667.3$ cm^{-1}, $\bar{\nu}_3 = 1383.3$ cm^{-1} and $\bar{\nu}_4 = 2439.3$ cm^{-1}.

6.28. Determine the molar specific heat of ammonia at a temperature of 300 K. Assume the ideal-gas formula and use the following data: the principal moments of inertia: $I_1 = 4.44 \times 10^{-40}$ g cm^2, $I_2 = I_3 = 2.816 \times 10^{-40}$ g cm^2; wave numbers of the various modes of vibration: $\bar{\nu}_1 = \bar{\nu}_2 = 3336$ cm^{-1}, $\bar{\nu}_3 = \bar{\nu}_4 = 950$ cm^{-1}; $\bar{\nu}_5 = 3414$ cm^{-1} and $\bar{\nu}_6 = 1627$ cm^{-1}.

Notes

[1] For a critique of this derivation, see Landsberg (1954a, 1961).

[2] The special case of fluctuations in the *ground state occupation number*, n_0, of a Bose–Einstein system has been discussed by Wergeland (1969) and by Fujiwara, ter Haar and Wergeland (1970).

[3] Clearly, only those velocities are relevant for which $u_z > 0$.

[4] As is well known, the presence of the nuclear spin gives rise to the so-called *hyperfine structure* in the electronic states. However, the intervals of this structure are such that for practically all temperatures of interest they are small in comparison with kT; for concreteness, these intervals correspond to T-values of the order of 10^{-1}–10^0 K. Accordingly, in the evaluation of the partition function $j(T)$, the hyperfine splitting of the electronic state may be disregarded while the multiplicity introduced by the nuclear spin may [illegible] taken into account through a degeneracy factor.

[5] It seems worth[illegible] to note here that the values of $\Delta\varepsilon_J/k$ for the components of the normal triplet term of oxygen are 230 K and 320 K, while those for the normal quintuplet term of iron range from 600 K to 1400 K.

[6] An odd case arises with oxygen. The separation between its normal term $^3\Sigma$ and the first excited term $^1\Delta$ is about 11,250 K, whereas the dissociation energy is about 55,000 K. The relevant factor $e^{-\varepsilon_1/kT}$, therefore, can be quite significant even when the factor $e^{-E_{\text{diss}}/kT}$ is not, say for $T \sim 2000-6000$ K.

[7] Strictly speaking, the term in question splits into two levels–the so-called Λ-doublet. The separation of the levels, however, is such that we can safely neglect it.

[8] The separation of the resulting levels is again negligible from the thermodynamic point of view; as an example, one may cite the very narrow triplet term of O_2.

[9] It may be pointed out that the vibrational motion of a molecule is influenced by the centrifugal force arising from the molecular rotation. This leads to an interaction between the rotational and the vibrational modes. However, unless the temperature is too high, this interaction can be neglected and the two modes treated independently of one another.

[10] In principle, these two effects are of the same order of magnitude.

[11] It seems instructive to outline here the purely classical derivation of the rotational partition function. Specifying the rotation of the molecule by the angles (θ, ϕ) and the corresponding momenta (p_θ, p_ϕ), the kinetic energy assumes the form

$$\varepsilon_{\text{rot}} = \frac{1}{2I} p_\theta^2 + \frac{1}{2I \sin^2\theta} p_\phi^2,$$

whence

$$j_{\text{rot}}(T) = \frac{1}{h^2} \int e^{-\varepsilon_{\text{rot}}/kT} (dp_\theta \, dp_\phi \, d\theta \, d\phi) = \frac{IkT}{\pi \hbar^2} \int_0^{\phi_{\max}} d\phi.$$

For heteronuclear molecules $\phi_{\max} = 2\pi$, while for homonuclear ones $\phi_{\max} = \pi$.

[12] See, for example, Schiff (1968), sec. 41.

[13] For example, the symmetry number γ for H_2O (isosceles triangle) is 2, for NH_3 (regular triangular pyramid) it is 3, while for CH_4 (tetrahedron) and C_6H_6 (regular hexagon) it is 12. For heteronuclear molecules, the symmetry number is unity.

[14] In the case of a collinear molecule, such as N_2O and CO_2, there are only two degrees of freedom for rotation; consequently, $j_{\text{rot}}^C(T)$ is given by $(2IkT/\hbar^2)$, where I is the (common) value of the two moments of inertia of the molecule; see eqn. (17). Of course, we must also take into account the symmetry number γ. In the examples quoted here, the molecule N_2O, being spatially asymmetric (NNO), has symmetry number 1, while the molecule CO_2, being spatially symmetric (OCO), has symmetry number 2.

[15] For example, of the four frequencies characterizing the normal modes of vibration of the collinear molecule OCO, two that correspond to the (transverse) bending modes, namely $\underset{\downarrow}{\mathrm{O}}\,\overset{\uparrow}{\mathrm{C}}\,\underset{\downarrow}{\mathrm{O}}$, are equal, while the others that correspond to (longitudinal) oscillations along the molecular axis, namely ←O C→ ←O and ←O C O→, are different; see Problem 6.27.

[16] This formula was first given by Boltzmann (1879). For a critical study of its derivation, see Walton (1969).

CHAPTER 7

IDEAL BOSE SYSTEMS

IN CONTINUATION of Secs 6.1–6.3, we shall now investigate in detail the physical behavior of a class of systems in which, while the intermolecular interactions are still negligible, the effects of quantum statistics (which arise from the indistinguishability of the particles) assume an increasingly important role. This means that the temperature T and the particle density n of the system no longer conform to the criterion

$$n\lambda^3 \equiv \frac{nh^3}{(2\pi mkT)^{3/2}} \ll 1, \qquad (6.3.5)$$

where $\lambda \{\equiv h/(2\pi mkT)^{1/2}\}$ is the *mean thermal wavelength* of the particles. In fact, the quantity $(n\lambda^3)$ turns out to be a very appropriate parameter, in terms of which the various physical properties of the system can be adequately expressed. In the limit $(n\lambda^3) \to 0$, all physical properties go over smoothly to their classical counterparts. For small, but not negligible, values of $(n\lambda^3)$, the various quantities pertaining to the system can be expanded as power series in this parameter; from these expansions one obtains the first glimpse of the manner in which departure from classical behavior sets in. When $(n\lambda^3)$ becomes of the order of unity, the behavior of the system becomes significantly different from the classical one and is characterized by typical quantum effects. A study of the system under these circumstances brings us face to face with a set of phenomena unknown in classical statistics.

It is evident that a system is more likely to display quantum behavior when it is at a relatively low temperature and/or has a relatively high density of particles.[1] Moreover, the smaller the particle mass the larger the quantum effects.

Now, when $(n\lambda^3)$ is of the order of unity, then not only does the behavior of a system exhibit significant departure from typical classical behavior but it is also influenced by whether the particles constituting the system obey Bose–Einstein statistics or Fermi–Dirac statistics. Under these circumstances, the properties of the two kinds of systems are themselves very different. In the present chapter we propose to consider systems belonging to the first category while the succeeding chapter will deal with systems belonging to the second category.

7.1. Thermodynamic behavior of an ideal Bose gas

We obtained, in Secs 6.1 and 6.2, the following formulae for an ideal Bose gas:

$$\frac{PV}{kT} \equiv \ln \mathcal{Q} = -\sum_{\varepsilon} \ln (1 - ze^{-\beta\varepsilon}) \tag{1}$$

and

$$N \equiv \sum_{\varepsilon} \langle n_{\varepsilon} \rangle = \sum_{\varepsilon} \frac{1}{z^{-1}e^{\beta\varepsilon} - 1}, \tag{2}$$

where $\beta = 1/kT$, while z is the fugacity of the gas which is related to the chemical potential μ through the formula

$$z \equiv \exp(\mu/kT). \tag{3}$$

As noted earlier, $ze^{-\beta\varepsilon}$, for all ε, is less than unity. In view of the fact that, for large V, the spectrum of the single-particle states is almost a continuous one, summations on the right-hand sides of eqns (1) and (2) may be replaced by integrations. In doing so, we make use of the asymptotic expression (2.4.7) for the density of states $a(\varepsilon)$ in the neighborhood of a given energy ε, namely[2]

$$a(\varepsilon)\,d\varepsilon = (2\pi V/h^3)(2m)^{3/2}\varepsilon^{1/2}\,d\varepsilon. \tag{4}$$

We, however, note that by substituting this expression into our integrals we are inadvertently giving a weight *zero* to the energy level $\varepsilon = 0$. This is wrong because in a quantum-mechanical treatment we must give a statistical weight unity to each nondegenerate single-particle state in the system. It is, therefore, advisable to take this particular state out of the sum in question before carrying out the integration. We thus obtain

$$\frac{P}{kT} = -\frac{2\pi}{h^3}(2m)^{3/2} \int_0^{\infty} \varepsilon^{1/2} \ln (1 - ze^{-\beta\varepsilon})\,d\varepsilon - \frac{1}{V} \ln (1 - z) \tag{5}$$

and

$$\frac{N}{V} = \frac{2\pi}{h^3}(2m)^{3/2} \int_0^{\infty} \frac{\varepsilon^{1/2}\,d\varepsilon}{z^{-1}e^{\beta\varepsilon} - 1} + \frac{1}{V}\frac{z}{1 - z}; \tag{6}$$

of course, the lower limit of the integrals can still be taken as 0, because the state $\varepsilon = 0$ is not going to contribute towards the integrals anyway.

Before proceeding further, a word about the relative importance of the last terms in eqns (5) and (6). For $z \ll 1$, which corresponds to situations not far removed from the classical limit, each of these terms is of order $1/N$ and, therefore, negligible. However, as z increases and assumes values close to unity, the term $z/(1 - z)V$ in (6), which is identically equal to N_0/V (N_0 being the number of particles in the ground state $\varepsilon = 0$), can well become a significant fraction of the quantity N/V; this accumulation of a macroscopic fraction of the given particles into a single state $\varepsilon = 0$ leads to the phenomenon of *Bose–Einstein condensation*. Nevertheless, since $z/(1 - z) = N_0$ and hence $z = N_0/(N_0 + 1)$, the term $\{-V^{-1} \ln (1 - z)\}$ in (5) is equal to $\{V^{-1} \ln (N_0 + 1)\}$, which is at most

$O(N^{-1}\ln N)$; this term is, therefore, negligible for *all* values of z and hence may be dropped altogether.

We now obtain from eqns (5) and (6), on substituting $\beta\varepsilon = p^2/(2mkT) = x$,

$$\frac{P}{kT} = -\frac{2\pi(2mkT)^{3/2}}{h^3}\int_0^\infty x^{1/2}\ln(1 - ze^{-x})\,dx = \frac{1}{\lambda^3}g_{5/2}(z) \tag{7}$$

and

$$\frac{N - N_0}{V} = \frac{2\pi(2mkT)^{3/2}}{h^3}\int_0^\infty \frac{x^{1/2}\,dx}{z^{-1}e^x - 1} = \frac{1}{\lambda^3}g_{3/2}(z), \tag{8}$$

where

$$\lambda = h/(2\pi mkT)^{1/2}, \tag{9}$$

while $g_\nu(z)$ are *Bose–Einstein functions* defined by, see Appendix D,

$$g_\nu(z) = \frac{1}{\Gamma(\nu)}\int_0^\infty \frac{x^{\nu-1}\,dx}{z^{-1}e^x - 1} = z + \frac{z^2}{2^\nu} + \frac{z^3}{3^\nu} + \cdots; \tag{10}$$

note that to write (7) in terms of the function $g_{5/2}(z)$ we first carried out an integration by parts. Equations (7) and (8) are our basic results; on elimination of z, they would give us the *equation of state* of the system.

The internal energy of this system is given by

$$\begin{aligned} U &\equiv -\left(\frac{\partial}{\partial\beta}\ln\mathcal{Q}\right)_{z,V} = kT^2\left\{\frac{\partial}{\partial T}\left(\frac{PV}{kT}\right)\right\}_{z,V} \\ &= kT^2 V g_{5/2}(z)\left\{\frac{d}{dT}\left(\frac{1}{\lambda^3}\right)\right\} = \frac{3}{2}kT\frac{V}{\lambda^3}g_{5/2}(z); \end{aligned} \tag{11}$$

here, use has been made of eqn. (7) and of the fact that $\lambda \propto T^{-1/2}$. Thus, quite generally, our system satisfies the relationship

$$P = \tfrac{2}{3}(U/V). \tag{12}$$

For small values of z, we can make use of expansion (10); at the same time, we can neglect N_0 in comparison with N. An elimination of z between eqns (7) and (8) can then be carried out by first inverting the series in (8) to obtain an expansion for z in powers of $(n\lambda^3)$ and then substituting this expansion into the series appearing in (7). The equation of state thereby takes the form of the *virial expansion*,

$$\frac{PV}{NkT} = \sum_{l=1}^{\infty} a_l\left(\frac{\lambda^3}{v}\right)^{l-1}, \tag{13}$$

where $v(\equiv 1/n)$ is the volume per particle; the coefficients a_l, which are referred to as the *virial coefficients* of the system, turn out to be

$$\left.\begin{aligned} a_1 &= 1, \\ a_2 &= -\frac{1}{4\sqrt{2}} = -0.17678, \\ a_3 &= -\left(\frac{2}{9\sqrt{3}} - \frac{1}{8}\right) = -0.00330, \\ a_4 &= -\left(\frac{3}{32} + \frac{5}{32\sqrt{2}} - \frac{1}{2\sqrt{6}}\right) = -0.00011, \end{aligned}\right\} \tag{14}$$

and so on. For the specific heat of the gas, we obtain

$$\begin{aligned} \frac{C_V}{Nk} &\equiv \frac{1}{Nk}\left(\frac{\partial U}{\partial T}\right)_{N,V} = \frac{3}{2}\left\{\frac{\partial}{\partial T}\left(\frac{PV}{Nk}\right)\right\}_v \\ &= \frac{3}{2}\sum_{l=1}^{\infty}\frac{5-3l}{2}a_l\left(\frac{\lambda^3}{v}\right)^{l-1} \\ &= \frac{3}{2}\left[1 + 0.0884\left(\frac{\lambda^3}{v}\right) + 0.0066\left(\frac{\lambda^3}{v}\right)^2 + 0.0004\left(\frac{\lambda^3}{v}\right)^3 + \cdots\right]. \end{aligned} \tag{15}$$

As $T \to \infty$ (and hence $\lambda \to 0$), both the pressure and the specific heat of the gas approach their classical values, namely nkT and $\frac{3}{2}Nk$, respectively. We also note that at finite, but large, temperatures the specific heat of the gas is larger than its limiting value; in other words, the (C_V, T)-curve has a negative slope at high temperatures. On the other hand, as $T \to 0$, the specific heat must go to zero. Consequently, it must pass through a maximum somewhere. As seen later, this maximum is in the nature of a cusp which appears at a critical temperature T_c; the derivative of the specific heat is found to be discontinuous at this temperature (see Fig. 7.4).

As the temperature of the system falls (and the value of the parameter λ^3/v grows), expansions such as (13) and (15) do not remain useful. We then have to work with formulae (7), (8) and (11) as such. The precise value of z is now obtained from eqn. (8), which may be rewritten as

$$N_e = V\frac{(2\pi mkT)^{3/2}}{h^3}g_{3/2}(z), \tag{16}$$

where N_e is the number of particles in the excited states ($\varepsilon \neq 0$); of course, unless z gets extremely close to unity, $N_e \simeq N$.[3] It is obvious that, for $0 \leq z \leq 1$, the function $g_{3/2}(z)$ increases monotonically with z and is *bounded*, its largest value being

$$g_{3/2}(1) = 1 + \frac{1}{2^{3/2}} + \frac{1}{3^{3/2}} + \cdots \equiv \zeta\left(\frac{3}{2}\right) \simeq 2.612; \tag{17}$$

see eqn. (D.5). Hence, for all z of interest,

$$g_{3/2}(z) \leq \zeta(\tfrac{3}{2}). \tag{18}$$

Consequently, for given V and T, the total (equilibrium) number of particles in all the excited states taken together is also bounded, i.e.

$$N_e \le V\frac{(2\pi mkT)^{3/2}}{h^3}\zeta\left(\frac{3}{2}\right). \tag{19}$$

Now, so long as the actual number of particles in the system is less than this limiting value, everything is well and good; practically all the particles in the system are distributed over the excited states and the precise value of z is determined by eqn. (16), with $N_e \simeq N$. However, if the actual number of particles exceeds this limiting value, then it is natural that the excited states will receive as many of them as they can hold, namely

$$N_e = V\frac{(2\pi mkT)^{3/2}}{h^3}\zeta\left(\frac{3}{2}\right), \tag{20}$$

while the rest will be pushed *en masse* into the ground state $\varepsilon = 0$ (whose capacity, under all circumstances, is practically unlimited):

$$N_0 = N - \left\{V\frac{(2\pi mkT)^{3/2}}{h^3}\zeta\left(\frac{3}{2}\right)\right\}. \tag{21}$$

The precise value of z is now determined by the formula

$$z = \frac{N_0}{N_0+1} \simeq 1 - \frac{1}{N_0}, \tag{22}$$

which, for all practical purposes, is unity. This curious phenomenon of a macroscopically large number of particles accumulating in a single quantum state ($\varepsilon = 0$) is generally referred to as the phenomenon of *Bose–Einstein condensation*. In a certain sense, this phenomenon is akin to the familiar process of a vapor condensing into the liquid state, which takes place in the ordinary physical space. Conceptually, however, the two processes are very different. Firstly, the phenomenon of Bose–Einstein condensation is purely of a quantum origin (occurring even in the absence of intermolecular forces); secondly, it takes place at best in the momentum space and not in the coordinate space.[4]

The condition for the onset of Bose–Einstein condensation is

$$N > VT^{3/2}\frac{(2\pi mk)^{3/2}}{h^3}\zeta\left(\frac{3}{2}\right) \tag{23}$$

or, if we hold N and V constant and vary T,

$$T < T_c = \frac{h^2}{2\pi mk}\left\{\frac{N}{V\zeta\left(\frac{3}{2}\right)}\right\}^{2/3}; \tag{24^5}$$

here, T_c denotes a characteristic temperature that depends upon the particle mass m and the particle density N/V in the system. Accordingly, for $T < T_c$, the system may be looked upon as a mixture of two "phases":

(i) a *normal* phase, consisting of $N_e\ \{= N(T/T_c)^{3/2}\}$ particles distributed over the excited states ($\varepsilon \neq 0$), and

(ii) a *condensed* phase, consisting of $N_0\ \{= (N - N_e)\}$ particles accumulated in the ground state ($\varepsilon = 0$).

Figure 7.1 shows the manner in which the complementary fractions (N_e/N) and (N_0/N) vary with T. For $T > T_c$, we have the normal phase alone; the number of particles in the ground state, viz. $z/(1-z)$, is $O(1)$, which is completely negligible in comparison with the total number N. Clearly, the situation is ***singular*** at $T = T_c$. For later reference, we note that, at $T \to T_c$ from below, the condensate fraction vanishes as follows:

$$\frac{N_0}{N} = 1 - \left(\frac{T}{T_c}\right)^{3/2} \approx \frac{3}{2}\frac{T_c - T}{T_c}. \tag{25}$$

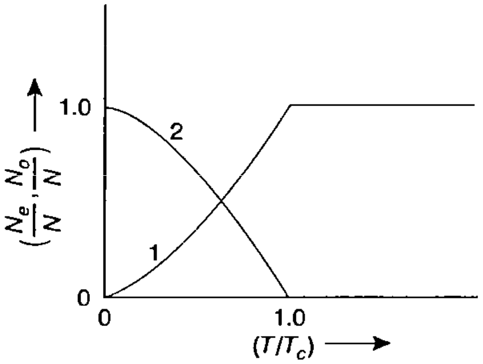


FIG. 7.1. Fractions of the normal phase and the condensed phase in an ideal Bose gas as a function of the temperature parameter (T/T_c).

A knowledge of the variation of z with T is also of interest here. It is, however, simpler to consider the variation of z with (v/λ^3), the latter being proportional to $T^{3/2}$. For $0 \le (v/\lambda^3) \le (2.612)^{-1}$, which corresponds to $0 \le T \le T_c$, the parameter $z \simeq 1$; see eqn. (22). For $(v/\lambda^3) > (2.612)^{-1}$, $z < 1$ and is determined by the relationship

$$g_{3/2}(z) = (\lambda^3/v) < 2.612; \tag{26}$$

[6]

see eqn. (8). For $(v/\lambda^3) \gg 1$, we have: $g_{3/2}(z) \ll 1$ and, hence, $z \ll 1$. Under these circumstances, $g_{3/2}(z) \simeq z$; see eqn. (10). Therefore, in this region, $z \simeq (v/\lambda^3)^{-1}$, in agreement with the classical case.[7] Figure 7.2 shows the variation of z with (v/λ^3).

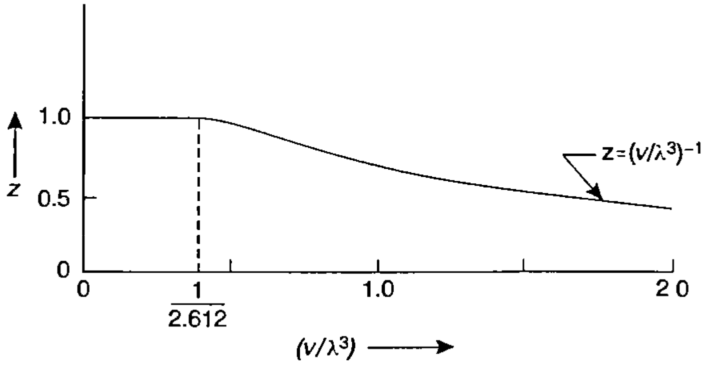


FIG. 7.2. The fugacity of an ideal Bose gas as a function of (v/λ^3)

Next, we examine the (P, T)-diagram of this system, i.e. the variation of P with T, *keeping v fixed*. Now, for $T < T_c$, the pressure is given by eqn. (7), with z replaced by unity, i.e.

$$P(T) = \frac{kT}{\lambda^3}\zeta\left(\frac{5}{2}\right), \tag{27}$$

which is proportional to $T^{5/2}$ and is *independent* of v–implying infinite compressibility. At the transition point the value of the pressure is

$$P(T_c) = \left(\frac{2\pi m}{h^2}\right)^{3/2}(kT_c)^{5/2}\zeta\left(\frac{5}{2}\right); \tag{28}$$

with the help of (24), this can be written as

$$P(T_c) = \frac{\zeta\left(\frac{5}{2}\right)}{\zeta\left(\frac{3}{2}\right)}\left(\frac{N}{V}kT_c\right) \simeq 0.5134\left(\frac{N}{V}kT_c\right) \tag{29}$$

Thus, the pressure (exerted by the particles) of an ideal Bose gas at the transition temperature T_c is about one-half of that (exerted by the particles) of an equivalent Boltzmannian gas.[8] For $T > T_c$, the pressure is given by

$$P = \frac{N}{V}kT\frac{g_{5/2}(z)}{g_{3/2}(z)}, \tag{30}$$

where $z(T)$ is determined by the implicit relationship

$$g_{3/2}(z) = \frac{\lambda^3}{v} = \frac{N}{V}\frac{h^3}{(2\pi mkT)^{3/2}}. \tag{26a}$$

Unless T is very high, the expression for P cannot be expressed in any simpler terms; of course, for $T \gg T_c$, the virial expansion (13) can be made use of. As $T \to \infty$, the pressure approaches the classical value NkT/V. All these features are shown in Fig. 7.3. The transition line in the figure portrays eqn. (27). The actual (P, T)-curve follows this line from $T = 0$ up to $T = T_c$ and thereafter departs, tending asymptotically to the classical limit. It may be pointed out that the region to the right of the transition line belongs to the normal phase alone, the line itself belongs to the mixed phase, while the region to the left is inaccessible to the system.

In view of the direct relationship between the internal energy of the gas and its pressure, see eqn. (12), Fig. 7.3 depicts equally well the variation of U with T (of course, with v fixed). Its slope should, therefore, be a measure of the specific heat $C_V(T)$ of the gas. We readily observe that the specific heat is vanishingly small at low temperatures, rises with T until it reaches a maximum at $T = T_c$; thereafter, it decreases, tending asymptotically to the constant classical value. Analytically, for $T \leq T_c$, we obtain [see eqns (15) and (27)]

$$\frac{C_V}{Nk} = \frac{3}{2}\frac{V}{N}\zeta\left(\frac{5}{2}\right)\frac{d}{dT}\left(\frac{T}{\lambda^3}\right) = \frac{15}{4}\zeta\left(\frac{5}{2}\right)\frac{v}{\lambda^3}, \tag{31}$$

which is proportional to $T^{3/2}$. At $T = T_c$, we have

$$\frac{C_V(T_c)}{Nk} = \frac{15}{4}\frac{\zeta\left(\frac{5}{2}\right)}{\zeta\left(\frac{3}{2}\right)} \simeq 1.925, \tag{32}$$

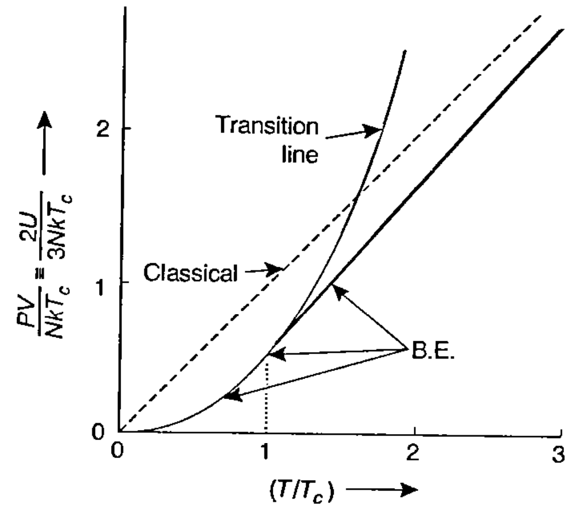


FIG. 7.3. The pressure and the internal energy of an ideal Bose gas as a function of the temperature parameter (T/T_c).

which is significantly higher than the classical value 1.5. For $T > T_c$, we obtain an implicit formula. First of all,

$$\frac{C_V}{Nk} = \left[\frac{\partial}{\partial T}\left(\frac{3}{2}T\frac{g_{5/2}(z)}{g_{3/2}(z)}\right)\right]_v; \tag{33}$$

see eqns (11) and (26). To carry out the differentiation, we need to know $(\partial z/\partial T)_v$; this can be obtained from eqn. (26) with the help of the recurrence relation (D.10). On one hand, since $g_{3/2}(z) \propto T^{-3/2}$,

$$\left[\frac{\partial}{\partial T}g_{3/2}(z)\right]_v = -\frac{3}{2T}g_{3/2}(z); \tag{34}$$

on the other,

$$z\frac{\partial}{\partial z}g_{3/2}(z) = g_{1/2}(z). \tag{35}$$

Combining the two results, we obtain

$$\frac{1}{z}\left(\frac{\partial z}{\partial T}\right)_v = -\frac{3}{2T}\frac{g_{3/2}(z)}{g_{1/2}(z)}. \tag{36}$$

Equation (33) now gives

$$\frac{C_V}{Nk} = \frac{15}{4}\frac{g_{5/2}(z)}{g_{3/2}(z)} - \frac{9}{4}\frac{g_{3/2}(z)}{g_{1/2}(z)}; \tag{37}$$

the value of z, as a function of T, is again to be determined from eqn. (26). In the limit $z \to 1$, the second term in (37) vanishes because of the divergence of $g_{1/2}(z)$, while the first term gives exactly the result appearing in (32). The specific heat is, therefore, continuous at the transition point. Its derivative is, however, discontinuous, the magnitude of the discontinuity being

$$\left(\frac{\partial C_V}{\partial T}\right)_{T=T_c-0} - \left(\frac{\partial C_V}{\partial T}\right)_{T=T_c+0} = \frac{27Nk}{16\pi T_c}\left\{\zeta\left(\frac{3}{2}\right)\right\}^2 \simeq 3.665\frac{Nk}{T_c}; \tag{38}$$

see Problem 5. For $T > T_c$, the specific heat decreases steadily towards the limiting value

$$\left(\frac{C_V}{Nk}\right)_{z\to 0} = \frac{15}{4} - \frac{9}{4} = \frac{3}{2}. \tag{39}$$

Figure 7.4 shows all these features of the (C_V, T)-relationship. It may be noted that it was the similarity of this curve with the experimental one for liquid He^4 (Fig. 7.5) that prompted F. London to suggest, in 1938, that the curious phase transition that occurs in liquid He^4 at a temperature of about 2.19 K might be a manifestation of the *Bose–Einstein condensation* taking place in the liquid. Indeed, if we substitute, in (24), data for liquid He^4, namely $m = 6.65 \times 10^{-24}$ g and $V = 27.6$ cm^3/mole, we obtain for T_c a value of about 3.13 K, which is not drastically different from the observed transition temperature of the liquid. Moreover, the interpretation of the phase transition in liquid He^4 as Bose–Einstein condensation provides a theoretical basis for the *two-fluid model* of this liquid, which was empirically put forward by Tisza (1938) in order to explain the physical behavior of the liquid below the transition temperature. According to London,

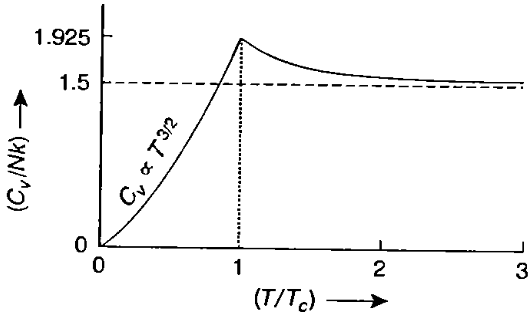


FIG. 7.4. The specific heat of an ideal Bose gas as a function of the temperature parameter (T/T_c).

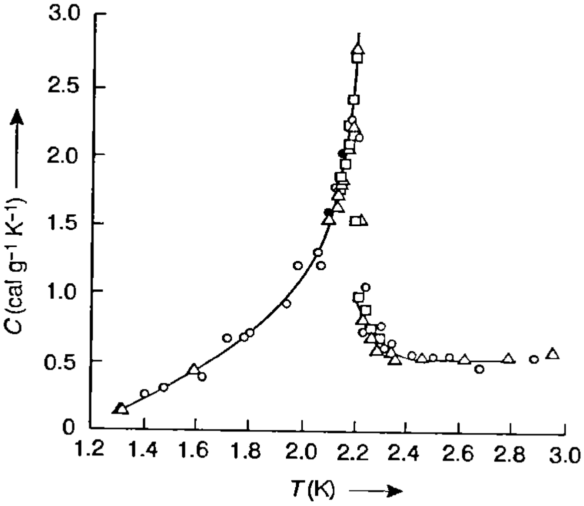


FIG. 7.5. The specific heat of liquid He^4 under its own vapor pressure (after Keesom and co-workers).

the N_0 particles that occupy a single, entropyless state ($\varepsilon = 0$) could be identified with the "superfluid component" of the liquid and the N_e particles that occupy the excited states ($\varepsilon \neq 0$) with the "normal component". As required in the model of Tisza, the superfluid fraction makes its appearance at the transition temperature T_c, and builds up at the cost of the normal fraction until at $T = 0$ the whole of the fluid becomes superfluid; cf. Fig. 7.1. Of course, the actual temperature dependence of these fractions, and of other physical quantities pertaining to liquid He^4, is considerably different from what the simple-minded ideal Bose gas suggests. London had expected that the inclusion of intermolecular interactions would improve the quantitative agreement. Although this expectation has been partially vindicated, there have been other advances in the field which provide alternative ways of looking at the helium problem; see Sec. 7.5. Nevertheless, many of the features provided by London's interpretation of this phenomenon continue to be of value.

Historically, the experimental measurements of the specific heat of liquid He^4, which led to the discovery of this so-called He I–He II transition, were first made by Keesom in 1927–8. Struck by the shape of the (C_V, T)-curve, Keesom gave this transition the name λ-transition; as a result, the term transition temperature (or transition point) also came to be known as λ-temperature (or λ-point).

We shall now look at the *isotherms* of the ideal Bose gas, i.e. the variation of the pressure of the gas with its volume, keeping T fixed. The Bose–Einstein condensation now sets in at a characteristic volume v_c, given by

$$v_c = \lambda^3/\zeta\left(\tfrac{3}{2}\right); \tag{40}$$

see (23). We note that $v_c \propto T^{-3/2}$. For $v < v_c$, the pressure of the gas is independent of v and is given by

$$P_0 = \frac{kT}{\lambda^3}\zeta\left(\frac{5}{2}\right) \tag{41}$$

see (27). The region of the mixed phase, in the (P, v)-diagram, is marked by a boundary line (called the *transition line*) given by the equation

$$P_0 v_c^{5/3} = \frac{h^2}{2\pi m}\frac{\zeta\left(\frac{5}{2}\right)}{\left\{\zeta\left(\frac{3}{2}\right)\right\}^{5/3}} = \text{const.}; \tag{42}$$

see Fig. 7.6. Clearly, the region to the left of this line belongs to the mixed phase, while the region to the right belongs to the normal phase alone.

Finally, we examine the *adiabats* of the ideal Bose gas. For this, we need an expression for the entropy of the system. Making use of the thermodynamic formula

$$U - TS + PV \equiv N\mu \tag{43}$$

and the expressions for U and P obtained above, we get

$$\frac{S}{Nk} \equiv \frac{U + PV}{NkT} - \frac{\mu}{kT} = \begin{cases} \dfrac{5}{2}\dfrac{g_{5/2}(z)}{g_{3/2}(z)} - \ln z & \text{for} \quad T > T_c, \qquad (44a) \\[2ex] \dfrac{5}{2}\dfrac{v}{\lambda^3}\zeta\left(\dfrac{5}{2}\right) & \text{for} \quad T \leq T_c; \qquad (44b) \end{cases}$$

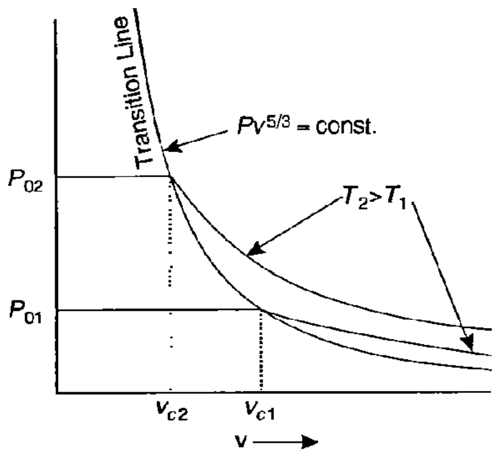


FIG. 7.6. The isotherms of an ideal Bose gas.

again, the value of $z(T)$, for $T > T_c$, is to be obtained from eqn. (26). Now, a reversible adiabatic process implies the constancy of S and N. For $T > T_c$, this implies the constancy of z as well and in turn, by (26), the constancy of (v/λ^3). For $T \leq T_c$, it again implies the same. We thus obtain, quite generally, the following relationship between the volume and the temperature of the system when it undergoes a reversible adiabatic process:

$$vT^{3/2} = \text{const.} \tag{45}$$

The corresponding relationship between the pressure and the temperature is

$$P/T^{5/2} = \text{const.}; \tag{46}$$

see eqns (7) and (27). Eliminating T, we obtain

$$Pv^{5/3} = \text{const.} \tag{47}$$

as the equation for an adiabat of the ideal Bose gas.

Incidentally, the foregoing results are exactly the same as for an ideal classical gas. There is, however, a significant difference between the two cases, i.e. while the exponent $\frac{5}{3}$ in formula (47) is identically equal to the ratio of the specific heats C_P and C_V in the case of the ideal classical gas, it is not so in the case of the ideal Bose gas. For the latter, this ratio is given by

$$\gamma \equiv \frac{C_P}{C_V} = 1 + \frac{4}{9}\frac{C_V}{Nk}\frac{g_{1/2}(z)}{g_{3/2}(z)} \tag{48a}$$

$$= \frac{5}{3}\frac{g_{5/2}(z)g_{1/2}(z)}{\{g_{3/2}(z)\}^2}; \tag{48b}$$

see Problems (7.4) and (7.5). It is only for $T \gg T_c$ that $\gamma \simeq \frac{5}{3}$. At any finite temperature, $\gamma > \frac{5}{3}$ and as $T \to T_c, \gamma \to \infty$. Equation (47), on the other hand, holds for *all* T.

In the region of the mixed phase ($T < T_c$), the entropy of the gas may be written as

$$S = N_e \cdot \frac{5}{2} k \frac{\zeta\left(\frac{5}{2}\right)}{\zeta\left(\frac{3}{2}\right)} \propto N_e; \tag{49}$$

see eqns (20) and (44b). As expected, the N_0 particles (that constitute the "condensate") do not contribute towards the entropy of the system, while the N_e particles (that constitute the normal part) contribute an amount of $\frac{5}{2}k\zeta(\frac{5}{2})/\zeta(\frac{3}{2})$ per particle.

7.2. Thermodynamics of the black-body radiation

One of the most important applications of Bose–Einstein statistics is to investigate the equilibrium properties of the *black-body radiation*. We consider a radiation cavity of volume V at temperature T. Historically, this system has been looked upon from two, practically identical but conceptually different, points of view:

(i) as an assembly of *harmonic oscillators* with quantized energies $(n_s + \frac{1}{2})\hbar\omega_s$, where $n_s = 0, 1, 2, \ldots$, and ω_s is the (angular) frequency of an oscillator, or

(ii) as a gas of identical and indistinguishable quanta–the so-called *photons* –the energy of a photon (corresponding to the frequency ω_s of the radiation mode) being $\hbar\omega_s$.

The first point of view is essentially the one adopted by Planck (1900), except that we have also included here the zero-point energy of the oscillator; for the thermodynamics of the radiation, this energy is of no great consequence and may be dropped altogether. The oscillators, being distinguishable from one another (by the very values of ω_s), would obey Maxwell–Boltzmann statistics; however, the expression for the single-oscillator partition function $Q_1(V, T)$ would be different from the classical expression because now the energies accessible to the oscillator are discrete, rather than continuous; cf. eqns (3.8.2) and (3.8.14). The expectation value of the energy of a Planck oscillator of frequency ω_s is then given by eqn. (3.8.20), excluding the zero-point term $\frac{1}{2}\hbar\omega_s$:

$$\langle \varepsilon_s \rangle = \frac{\hbar\omega_s}{e^{\hbar\omega_s/kT} - 1}. \tag{1}$$

Now, the number of normal modes of vibration per unit volume of the cavity in the frequency range $(\omega, \omega + d\omega)$ is given by the *Rayleigh expression*

$$2 \cdot 4\pi \left(\frac{1}{\lambda}\right)^2 d\left(\frac{1}{\lambda}\right) = \frac{\omega^2\, d\omega}{\pi^2 c^3}, \tag{2}$$

where the factor 2 has been included to take into account the duplicity of the transverse modes;[9] the symbol c here denotes the speed of light. By eqns (1) and (2), the energy density associated with the frequency range $(\omega, \omega + d\omega)$ is given by

$$u(\omega)\, d\omega = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3\, d\omega}{e^{\hbar\omega/kT} - 1}, \tag{3}$$

which is *Planck's formula* for the distribution of energy over the black-body spectrum. Integrating (3) over all values of ω, we obtain the total energy density in the cavity.

The second point of view originated with Bose (1924) and Einstein (1924, 1925). Bose investigated the problem of the "distribution of *photons* over the various energy levels" in the system; however, instead of worrying about the allocation of the various photons to the various energy levels (as one would have ordinarily done), he concentrated on the statistics of the energy levels themselves! He examined questions such as the "probability of an energy level $\varepsilon_s (= \hbar\omega_s)$ being occupied by n_s photons at a time", "the mean values of n_s and ε_s", etc. The statistics of the energy levels is indeed Boltzmannian; the mean values of n_s and ε_s, however, turn out to be

$$\begin{aligned}\langle n_s \rangle &= \sum_{n_s=0}^{\infty} n_s e^{-n_s \hbar\omega_s/kT} \Bigg/ \sum_{n_s=0}^{\infty} e^{-n_s \hbar\omega_s/kT} \\ &= \frac{1}{e^{\hbar\omega_s/kT} - 1}\end{aligned} \tag{4}$$

and hence

$$\langle \varepsilon_s \rangle = \hbar\omega_s \langle n_s \rangle = \frac{\hbar\omega_s}{e^{\hbar\omega_s/kT} - 1}, \tag{5}$$

identical with our earlier result (1). To obtain the number of photon states with momenta lying between $\hbar\omega/c$ and $\hbar(\omega + d\omega)/c$, Bose made use of the connection between this number and the "volume of the relevant region of the phase space", with the result

$$g(\omega)\,d\omega \approx 2 \cdot \frac{V}{h^3} \left\{ 4\pi \left(\frac{\hbar\omega}{c} \right)^2 \left(\frac{\hbar\, d\omega}{c} \right) \right\} = \frac{V\omega^2\, d\omega}{\pi^2 c^3}, \qquad (6)$$[10]

which is also identical with our earlier result (2). Thus, he finally obtained the distribution formula of Planck. It must be noted here that, *although emphasis lay elsewhere*, the mathematical steps that led Bose to the final result went literally parallel to the ones occurring in the oscillator approach!

Einstein, on the other hand, went deeper into the problem and pondered over the statistics of both the photons and the energy levels, *taken together*. He inferred (from Bose's treatment) that the basic fact to keep in mind during the process of distributing photons over the various energy levels is that the photons are *indistinguishable*—a fact that had been implicitly taken care of in Bose's treatment. Einstein's derivation of the desired distribution was essentially the same as given in Sec. 6.1, with one important difference, i.e. since the total number of photons in any given volume is indefinite, the constraint of a *fixed* N is no longer present. As a result, the Lagrange multiplier α does not enter into the discussion and to that extent the final formula for $\langle n_\varepsilon \rangle$ is simpler:[11]

$$\langle n_\varepsilon \rangle = \frac{1}{e^{\varepsilon/kT} - 1}; \tag{7}$$

cf. eqn. (6.1.18a) or (6.2.22). The foregoing result is identical with (4), with $\varepsilon = \hbar\omega_s$. The subsequent steps in Einstein's treatment were the same as in Bose's

Looking back at the two approaches, we note that there is a complete correspondence between them: "an oscillator in the eigenstate n_s, with energy $(n_s + \frac{1}{2})\hbar\omega_s$" in the first approach corresponds to "the occupation of the energy level $\hbar\omega_s$ by n_s photons" in the second approach, "the average energy $\langle\varepsilon_s\rangle$ of an oscillator" corresponds to "the mean occupation number $\langle n_s\rangle$ of the corresponding energy level", and so on.

Figure 7.7 shows a plot of the distribution function (3), which may be written in the dimensionless form

$$u'(x)\,dx = \frac{x^3\,dx}{e^x - 1}, \tag{8}$$

where

$$u'(x) = \frac{\pi^2\hbar^3c^3}{(kT)^4}u(x) \quad \text{and} \quad x = \frac{\hbar\omega}{kT}. \tag{9}$$

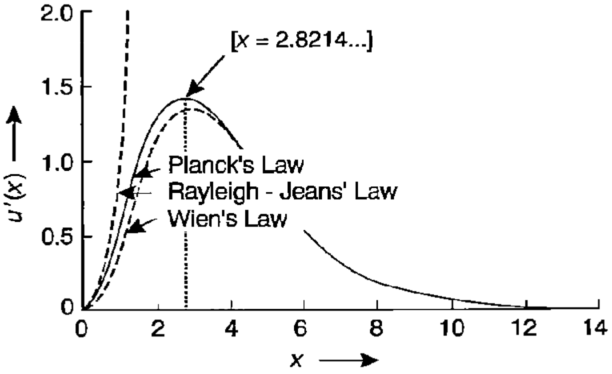


Fig. 7.7. The spectral distribution of energy in the black-body radiation. The solid curve represents the quantum-theoretical formula of Planck. The long-wavelength approximation of Rayleigh–Jeans and the short-wavelength approximation of Wien are also shown.

For long wavelengths ($x \ll 1$), formula (8) reduces to the classical approximation of Rayleigh (1900) and Jeans (1905), namely[12]

$$u'(x) \approx x^2, \tag{10}$$

while for short wavelengths ($x \gg 1$), it reduces to the rival formula of Wien (1896), namely

$$u'(x) \approx x^3e^{-x}. \tag{11}$$

For comparison, the limiting forms (10) and (11) are also included in the figure. We note that the areas under the Planck curve and the Wien curve are $\pi^4/15$ ($\simeq 6.49$) and 6, respectively. The Rayleigh–Jeans curve, however, suffers from a short-wavelength catastrophe!

For the total energy density in the cavity, we obtain from eqns (8) and (9)

$$\frac{U}{V} = \int_0^\infty u(x)\,dx = \frac{(kT)^4}{\pi^2\hbar^3c^3}\int_0^\infty \frac{x^3\,dx}{e^x - 1}$$

$$= \frac{\pi^2 k^4}{15\hbar^3 c^3} T^4. \qquad (12)^{13}$$

If there is a small opening in the walls of the cavity, the photons will "effuse" through it. The net rate of flow of the radiation, per unit area of the opening, will be given by, see eqn. (6.4.12),

$$\frac{1}{4}\frac{U}{V}c = \frac{\pi^2 k^4}{60\hbar^3 c^2} T^4 = \sigma T^4, \qquad (13)$$

where

$$\sigma = \frac{\pi^2 k^4}{60\hbar^3 c^2} = 5.670 \times 10^{-8}\ \mathrm{W\ m^{-2}\ K^{-4}}. \qquad (14)$$

Equation (13) describes the *Stefan–Boltzmann* law of black-body radiation, σ being the *Stefan constant*. This law was deduced from experimental observations by Stefan in 1879; five years later, Boltzmann derived it from thermodynamic considerations.

For further study of thermodynamics, we evaluate the grand partition function of the photon gas. Using eqn. (6.2.17) with $z = 1$, we obtain

$$\ln \mathcal{Q}(V, T) \equiv \frac{PV}{kT} = -\sum_{\varepsilon} \ln\left(1 - e^{-\varepsilon/kT}\right). \qquad (15)$$

Replacing summation by integration and making use of the extreme relativistic formula

$$a(\varepsilon)\,d\varepsilon = 2V\frac{4\pi p^2\,dp}{h^3} = \frac{8\pi V}{h^3 c^3}\varepsilon^2\,d\varepsilon, \qquad (16)$$

we obtain, after an integration by parts,

$$\ln \mathcal{Q}(V, T) \equiv \frac{PV}{kT} = \frac{8\pi V}{3h^3c^3}\frac{1}{kT}\int_0^\infty \frac{\varepsilon^3\,d\varepsilon}{e^{\varepsilon/kT} - 1}.$$

By a change of variable, this becomes

$$PV = \frac{8\pi V}{3h^3c^3}(kT)^4 \int_0^\infty \frac{x^3\,dx}{e^x - 1}$$

$$= \frac{8\pi^5 V}{45h^3c^3}(kT)^4 = \frac{1}{3}U. \qquad (17)$$

We thus obtain the well-known result of the radiation theory, i.e., the pressure of the radiation is equal to one-third its energy density; see also eqns (6.4.3) and (6.4.4). Next, since the chemical potential of the system is zero, the Helmholtz free energy is equal to $-PV$:

$$A = -PV = -\tfrac{1}{3}U, \qquad (18)$$

whence

$$S \equiv \frac{U - A}{T} = \frac{4}{3}\frac{U}{T} \propto VT^3 \qquad (19)$$

and

$$C_V = T\left(\frac{\partial S}{\partial T}\right)_V = 3S. \tag{20}$$

If the radiation undergoes a reversible adiabatic change, the law governing the variation of T with V would be, see (19),

$$VT^3 = \text{const.} \tag{21}$$

Combining (21) with the fact that $P \propto T^4$, we obtain an equation for the *adiabats* of the system, namely

$$PV^{4/3} = \text{const.} \tag{22}$$

It should be noted, however, that the ratio C_P/C_V of the photon gas is not 4/3; it is infinite!

Finally, we derive an expression for the equilibrium number $\overline{N}$ of photons in the radiation cavity. We obtain

$$\begin{aligned}\overline{N} &= \frac{V}{\pi^2 c^3}\int_0^\infty \frac{\omega^2\, d\omega}{e^{\hbar\omega/kT}-1}\\ &= V\frac{2\zeta(3)(kT)^3}{\pi^2\hbar^3 c^3} \propto VT^3\end{aligned} \tag{23}$$

Instructive though it may be, formula (23) cannot be taken at its face value because in the present problem the magnitude of the *fluctuations* in the variable N, which is determined by the quantity $(\partial P/\partial V)^{-1}$, is infinitely large; see eqn. (4.5.7).

7.3. The field of sound waves

A problem mathematically similar to the one discussed in Sec. 7.2 arises from the vibrational modes of a macroscopic body, specifically a solid. As in the case of black-body radiation, the problem of the vibrational modes of a solid can be studied equally well by regarding the system as a collection of harmonic oscillators or by regarding it as an enclosed region containing a gas of sound quanta–the so-called *phonons*. To illustrate this point, we consider the Hamiltonian of a *classical* solid composed of N atoms whose positions in space are specified by the coordinates $(x_1, x_2, \ldots, x_{3N})$. In the state of lowest energy, the values of these coordinates may be denoted by $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{3N})$. Denoting the displacements $(x_i - \bar{x}_i)$ of the atoms from their equilibrium positions by the variables $\xi_i (i = 1, 2, \ldots, 3N)$, the kinetic energy of the system in configuration (x_i) is given by

$$K = \tfrac{1}{2}m\sum_{i=1}^{3N}\dot{x}_i^2 = \tfrac{1}{2}m\sum_{i=1}^{3N}\dot{\xi}_i^2, \tag{1}$$

and the potential energy by

$$\begin{aligned}\Phi \equiv \Phi(x_i) = \Phi(\bar{x}_i) &+ \sum_i \left(\frac{\partial\Phi}{\partial x_i}\right)_{(x_i)=(\bar{x}_i)}(x_i - \bar{x}_i)\\ &+ \sum_{i,j}\frac{1}{2}\left(\frac{\partial^2\Phi}{\partial x_i \partial x_j}\right)_{(x_i)=(\bar{x}_i)}(x_i-\bar{x}_i)(x_j-\bar{x}_j) + \cdots.\end{aligned} \tag{2}$$

The main term this expansion represents the (minimum) energy of the solid when all the atoms are at rest at their mean positions $\bar{x}_i$; this energy may be denoted by the symbol Φ_0. The next set of terms is identically zero because the function $\Phi(x_i)$ has a minimum at $(x_i) = (\bar{x}_i)$ and hence all its first derivatives vanish there. The second-order terms of the expansion represent the *harmonic component* of the vibrations of the atoms about their mean positions. If we assume that the overall amplitude of these vibrations is not large we may retain only the harmonic terms of the expansion and neglect all successive ones; we are then working in the so-called *harmonic approximation*. Thus, we may write

$$H = \Phi_0 + \left\{ \sum_i \tfrac{1}{2} m \dot{\xi}_i^2 + \sum_{i,j} \alpha_{ij} \xi_i \xi_j \right\}, \tag{3}$$

where

$$\alpha_{ij} = \frac{1}{2} \left(\frac{\partial^2 \Phi}{\partial x_i \partial x_j} \right)_{(x_i)=(\bar{x}_i)} \tag{4}$$

We now introduce a linear transformation, from the coordinates ξ_i to the so-called *normal coordinates* q_i, and choose the transformation matrix such that the new expression for the Hamiltonian does not contain any cross terms, i.e.

$$H = \Phi_0 + \sum_i \tfrac{1}{2} m \left(\dot{q}_i^2 + \omega_i^2 q_i^2 \right), \tag{5}$$

where $\omega_i (i = 1, 2, \ldots, 3N)$ are the *characteristic frequencies* of the *normal modes* of the system and are determined essentially by the quantities α_{ij} or, in turn, by the nature of the potential energy function $\Phi(x_i)$. Expression (5) suggests that the energy of the solid, over and above the (minimum) value Φ_0, may be considered as arising from a set of $3N$ one-dimensional, *non-interacting*, harmonic oscillators whose characteristic frequencies ω_i are determined by the interatomic interactions in the system.

Classically, each of the $3N$ normal modes of vibration corresponds to a wave of distortion of the lattice, i.e. a sound wave. Quantum-mechanically, these modes give rise to quanta, called *phonons*, in much the same way as the vibrational modes of the electromagnetic field give rise to photons. There is one important difference, however, i.e. while the number of normal modes in the case of an electromagnetic field is infinite, the number of normal modes (or the number of phonon energy levels) in the case of a solid is fixed by the number of lattice sites.[14] This introduces certain differences in the thermodynamic behavior of the sound field in contrast to the thermodynamic behavior of the radiation field; however, at low temperatures, where the high-frequency modes of the solid are not very likely to be excited, these differences become rather insignificant and we obtain a striking similarity between the two sets of results.

The thermodynamics of the solid can now be studied along the lines of Sec. 3.8. First of all, we note that the eigenvalues of the Hamiltonian (5) are

$$E\{n_i\} = \Phi_0 + \sum_i \left(n_i + \tfrac{1}{2} \right) \hbar \omega_i, \tag{6}$$

where the numbers n_i denote the "states of excitation" of the various oscillators (or, equally well, the occupation numbers of the various phonon levels in the system). The internal energy of the system is then given by

$$U(T) = \left\{ \Phi_0 + \sum_i \frac{1}{2}\hbar\omega_i \right\} + \sum_i \frac{\hbar\omega_i}{e^{\hbar\omega_i/kT} - 1}. \tag{7}$$

The expression within the curly brackets gives the energy of the solid at absolute zero. The term Φ_0 is negative and larger in magnitude than the total zero-point energy, $\sum_i \frac{1}{2}\hbar\omega_i$, of the oscillators: together, they determine the *binding energy* of the lattice. The last term in (7) represents the temperature-dependent part of the energy,[15] which determines the specific heat of the solid:

$$C_V(T) \equiv \left(\frac{\partial U}{\partial T}\right)_V = k \sum_i \frac{(\hbar\omega_i/kT)^2 e^{\hbar\omega_i/kT}}{(e^{\hbar\omega_i/kT} - 1)^2}. \tag{8}$$

To proceed further, we need to know the frequency spectrum of the solid. To obtain this from first principles is not an easy task. Accordingly, one obtains this spectrum either through experiment or by making certain plausible assumptions about it. Einstein, who was the first to apply the quantum concept to the theory of solids (1907), assumed, for simplicity, that the frequencies ω_i are all equal. Denoting this (common) value by ω_E, the specific heat of the solid is given by

$$C_V(T) = 3NkE(x), \tag{9}$$

where $E(x)$ is the so-called *Einstein function*:

$$E(x) = \frac{x^2 e^x}{(e^x - 1)^2}, \tag{10}$$

with

$$x = \hbar\omega_E/kT = \Theta_E/T. \tag{11}$$

The dashed curve in Fig. 7.8 depicts the variation of the specific heat with temperature, as given by the Einstein formula (9). At sufficiently high temperatures, where $T \gg \Theta_E$ and hence $x \ll 1$, the Einstein result tends towards the classical one,

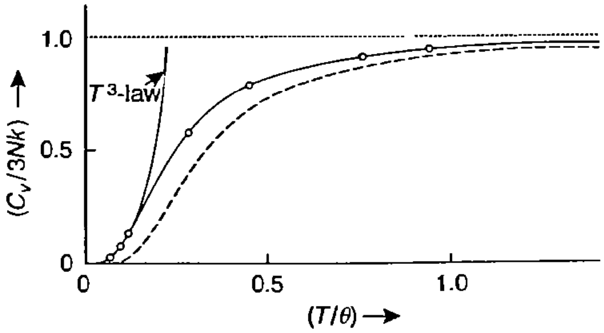


FIG. 7.8. The specific heat of a solid, according to the Ei[illegible]n model: – – –, and according to the Debye model: ——. The circles denote [illegible]xperimental results for copper.

viz. $3Nk$.[16] At sufficiently low temperatures, where $T \ll \Theta_E$ and hence $x \gg 1$, the specific heat falls exponentially fast and tends to zero as $T \to 0$. The theoretical rate of fall, however, turns out to be too fast in comparison with the observed rate. Nevertheless, Einstein's approach did at least provide a theoretical basis for understanding the observed departure of the specific heat of solids from the classical law of Dulong and Petit, whereby $C_V = 3R \simeq 5.96$ calories per mole per degree.

Debye (1912), on the other hand, allowed a *continuous spectrum* of frequencies, cut off at an upper limit ω_D such that the total number of normal modes of vibration is $3N$, that is

$$\int_0^{\omega_D} g(\omega)\,d\omega = 3N, \tag{12}$$

where $g(\omega)\,d\omega$ denotes the number of normal modes of vibration whose frequency lies in the range $(\omega, \omega + d\omega)$. For $g(\omega)$, Debye adopted the Rayleigh expression (7.2.2), modified so as to suit the problem under study. Writing c_L for the velocity of propagation of the longitudinal modes and c_T for that of the transverse modes (and noting that, for any frequency ω, the transverse mode is doubly degenerate), eqn. (12) becomes

$$\int_0^{\omega_D} V\left(\frac{\omega^2\,d\omega}{2\pi^2 c_L^3} + \frac{\omega^2\,d\omega}{\pi^2 c_T^3}\right) = 3N, \tag{13}$$

whence one obtains for the cut-off frequency

$$\omega_D^3 = 18\pi^2 \frac{N}{V}\left(\frac{1}{c_L^3} + \frac{2}{c_T^3}\right)^{-1} \tag{14}$$

Accordingly, the Debye spectrum may be written as

$$g(\omega) = \begin{cases} \dfrac{9N}{\omega_D^3}\omega^2 & \text{for} \quad \omega \le \omega_D, \\ 0 & \text{for} \quad \omega > \omega_D. \end{cases} \tag{15}$$

Before we proceed further to calculate the specific heat of solids on the basis of the Debye spectrum, two remarks seem to be in order. First, the Debye spectrum is only an idealization of the actual situation obtaining in a solid; it may, for instance, be compared with a typical spectrum such as the one shown in Fig. 7.9. While for low-frequency modes (the so-called *acoustic* modes) the Debye approximation is reasonable, serious discrepancies are seen in the case of high-frequency modes (the so-called *optical* modes). At any rate, for "averaged" quantities, such as the specific heat, the finer details of the spectrum are not very important. Second, the longitudinal and the transverse modes of the solid should have their own cut-off frequencies, $\omega_{D,L}$ and $\omega_{D,T}$ say, rather than a common cut-off at ω_D, for the simple reason that, of the $3N$ normal modes of the lattice, exactly N are longitudinal and $2N$ transverse. Accordingly, we should have, instead of (13),

$$\int_0^{\omega_{D,L}} V\frac{\omega^2\,d\omega}{2\pi^2 c_L^3} = N \quad \text{and} \quad \int_0^{\omega_{D,T}} V\frac{\omega^2\,d\omega}{\pi^2 c_T^3} = 2N. \tag{16}$$

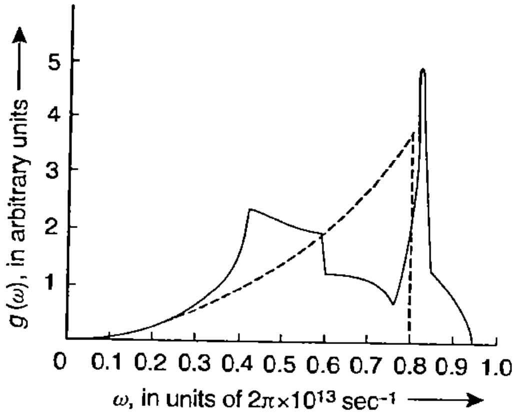


FIG. 7.9. The normal-mode frequency distribution $g(\omega)$ for aluminum. The solid curve is derived from x-ray scattering measurements (Walker, 1956) while the dashed curve represents the corresponding Debye approximation.

We note that the two cut-offs correspond to a *common wavelength* $\lambda_{\text{min}}\{= (4\pi V/3N)^{1/3}\}$, which is comparable to the *mean interatomic distance* in the solid. This is quite reasonable because, for wavelengths shorter than λ_{min}, it would be rather meaningless to speak of a wave of atomic displacements.

In the Debye approximation, formula (8) gives

$$C_V(T) = 3NkD(x_0), \tag{17}$$

where $D(x_0)$ is the so-called *Debye function*:

$$D(x_0) = \frac{3}{x_0^3}\int_0^{x_0} \frac{x^4 e^x\,dx}{(e^x-1)^2}, \tag{18}$$

with

$$x_0 = \frac{\hbar\omega_D}{kT} = \frac{\Theta_D}{T}, \tag{19}$$

Θ_D being the so-called *Debye temperature* of the solid. Integrating by parts, the expression for the Debye function becomes

$$D(x_0) = -\frac{3x_0}{e^{x_0}-1} + \frac{12}{x_0^3}\int_0^{x_0} \frac{x^3\,dx}{e^x-1}. \tag{20}$$

For $T \gg \Theta_D$, which means $x_0 \ll 1$, the function $D(x_0)$ may be expressed as a power series in x_0:

$$D(x_0) = 1 - \frac{x_0^2}{20} + \cdots. \tag{21}$$

Thus, as $T \to \infty$, $C_V \to 3Nk$; moreover, according to this theory, the classical result should be applicable to within $\frac{1}{2}$ per cent so long as $T > 3\Theta_D$. For $T \ll \Theta_D$, which means $x_0 \gg 1$, the function $D(x_0)$ may be written as

$$D(x_0) = \frac{12}{x_0^3}\int_0^{\infty} \frac{x^3\,dx}{e^x-1} + O\left(e^{-x_0}\right),$$

$$\approx \frac{4\pi^4}{5x_0^3} = \frac{4\pi^4}{5}\left(\frac{T}{\Theta_D}\right)^3 \tag{22}$$

Thus, at low temperatures the specific heat of the solid follows the *Debye T^3-law*:

$$C_V = \frac{12\pi^4}{5}Nk\left(\frac{T}{\Theta_D}\right)^3 = 464.4\left(\frac{T}{\Theta_D}\right)^3 \text{ cal mole}^{-1}\text{K}^{-1}. \tag{23}$$

It is clear from eqn. (23) that a measurement of the low-temperature specific heat of a solid should enable us not only to check the validity of the T^3-law but also to obtain an empirical value of the Debye temperature Θ_D.[17] The value of Θ_D can also be obtained by computing the cut-off frequency ω_D from a knowledge of the parameters N/V, c_L and c_T; see eqns (14) and (19). The closeness of these estimates is further evidence in favor of Debye's theory. Once Θ_D is known, the whole temperature range can be covered theoretically by making use of the tabulated values of the function $D(x_0)$.[18] A typical case was shown in Fig. 7.8. We saw that not only was the T^3-law obeyed at low temperatures, but also the agreement between theory and experiment was good throughout the range of observations.

As another illustration of agreement in the low-temperature regime, we include here another plot, Fig. 7.10, which is based on data obtained with the KCl crystal at temperatures below 5 K (Keesom and Pearlman, 1953). Here, the observed values of C_V/T are plotted against T^2. It is evident that the data fall quite well on a straight line from whose slope the value of Θ_D can be determined. One thus obtains, for KCl, $\Theta_D = 233 \pm 3$ K, which is in reasonable agreement with the values of 230–246 K coming from various estimates of the relevant elastic constants.

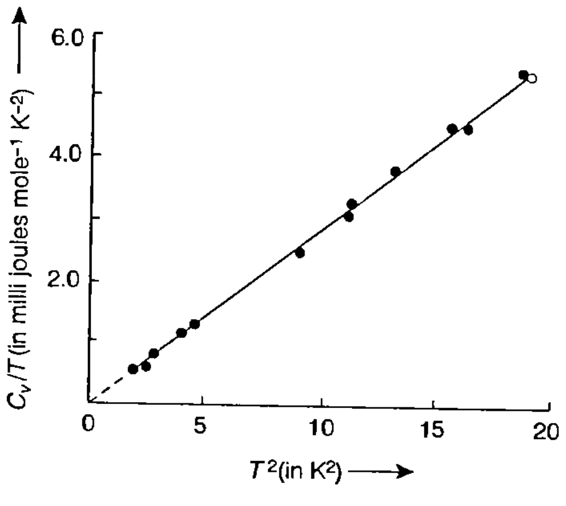


FIG. 7.10. A plot of (C_V/T) versus T^2 for KCl, showing the validity of the Debye T^3-law. The experimental points are due to Keesom and Pearlman (1953).

In Table 7.1 we list the values of Θ_D for several crystals, as derived from the specific heat measurements and from the values of the elastic constants.

In general, if the specific heat measurements of a given system conform to a T^3-law one may infer that the (thermal) excitations in the system are accounted for *solely* by phonons. We expect something similar to happen in liquids as

TABLE 7.1. THE VALUES OF THE DEBYE TEMPERATURE Θ_D FOR DIFFERENT CRYSTALS

Crystal	Pb	Ag	Zn	Cu	Al	C	NaCl	KCl	MgO
Θ_D from the specific heat measurements	88	215	308	345	398	~1850	308	233	~850
Θ_D from the elastic constants	73	214	305	332	402		320	240	~950

well, with two important differences. First, since liquids cannot withstand shear stress they cannot sustain transverse modes of vibration; a liquid composed of N atoms will, therefore, have only N (longitudinal) modes of vibration. Second, the normal modes of a liquid cannot be expected to be strictly harmonic; consequently, in addition to phonons, we might have other types of excitation such as *vortex flow*, *turbulence* (or even a modified kind of excitations, such as *rotons* in liquid He^4).

Now, helium is the only substance that remains liquid at temperatures low enough to exhibit the T^3-behavior. In the case of the lighter isotope, He^3, the results are strongly influenced by the Fermi–Dirac statistics; as a result, a specific heat proportional to the *first* power of T dominates the scene (see Sec. 8.1). In the case of the heavier isotope, He^4, the low-temperature situation is completely governed by phonons; accordingly, we expect its specific heat to be given by, see eqns (16) and (23),

$$C_V = \frac{4\pi^4}{5} Nk \left(\frac{kT}{\hbar\omega_D}\right)^3 \tag{24}$$

where

$$\omega_D = \left(\frac{6\pi^2 N}{V}\right)^{1/3} c, \tag{25}$$

c being the velocity of sound in the liquid. The specific heat per unit mass of the liquid is then given by

$$c_V = \frac{2\pi^2 k^4}{15\rho\hbar^3 c^3} T^3, \tag{26}$$

where ρ is the mass density. Substituting $\rho = 0.1455$ g/cm^3 and $c = 238$ m/sec, the foregoing result becomes

$$c_V = 0.0209 T^3 \text{ joule g}^{-1}\text{K}^{-1} \tag{27}$$

The experimental measurements of Wiebes *et al.* (1957), for $0 < T < 0.6$ K, conformed to the expression

$$c_V = (0.0204 \pm 0.0004) T^3 \text{ joule g}^{-1}\text{K}^{-1} \tag{28}$$

The agreement between the theoretical result and the experimental observations is clearly good.

7.4. Inertial density of the sound field

For further understanding of the low-temperature behavior of liquid He^4, we determine the "inertial mass" associated with a gas of sound quanta in thermal equilibrium. For this, we consider "a phonon gas in mass motion", for then by determining the relationship between the momentum $\boldsymbol{P}$ of the gas and the velocity $\boldsymbol{v}$ of its mass motion we can readily evaluate the property in question. Now, since the total number of phonons in the system is indefinite, the problem is free from the constraint of a fixed N; consequently, the undetermined multiplier α may be taken to be identically zero. However, we now have a new constraint on the system, namely that of a fixed total momentum $\boldsymbol{P}$, additional to the constraint of the fixed total energy E. Under these constraints, the mean occupation number of the phonon level $\varepsilon(\boldsymbol{p})$ would be

$$\langle n(\boldsymbol{p})\rangle = \frac{1}{\exp(\beta\varepsilon + \boldsymbol{\gamma}\cdot\boldsymbol{p}) - 1}. \tag{1}$$

As usual, the parameter β is equal to $1/kT$. To determine $\boldsymbol{\gamma}$ it seems natural to evaluate the drift velocity of the gas. Choosing the z-axis to be in the direction of the mass motion, the magnitude v of the drift velocity will be given by "the mean value of the component u_z of the individual phonon velocities":

$$v = \langle u\cos\theta\rangle \tag{2}$$

Now, for phonons

$$\varepsilon = pc \quad \text{and} \quad u \equiv \frac{d\varepsilon}{dp} = c, \tag{3}$$

where c is the *velocity of sound* in the medium. Moreover, by reasons of symmetry, we expect the undetermined vector $\boldsymbol{\gamma}$ to be either parallel or antiparallel to the direction of mass motion; hence, we may write

$$\boldsymbol{\gamma}\cdot\boldsymbol{p} = \gamma_z p_z = \gamma_z p\cos\theta. \tag{4}$$

In view of eqns (1), (3) and (4), eqn. (2) becomes

$$v = \frac{\int\limits_0^\infty\int\limits_0^\pi [\exp\{\beta pc(1 + (\gamma_z/\beta c)\cos\theta)\} - 1]^{-1}(c\cos\theta)(p^2\,dp\,2\pi\sin\theta\,d\theta)}{\int\limits_0^\infty\int\limits_0^\pi [\exp\{\beta pc(1 + (\gamma_z/\beta c)\cos\theta)\} - 1]^{-1}(p^2\,dp\,2\pi\sin\theta\,d\theta)} \tag{5}$$

Making the substitutions

$$\cos\theta = \eta, \quad p(1 + (\gamma_z/\beta c)\eta) = p'$$

and cancelling away the integrations over p', we obtain

$$v = c\frac{\int\limits_{-1}^{1}(1 + (\gamma_z/\beta c)\eta)^{-3}\eta\,d\eta}{\int\limits_{-1}^{1}(1 + (\gamma_z/\beta c)\eta)^{-3}\,d\eta} = -\gamma_z/\beta.$$

It follows that

$$\gamma = -\beta \boldsymbol{v}. \tag{6}$$

Accordingly, the expression for the mean occupation number becomes

$$\langle n(\boldsymbol{p})\rangle = \frac{1}{\exp\{\beta(\varepsilon - \boldsymbol{v}\cdot\boldsymbol{p})\} - 1}. \tag{7}$$

A comparison of (7) with the corresponding result in the rest frame of the gas, namely

$$\langle n_0(\boldsymbol{p}_0)\rangle = \frac{1}{\exp(\beta\varepsilon_0) - 1}, \tag{8}$$

shows that the change caused by the imposition of mass motion on the system is nothing but a straightforward manifestation of *Galilean transformation* between the two frames of reference.

Alternatively, eqn. (7) may be written as

$$\langle n(\boldsymbol{p})\rangle = \frac{1}{\exp(\beta p'c) - 1} = \frac{1}{\exp\{\beta pc(1 - (v/c)\cos\theta)\} - 1}. \tag{9}$$

As such, formula (9) lays down a serious restriction on the drift velocity v, i.e. it must not exceed c, the velocity of the phonons, for otherwise some of the occupation numbers would become negative! Actually, as our subsequent analysis will show, the formalism developed in this section breaks down as v approaches c. The velocity c may, therefore, be regarded as a *critical velocity* for the flow of the phonon gas:

$$(v_c)_{\text{ph}} = c. \tag{10}$$

The relevance of this result to the problem of superfluidity in liquid helium II will be seen in the following section.

Next we calculate the total momentum $\boldsymbol{P}$ of the phonon gas:

$$\boldsymbol{P} = \sum_{\boldsymbol{p}} \langle n(\boldsymbol{p})\rangle \boldsymbol{p}. \tag{11}$$

Indeed, the vector $\boldsymbol{P}$ will be parallel to the vector $\boldsymbol{v}$, the latter being already in the direction of the z-axis. We have, therefore, to calculate only the z-component of the momentum:

$$\begin{aligned}
P = P_z &= \sum_{\boldsymbol{p}} \langle n(\boldsymbol{p})\rangle p_z \\
&= \int_0^\infty \int_0^\pi \frac{p\cos\theta}{\exp\{\beta pc(1 - (v/c)\cos\theta)\} - 1}\left(\frac{Vp^2\,dp\,2\pi\sin\theta\,d\theta}{h^3}\right) \\
&= \frac{2\pi V}{h^3}\int_0^\infty \frac{p'^3\,dp'}{\exp(\beta p'c) - 1}\int_0^\pi \{1 - (v/c)\cos\theta\}^{-4}\cos\theta\sin\theta\,d\theta \\
&= V\frac{16\pi^5}{45h^3c^3\beta^4}\cdot\frac{v/c^2}{(1 - v^2/c^2)^3}.
\end{aligned} \tag{12}$$

The total energy E of the gas is given by

$$E = \sum_p \langle n(p) \rangle pc$$

$$= \frac{2\pi Vc}{h^3} \int_0^\infty \frac{p'^3 \, dp'}{\exp(\beta p'c) - 1} \int_0^\pi \{1 - (v/c)\cos\theta\}^{-4} \sin\theta \, d\theta$$

$$= V \frac{4\pi^5}{15h^3c^3\beta^4} \cdot \frac{1 + \frac{1}{3}v^2/c^2}{(1 - v^2/c^2)^3}. \tag{13}$$

It is now natural to regard the ratio P/v as the "inertial mass" of the gas. The corresponding mass density ρ is, therefore, given by

$$\rho = \frac{P}{vV} = \frac{16\pi^5 k^4 T^4}{45h^3c^5} \frac{1}{(1 - v^2/c^2)^3}. \tag{14}$$

For $(v/c) \ll 1$, which is generally true, the mass density of the phonon gas is given by

$$(\rho_0)_{\text{ph}} = \frac{16\pi^5 k^4}{45h^3c^5} T^4 = \frac{4}{3c^2}(E_0/V). \tag{15}$$

Substituting the value of c for liquid He^4 (at low temperatures), the phonon mass density, as a fraction of the actual density of the liquid, is given by

$$(\rho_0)_{\text{ph}}/\rho_{\text{He}} = 1.22 \times 10^{-4} T^4; \tag{16}$$

thus, for example, at $T = 0.3$ K the value of this fraction turns out to be about 9.9×10^{-7}. Now, at a temperature like 0.3 K, phonons are the only excitations

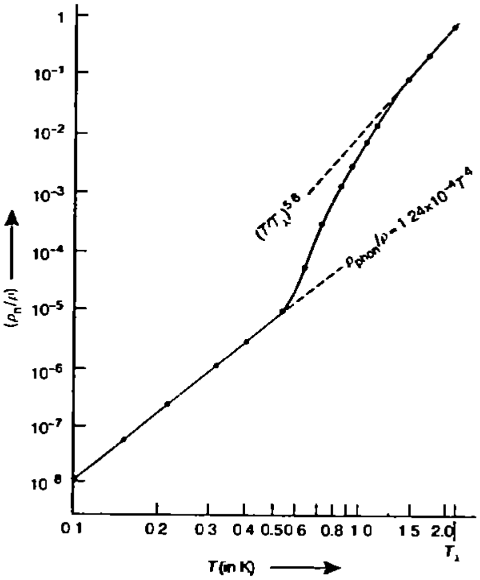


FIG. 7.11. The normal fraction (ρ_n/ρ), as obtained from experimental data on (i) the velocity of second sound and (ii) the entropy of liquid He II (after de Klerk, Hudson and Pellam, 1953).

in liquid He^4 that need to be considered; the calculated result should, therefore, correspond to the "ratio of the density ρ_n of the *normal fluid* in the liquid to the total density ρ of the liquid". It is practically impossible to make a *direct* determination of a fraction as small as that; however, *indirect* evaluations that make use of other experimentally viable properties of the liquid provide a striking confirmation of the foregoing result; see Fig. 7.11.

7.5. Elementary excitations in liquid helium II

Landau (1941, 1947) developed a simple theoretical scheme which explains reasonably well the behavior of liquid helium II at low temperatures not too close to the λ-point. According to this scheme, the liquid is treated as a weakly excited quantum-mechanical system, in which deviations from the ground state ($T = 0$ K) are described in terms of "a gas of elementary excitations" hovering over a *quiescent* background. The gas of excitations corresponds to the "normal fluid", while the quiescent background represents the "superfluid". At $T = 0$ K, there are no excitations at all ($\rho_n = 0$) and the whole of the fluid constitutes the superfluid background ($\rho_s = \rho_{\mathrm{He}}$). At higher temperatures, we may write

$$\rho_s(T) = \rho_{He}(T) - \rho_n(T), \tag{1}$$

so that at $T = T_\lambda$, $\rho_n = \rho_{\mathrm{He}}$ and $\rho_s = 0$. At $T > T_\lambda$, the liquid behaves in all respects as a normal fluid, commonly known as liquid helium I.

Guided by purely empirical considerations, Landau also proposed an energy–momentum relationship $\varepsilon(p)$ for the elementary excitations in liquid helium II. At low momenta, the relationship between ε and p was linear (which is characteristic of phonons), while at higher momenta it exhibited a non-monotonic character. The excitations were assumed to be bosons and, at low temperatures (when their number is not very large), mutually non-interacting; the macroscopic properties of the liquid could then be calculated by following a straightforward statistical-mechanical approach. It was found that Landau's theory could explain quite successfully the observed properties of liquid helium II over a temperature range of about 0–2 K; however, it still remained to be verified that the actual excitations in the liquid did, in fact, conform to the proposed energy spectrum.

Following a suggestion by Cohen and Feynman (1957), a number of experimental workers set out to investigate the spectrum of excitations in liquid helium II by scattering long-wavelength neutrons ($\lambda \gtrsim 4$ Å) from the liquid. At temperatures below 2 K, the most important scattering process is the one in which a neutron creates a *single* excitation in the liquid. By measuring the modified wavelength λ_f of the neutrons scattered at an angle ϕ, the energy ε and the momentum p of the excitation created in the scattering process could be determined on the basis of the relevant conservation laws:

$$\varepsilon = h^2(\lambda_i^{-2} - \lambda_f^{-2})/2m, \tag{2}$$

$$p^2 = h^2(\lambda_i^{-2} + \lambda_f^{-2} - 2\lambda_i^{-1}\lambda_f^{-1}\cos\phi), \tag{3}$$

where λ_i is the initial wavelength of the neutrons and m the neutron mass. By arying ϕ, or λ_i, one could map the entire spectrum of the excitations.

The first exhaustive investigation along these lines was carried out by Yarnell t *al.* (1959); their results, shown in Fig. 7.12, possess a striking resemblance to e empirical spectrum proposed by Landau. The more important features of the pectrum, which was obtained at a temperature of 1.1 K, are the following:

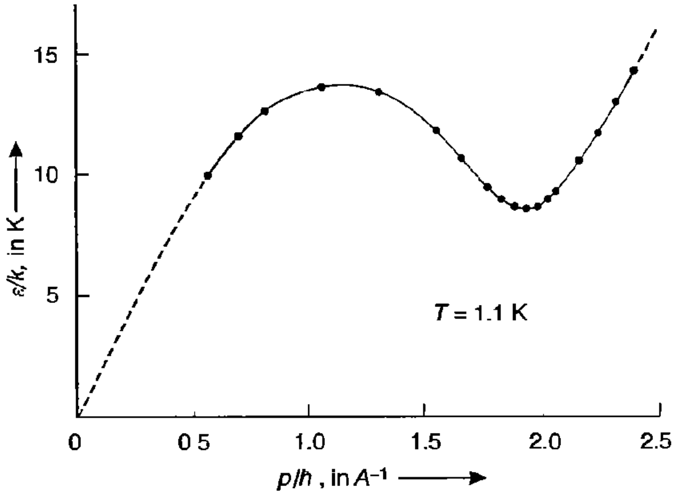


FIG. 7.12. The energy spectrum of the elementary excitations in liquid He II at 1.1 K (after Yarnell *et al.*, 1959); the dashed line emanating from the origin has a slope corresponding to the velocity of sound in the liquid, viz. (239±5) m/sec.

(i) If we fit a linear, *phonon-like* spectrum ($\varepsilon = pc$) to points in the vicinity of $p/\hbar = 0.55$ Å^{-1}, we obtain for c a value of (239±5) m/sec, which is in excellent agreement with the measured value of the velocity of sound in the liquid, viz. about 238 m/sec.

(ii) The spectrum passes through a maximum value of $\varepsilon/k = (13.92 \pm 0.10)$ K at $p/\hbar = (1.11 \pm 0.02)$ Å^{-1}.

(iii) This is followed by a minimum at $p/\hbar = (1.92 \pm 0.01)$ Å^{-1}, whose neighborhood may be represented by Landau's *roton* spectrum:

$$\varepsilon(p) = \Delta + \frac{(p - p_0)^2}{2\mu}, \tag{4}$$

with

$$\Delta/k = (8.65 \pm 0.04)\ \text{K}$$
$$p_0/\hbar = (1.92 \pm 0.01)\ \text{Å}^{-1} \qquad (5)^{19}$$

and

$$\mu = (0.16 \pm 0.01) m_{\text{He}}.$$

(iv) Above $p/\hbar \simeq 2.18$ Å^{-1}, the spectrum rises linearly, again with a slope equal to c.

Data were also obtained at temperatures of 1.6 K and 1.8 K. The spectrum was found to be of the same general shape as at 1.1 K; only the value of Δ was slightly lower.

In a later investigation, Henshaw and Woods (1961) extended the range of observation at both ends of the spectrum; their results are shown in Fig. 7.13. On the lower side, they carried out measurements down to $p/\hbar = 0.26$ Å^{-1} and found that the experimental points indeed lie on a straight line (of slope 237 m/sec). On the upper side, they pushed their measurements up to $p/\hbar = 2.68$ Å^{-1} and found that after passing through a minimum at 1.91 Å^{-1}, the curve rises with an increasing slope up to about 2.4 Å^{-1} at which point the second derivative $\partial^2\varepsilon/\partial p^2$ changes sign; the subsequent trend of the curve suggests the possible existence of a second maximum in the spectrum![20]

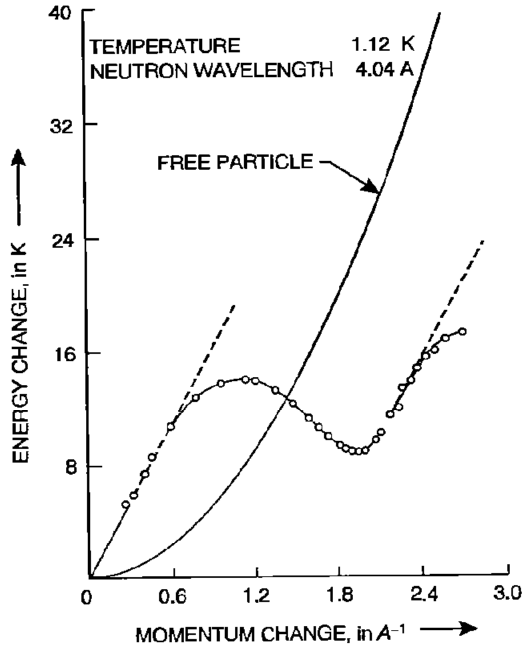


FIG. 7.13. The energy spectrum of the elementary excitations in liquid He II at 1.12 K (after Henshaw and Woods, 1961); the dashed straight lines have a common slope corresponding to the velocity of sound in the liquid, viz. 237 m/sec. The parabolic curve rising from the origin represents the energy spectrum, $\varepsilon(p) = p^2/2m$, of *free* helium atoms.

To evaluate the thermodynamics of liquid helium II, we first of all note tha at sufficiently low temperatures we have only low-lying excitations, namely the phonons. The thermodynamic behavior of the liquid is then governed by formulae derived in Secs 7.3 and 7.4. At temperatures higher than about 0.5 K, the seco group of excitations, namely the rotons (with momenta in the vicinity of p_0), al shows up. Between 0.5 K and about 1 K, the behavior of the liquid is governe by phonons and rotons together. Above 1 K, however, the phonon contributi to the various thermodynamic properties of the liquid become rather unimport then, rotons are the only excitations that need to be considered.

We shall now study the temperature dependence of the roton contributions the various thermodynamic properties of the liquid. In view of the continuit the energy spectrum, it is natural to expect that, like phonons, rotons also obey Bose–Einstein statistics. Moreover, their total number N in the system is quite indefinite; consequently, their chemical potential μ is identically zero. We then have for the mean occupation numbers of the rotons

$$\langle n(p)\rangle = \frac{1}{\exp\{\beta\varepsilon(p)\} - 1}, \tag{6}$$

where $\varepsilon(p)$ is given by eqns (4) and (5). Now, at all temperatures of interest (viz. $T \leq 2$ K), the minimum value of the term $\exp\{\beta\varepsilon(p)\}$, namely $\exp(\Delta/kT)$, is considerably larger than unity. We may, therefore, write

$$\langle n(p)\rangle \simeq \exp\{-\beta\varepsilon(p)\}. \tag{7}$$

The grand potential of the system of rotons is, therefore, given by

$$q(V,T) \equiv \frac{PV}{kT} = -\sum_p \ln[1 - \exp\{-\beta\varepsilon(p)\}] \simeq \sum_p \exp\{-\beta\varepsilon(p)\} \simeq \overline{N}, \tag{8}$$

where $\overline{N}$ is the "equilibrium" number of rotons in the system. The summation over p may be replaced by an integration, with the result

$$\frac{PV}{kT} = \overline{N} = \frac{V}{h^3}\int_0^\infty e^{-\left\{\Delta + \frac{(p-p_0)^2}{2\mu}\right\}/kT}(4\pi p^2\,dp). \tag{9}$$

Substituting $p = p_0 + (2\mu kT)^{1/2}x$, we get

$$\frac{PV}{kT} = \overline{N} = \frac{4\pi p_0^2 V}{h^3}e^{-\Delta/kT}(2\mu kT)^{1/2}\int e^{-x^2}\left\{1 + \frac{(2\mu kT)^{1/2}}{p_0}x\right\}^2 dx. \tag{10}$$

The "relevant" range of the variable x, the one that makes significant contribution towards the integral, is fairly symmetric about the value $x = 0$; consequently, the net effect of the *linear* term in the integrand is vanishingly small. The *quadratic* term too is unimportant because its coefficient $(2\mu kT)/p_0^2 \ll 1$. Thus, all we have to consider is the integral of $\exp(-x^2)$. Now one can readily verify that the limits of this integral are such that, without seriously affecting the value of the integral, they may be taken as $-\infty$ and $+\infty$; the value of the integral is then simply $\pi^{1/2}$. We thus obtain

$$\frac{PV}{kT} = \overline{N} = \frac{4\pi p_0^2 V}{h^3}(2\pi\mu kT)^{1/2}e^{-\Delta/kT}. \qquad (11)^{21}$$

The free energy of the roton gas is given by (since $\mu = 0$)

$$A = -PV = -\overline{N}kT \propto T^{3/2}e^{-\Delta/kT}, \tag{12}$$

whence

$$S = -\left(\frac{\partial A}{\partial T}\right)_V = -A\left\{\frac{3}{2T} + \frac{\Delta}{kT^2}\right\} = \overline{N}k\left\{\frac{3}{2} + \frac{\Delta}{kT}\right\}, \tag{13}$$

$$U = A + TS = \overline{N}\left(\Delta + \frac{1}{2}kT\right) \qquad (14)^{22}$$

and

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V = \bar{N}k\left\{\frac{3}{4} + \frac{\Delta}{kT} + \left(\frac{\Delta}{kT}\right)^2\right\} \tag{15}$$

Clearly, as $T \to 0$, all these results tend to zero (essentially exponentially).

We shall now determine the inertia of the roton gas. Proceeding as in Sec. 7.4, we obtain for a gas of excitations with energy spectrum $\varepsilon(p)$

$$\rho_0 = \frac{M_0}{V} = \lim_{v \to 0} \frac{1}{v} \int n(\varepsilon - \boldsymbol{v}\cdot\boldsymbol{p})p\frac{d^3p}{h^3}, \tag{16}$$

where $n(\varepsilon - \boldsymbol{v}\cdot\boldsymbol{p})$ is the mean occupation number of the state $\varepsilon(p)$, as observed in a frame of reference K with respect to which the gas is in *mass motion* with a drift velocity $\boldsymbol{v}$.[23] For small v, the function $n(\varepsilon - \boldsymbol{v}\cdot\boldsymbol{p})$ might be expanded as a Taylor series in $\boldsymbol{v}$ and only the terms $n(\varepsilon) - (\boldsymbol{v}\cdot\boldsymbol{p})\,\partial n(\varepsilon)/\partial\varepsilon$ retained. The integral over the first part denotes the momentum density of the system, as observed in the rest frame K_0, and is identically zero. We are thus left with

$$\begin{aligned}\rho_0 &= -\frac{1}{h^3}\int p^2\cos^2\theta\frac{\partial n(\varepsilon)}{\partial\varepsilon}(p^2\,dp\,2\pi\sin\theta\,d\theta)\\ &= -\frac{4\pi}{3h^3}\int_0^\infty \frac{\partial n(\varepsilon)}{\partial\varepsilon}p^4\,dp,\end{aligned} \tag{17}$$

which holds for *any* energy spectrum and for *any* statistics. For phonons, in particular, we obtain

$$\begin{aligned}(\rho_0)_{\text{ph}} &= -\frac{4\pi}{3h^3c}\int_0^\infty \frac{dn(p)}{dp}p^4\,dp\\ &= -\frac{4\pi}{3h^3c}\left[n(p)\cdot p^4\Big|_0^\infty - \int_0^\infty n(p)\cdot 4p^3\,dp\right]\\ &= \frac{4}{3c^2}\int_0^\infty n(p)\cdot pc\left(\frac{4\pi p^2\,dp}{h^3}\right) = \frac{4}{3c^2}(E_0)_{\text{ph}}/V,\end{aligned} \tag{18}$$

which is identical with our earlier result (7.4.15).

For rotons, $n(\varepsilon) \simeq \exp(-\beta\varepsilon)$; hence, $\partial n(\varepsilon)/\partial\varepsilon \simeq -\beta n(\varepsilon)$. Accordingly,

$$\begin{aligned}(\rho_0)_{\text{rot}} &= \frac{4\pi\beta}{3h^3}\int n(\varepsilon)p^4\,dp\\ &= \frac{\beta}{3}\langle p^2\rangle\frac{\bar{N}}{V} \simeq \frac{p_0^2}{3kT}\frac{\bar{N}}{V} \qquad (19)\\ &= \frac{4\pi p_0^4}{3h^3}\left(\frac{2\pi\mu}{kT}\right)^{1/2}e^{-\Delta/kT}; \qquad (20)\end{aligned}$$

At very low temperatures ($T < 0.3$ K), the roton cor[illegible]ution towards the inertia of the fluid is negligible in comparison with the phon[illegible]ontribution. At relatively higher temperatures ($T \sim 0.6$ K), the two contributions become comparable. At temperatures above 1 K, the roton contribution is far more dominant than the phonon contribution; at such temperatures, the roton density alone accounts for the density ρ_n of the normal fluid.

It would be instructive to determine the *critical temperature* T_c at which the theoretical value of the density ρ_n became equal to the actual density ρ_{He} of the liquid; this would imply the disappearance of the *superfluid component* from the liquid (and hence a transition from liquid He II to liquid He I). In this manner, we find: $T_c \simeq 2.5$ K, as opposed to the experimental value of T_λ, which is $\simeq 2.19$ K. The comparison is not too bad, considering the fact that in the present calculation we have assumed the roton gas to be a *non-interacting* system right up to the transition point; in fact, due to the presence of an exceedingly large number of excitations at higher temperatures, this assumption would not remain plausible.

Equation (19) suggests that a roton excitation possesses an *effective mass* $p_0^2/3kT$. Numerically, this is about 10–15 times the mass of a helium atom (and, hence, orders of magnitude larger than the parameter μ of the roton spectrum). However, the more important aspect of the roton effective mass is that it is *inversely* proportional to the temperature of the roton gas! Historically, this aspect was first discovered empirically by Landau (1947) on the basis of the experimental data on the velocity of second sound in liquid He II and its specific heat. Now, since the effective mass of an excitation is generally determined by the quantity $\langle p^2\rangle/3kT$, Landau concluded that the quantity $\langle p^2\rangle$ of the relevant excitations in the liquid must be temperature-independent. Thus, as the temperature of the liquid rises the mean value of p^2 (of the excitations) must stay constant; this value may be denoted by p_0^2. The mean value of ε, on the other hand, must rise with temperature. The only way to reconcile the two was to invoke a *nonmonotonic* spectrum with a minimum at $p = p_0$.

Finally, we would like to touch upon the question of the *critical velocity* of superflow. For this we consider a mass M of excitation-free superfluid in mass motion; its kinetic energy E and momentum $\boldsymbol{P}$ are given by $\frac{1}{2}Mv^2$ and $M\boldsymbol{v}$, respectively. Any changes in these quantities are related as follows:

$$\delta E = (\boldsymbol{v}\cdot\delta\boldsymbol{P}). \tag{21}$$

Supposing that these changes came about as a result of the creation of an excitation $\varepsilon(\boldsymbol{p})$ in the fluid, we must have, by the principles of conservation,

$$\delta E = -\varepsilon \quad \text{and} \quad \delta\boldsymbol{P} = -\boldsymbol{p}. \tag{22}$$

Equations (21) and (22) lead to the result

$$\varepsilon = (\boldsymbol{v}\cdot\boldsymbol{p}) \leq vp. \tag{23}$$

Thus, it is impossible to create an excitation $\varepsilon(p)$ in the fluid unless the drift velocity v of the fluid is greater than, or at least equal to, the quantity (ε/p). Accordingly, if v is less than even the lowest value of ε/p, no excitation at all can be created in the fluid, which will therefore maintain its superfluid character. We thus obtain a condition for the maintenance of superfluidity, viz.

$$v < v_c = (\varepsilon/p)_{\mathrm{min}}, \tag{24}$$

which is known as the *Landau criterion* for superf[illegible] The velocity v_c is called the *critical velocity* of superflow; it marks an "upper limit" to the flow velocities at which the fluid exhibits superfluid behavior. The observed magnitude of the critical velocity varies significantly with the geometry of the channel employed; as a rule, the narrower the channel the larger the critical velocity. The observed values of v_c range from about 0.1 cm/sec to about 70 cm/sec.

The theoretical estimates of v_c are clearly of interest. On one hand, we find that if the excitations obey the ideal-gas relationship, viz. $\varepsilon = p^2/2m$, then the critical velocity turns out to be exactly zero. Any velocity v is then greater than the critical velocity; accordingly, no superflow is possible at all. This is a very significant result, for it brings out very clearly the fact that interatomic interactions in the liquid, which give rise to an excitation spectrum different from the one characteristic of the ideal gas, play a fundamental role in bringing about the phenomenon of superfluidity. Thus, while an ideal Bose gas does undergo the phenomenon of Bose–Einstein condensation, it cannot support the phenomenon of superfluidity as such. On the other hand, we find that (i) for phonons, $v_c = c \simeq 2.4 \times 10^4$ cm/sec and (ii) for rotons, $v_c = \{(p_0^2 + 2\mu\Delta)^{1/2} - p_0\}/\mu \simeq \Delta/p_0 \simeq 6.3 \times 10^3$ cm/sec, which are too high in comparison with the observed values of v_c. In fact, there is another type of collective excitations which can appear in liquid helium II, viz. *quantized vortex rings*, with an energy–momentum relationship of the form: $\varepsilon \propto p^{1/2}$. The critical velocity for the creation of these rings turns out to be numerically consistent with the experimental findings; not only that, the dependence of v_c on the geometry of the channel can also be understood in terms of the size of the rings created.

For a review of this topic, especially in regard to Feynman's contributions, see Mehra and Pathria (1994); see also Secs 10.4–10.6 of this text.

Problems

7.1. By considering the order of magnitude of the occupation numbers $\langle n_\varepsilon \rangle$, show that it makes no difference to the final results of Sec. 7.1 if we combine a *finite* number of $(\varepsilon \neq 0)$-terms of the sum (7.1.2) with the $(\varepsilon = 0)$-part of eqn. (7.1.6) or include them in the integral over ε.

7.2. Deduce the virial expansion (7.1.13) from eqns (7.1.7) and (7.1.8), and verify the quoted values of the virial coefficients.

7.3. Combining eqns (7.1.24) and (7.1.26), and making use of the first two terms of formula (D. 9), show that, as T approaches T_c from above, the parameter $\alpha(= -\ln z)$ of the ideal Bose gas assumes the form

$$\alpha \approx \frac{1}{\pi}\left(\frac{3\zeta(3/2)}{4}\right)^2 \left(\frac{T - T_c}{T_c}\right)^2$$

7.4. Show that for an ideal Bose gas

$$\frac{1}{z}\left(\frac{\partial z}{\partial T}\right)_P = -\frac{5}{2T}\frac{g_{5/2}(z)}{g_{3/2}(z)};$$

cf. eqn. (7.1.36). Hence show that

$$\gamma \equiv \frac{C_P}{C_V} = \frac{(\partial z/\partial T)_P}{(\partial z/\partial T)_v} = \frac{5}{3}\frac{g_{5/2}(z)g_{1/2}(z)}{\{g_{3/2}(z)\}^2},$$

as in eqn. (7.1.48b). Check that, as T approaches T_c from above, both γ and C_P diverge as $(T - T_c)^{-1}$.

7.5. (a) Show that the isothermal compressibility κ_T and the adiabatic compressibility κ_S of an ideal Bose gas are given by

$$\kappa_T = \frac{1}{nkT}\frac{g_{1/2}(z)}{g_{3/2}(z)}, \quad \kappa_S = \frac{3}{5nkT}\frac{g_{3/2}(z)}{g_{5/2}(z)},$$

where $n(= N/V)$ is the particle density in the gas. Note that, as $z \to 0$, κ_T and κ_S approach their respective classical values, viz. $1/P$ and $1/\gamma P$. How do they behave as $z \to 1$?

(b) Making use of the thermodynamic relations

$$C_P - C_V = T\left(\frac{\partial P}{\partial T}\right)_V \left(\frac{\partial V}{\partial T}\right)_P = TV\kappa_T\left(\frac{\partial P}{\partial T}\right)_V^2$$

and

$$C_P/C_V = \kappa_T/\kappa_S,$$

derive eqns (7.1.48a) and (7.1.48b).

7.6. Show that for an ideal Bose gas the temperature derivative of the specific heat C_V is given by

$$\frac{1}{Nk}\left(\frac{\partial C_V}{\partial T}\right)_V = \begin{cases} \dfrac{1}{T}\left[\dfrac{45}{8}\dfrac{g_{5/2}(z)}{g_{3/2}(z)} - \dfrac{9}{4}\dfrac{g_{3/2}(z)}{g_{1/2}(z)} - \dfrac{27}{8}\dfrac{\{g_{3/2}(z)\}^2 g_{-1/2}(z)}{\{g_{1/2}(z)\}^3}\right] & \text{for} \quad T > T_c, \\ \dfrac{45}{8}\dfrac{v}{T\lambda^3}\zeta\left(\dfrac{5}{2}\right) & \text{for} \quad T < T_c. \end{cases}$$

Using these results and the main term of formula (D.9), verify eqn. (7.1.38).

7.7. Evaluate the quantities $(\partial^2 P/\partial T^2)_v$, $(\partial^2 \mu/\partial T^2)_v$ and $(\partial^2 \mu/\partial T^2)_P$ for an ideal Bose gas and check that your results satisfy the thermodynamic relationships

$$C_V = VT\left(\frac{\partial^2 P}{\partial T^2}\right)_v - NT\left(\frac{\partial^2 \mu}{\partial T^2}\right)_v$$

and

$$C_P = -NT\left(\frac{\partial^2 \mu}{\partial T^2}\right)_P.$$

Examine the behavior of these quantities as $T \to T_c$ from above and from below.

7.8. The velocity of sound in a fluid is given by the formula

$$w = \sqrt{(\partial P/\partial \rho)_s},$$

where ρ is the mass density of the fluid. Show that for an ideal Bose gas

$$w^2 = \frac{5kT}{3m}\frac{g_{5/2}(z)}{g_{3/2}(z)} = \frac{5}{9}\langle u^2 \rangle,$$

where $\langle u^2 \rangle$ is the mean square speed of the particles in the gas.

7.9. Show that for an ideal Bose gas

$$\langle u \rangle \langle \frac{1}{u} \rangle = \frac{4}{\pi}\frac{g_1(z)g_2(z)}{\{g_{3/2}(z)\}^2},$$

u being the speed of a particle. Examine and interpret the limiting cases $z \to 0$ and $z \to 1$; cf. Problem 6.6.

7.10. Consider an ideal Bose gas in a uniform gravitational field (of acceleration g). Show that the phenomenon of Bose–Einstein condensation sets in at a temperature T_c given by

$$T_c \simeq T_c^0\left[1 + \frac{8}{9}\frac{1}{\zeta\left(\frac{3}{2}\right)}\left(\frac{\pi mgL}{kT_c^0}\right)^{1/2}\right],$$

where L is the height of the container and $mgL \ll kT_c^0$. Also show that the condensation is accompanied by a discontinuity in the specific heat C_V of the gas:

$$(\Delta C_V)_{T=T_c} \simeq -\frac{9}{8\pi}\zeta\left(\frac{3}{2}\right) Nk \left(\frac{\pi mgL}{kT_c^0}\right)^{1/2}$$

see Eisenschitz (1958).

7.11. Consider an ideal Bose gas consisting of molecules with internal degrees of freedom. Assuming that, besides the ground state $\varepsilon_0 = 0$, only the first excited state ε_1 of the *internal* spectrum needs to be taken into account, determine the condensation temperature of the gas as a function of ε_1. Show that, for $(\varepsilon_1/kT_c^0) \gg 1$,

$$\frac{T_c}{T_c^0} \simeq 1 - \frac{\frac{2}{3}}{\zeta\left(\frac{3}{2}\right)} e^{-\varepsilon_1/kT_c^0}$$

while, for $(\varepsilon_1/kT_c^0) \ll 1$,

$$\frac{T_c}{T_c^0} \simeq \left(\frac{1}{2}\right)^{2/3}\left[1 + \frac{2^{4/3}}{3\zeta\left(\frac{3}{2}\right)}\left(\frac{\pi\varepsilon_1}{kT_c^0}\right)^{1/2}\right].$$

[*Hint*: To obtain the last result, use the first two terms of formula (D.9).]

7.12. Consider an ideal Bose gas in the grand canonical ensemble and study fluctuations in the total number of particles N and the total energy E Discuss, in particular, the situation when the gas becomes highly degenerate.

7.13. Consider an ideal Bose gas confined to a region of area A in *two* dimensions. Express the number of particles in the excited states, N_e, and the number of particles in the ground state, N_0, in terms of z, T and A, and show that the system does not exhibit Bose–Einstein condensation unless $T \to 0$ K.

Refine your argument to show that, if the area A and the total number of particles N are held fixed and we require both N_e and N_0 to be of order N, we do achieve condensation when

$$T \sim \frac{h^2}{mkl^2}\frac{1}{\ln N},$$

where $l[\sim \sqrt{(A/N)}]$ is the mean interparticle distance in the system. Of course, if both A and $N \to \infty$, keeping l fixed, then the desired T does go to zero.

7.14. Consider an n-dimensional Bose gas whose single-particle energy spectrum is given by $\varepsilon \propto p^s$, where s is some positive number. Discuss the onset of Bose–Einstein condensation in this system, especially its dependence on the numbers n and s. Study the thermodynamic behavior of the system and show that, quite generally,

$$P = \frac{s}{n}\frac{U}{V}, \quad C_V(T \to \infty) = \frac{n}{s}Nk, \quad \text{and} \quad C_P(T \to \infty) = \left(\frac{n}{s} + 1\right)Nk.$$

7.15. The (canonical) partition function of the black-body radiation may be written as

$$Q(V, T) = \prod_\omega Q_1(\omega, T),$$

so that

$$\ln Q(V, T) = \sum_\omega \ln Q_1(\omega, T) \approx \int_0^\infty \ln Q_1(\omega, T) g(\omega)\, d\omega;$$

here, $Q_1(\omega, T)$ is the single-oscillator partition function given by eqn. (3.8.14) and $g(\omega)$ is the density of states given by eqn. (7.2.2). Using this information, evaluate the Helmholtz free energy of the system and derive other thermodynamic properties such as the pressure P and the (thermal) energy density U/V. Compare your results with the ones derived in Sec. 7.2 from the grand potential of the system.

7.16. Show that the mean energy per photon in a black-body ra on cavity is very nearly $2.7kT$.

7.17. Considering the volume dependence of the frequencies the vibrational modes of the radiation field, establish relation (7.2.17) between the pressure P and the energy density U/V.

7.18. The sun may be regarded as a black body at a temperature of 5800 K. Its diameter is about 1.4×10^9 m while its distance from the earth is about 1.5×10^{11} m.

(i) Calculate the total radiant intensity (in W/m^2) of sunlight at the surface of the earth.

(ii) What pressure would it exert on a perfectly absorbing surface placed normal to the rays (of the sun)?

(iii) If a flat surface on a satellite, which faces the sun, were an ideal absorber and emitter, what equilibrium temperature would it ultimately attain?

7.19. Figure 7.14 is a plot of $C_V(T)$ against T for a solid, the limiting value $C_V(\infty)$ being the classical result $3Nk$. Show that the shaded area in the figure, namely

$$\int_0^\infty \{C_V(\infty) - C_V(T)\}\, dT,$$

is exactly equal to the zero-point energy of the solid. Interpret the result physically.

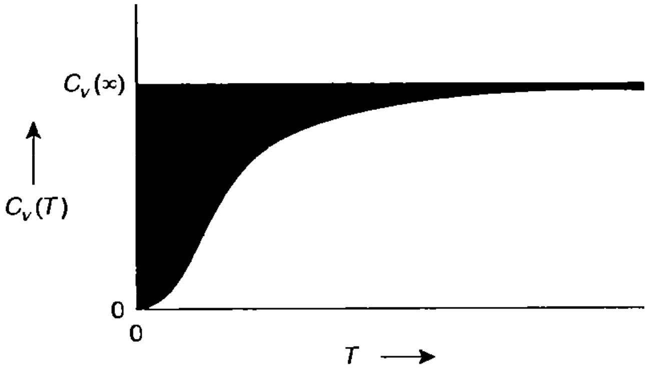


FIG. 7.14

7.20. Show that the total zero-point energy of a Debye solid is equal to $\frac{9}{8}Nk\Theta_D$.

[Note that this implies, for each vibrational mode of the solid, a mean zero-point energy $\frac{3}{8}k\Theta_D$, i.e. $\bar{\omega} = \frac{3}{4}\omega_D$.]

7.21. Show that, for $T \ll \Theta_D$, the quantity $(C_P - C_V)$ of a Debye solid varies as T^7 and hence the ratio $(C_P/C_V) \simeq 1$.

7.22. Determine the temperature T, in terms of the Debye temperature Θ_D, at which one-half of the oscillators in a Debye solid are expected to be in the excited states.

7.23. Determine the value of the parameter Θ_D for liquid He^4 from the empirical result (7.3.28).

7.24. (a) Compare the "mean thermal wavelength" λ_T of neutrons at a typical room temperature with the "minimum wavelength" λ_{min} of phonons in a typical crystal.

(b) Show that the frequency ω_D for a sodium chloride crystal is of the same order of magnitude as the frequency of an electromagnetic wave in the infrared.

7.25. Proceeding under the conditions (7.3.16) rather than (7.3.13), show that

$$C_V(T) = Nk\left\{D(x_{0,L}) + 2D(x_{0,T})\right\},$$

where $x_{0,L} = (\hbar\omega_{D,L}/kT)$ and $x_{0,T} = (\hbar\omega_{D,T}/kT)$. Compare this result with eqn. (7.3.17), and estimate the nature and the magnitude of the error involved in the latter.

7.26. A mechanical system consisting of n identical masses (each of mass m) connected in a straight line by identical springs (of stiffness K) has natural vibrational frequencies given by

$$\omega_r = 2\sqrt{\left(\frac{K}{m}\right)}\sin\left(\frac{r}{n}\cdot\frac{\pi}{2}\right); \quad r = 1, 2, \ldots, (n-1).$$

Correspondingly, a *linear* molecule composed of n identical atoms may be expected to have a vibrational spectrum given

$$\nu_r = \nu_c \sin\left(\frac{r}{n}\cdot\frac{\pi}{2}\right); \quad r = 1, 2, \ldots, (n-1)$$

where ν_c is a characteristic vibrational frequency of the molecu [illegible] how that this leads to a vibrational specific heat per molecule that varies as T^1 at low temperatures a [illegible] nds to the limiting value $(n-1)k$ at high temperatures.

7.27. Assuming the dispersion relation $\omega = Ak^s$, where ω is the angular frequency and k the wave number of a vibrational mode existing in a solid, show that the respective contribution towards the specific heat of the solid *at low temperatures* is proportional to $T^{3/s}$.

[Note that while $s = 1$ corresponds to the case of elastic waves in a lattice, $s = 2$ applies to spin waves propagating in a ferromagnetic system.]

7.28. Assuming the excitations to be phonons ($\omega = Ak$), show that their contribution towards the specific heat of an n-dimensional Debye system is proportional to T^n.

[Note that the elements selenium and tellurium form crystals in which atomic chains are arranged in parallel, so that in a certain sense they behave as one-dimensional; accordingly, over a certain range of temperatures, the T^1-law holds. For a similar reason, graphite obeys a T^2-law over a certain range of temperatures.]

7.29. The (minimum) potential energy of a solid, when all its atoms are "at rest" at their equilibrium positions, may be denoted by the symbol $\Phi_0(V)$, where V is the volume of the solid. Similarly, the normal frequencies of vibration, $\omega_i (i = 1, 2, \ldots, 3N-6)$, may be denoted by the symbols $\omega_i(V)$. Show that the pressure of the solid is given by

$$P = -\frac{\partial \Phi_0}{\partial V} + \gamma \frac{U}{V},$$

where U is the internal energy of the solid arising from the vibrations of the atoms, while γ is the *Grüneisen constant*:

$$\gamma = -\frac{\partial \ln \omega}{\partial \ln V} \approx \frac{1}{3}.$$

Assuming that, for $V \simeq V_0$,

$$\Phi_0(V) = \frac{(V - V_0)^2}{2\kappa_0 V_0},$$

where κ_0 and V_0 are constants and $\kappa_0 C_V T \ll V_0$, show that the coefficient of thermal expansion (at constant pressure $P \simeq 0$) is given by

$$\alpha \equiv \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_{N,P} = \frac{\gamma \kappa_0 C_V}{V_0}$$

Also show that

$$C_P - C_V = \frac{\gamma^2 \kappa_0 C_V^2 T}{V_0}.$$

7.30. Apply the general formula (6.4.3) for the kinetic pressure of a gas, namely

$$P = \tfrac{1}{3} n \langle pu \rangle,$$

to a gas of rotons and verify that the result so obtained agrees with the Boltzmannian relationship $P = nkT$.

7.31. Show that the free energy A and the inertial density ρ of a roton gas in *mass motion* are given by

$$A(v) = A(0)\frac{\sinh x}{x}$$

and

$$\rho(v) = \rho(0)\frac{3(x\cosh x - \sinh x)}{x^3},$$

where $x = vp_0/kT$.

7.32. Integrating (7.5.17) by parts, show that the effective mass of an excitation, whose energy–momentum relationship is denoted by $\varepsilon(p)$, is given by

$$m_{\text{eff}} = \left\langle \frac{1}{3p^2}\left\{\frac{d}{dp}\left(p^4 \frac{dp}{d\varepsilon}\right)\right\}\right\rangle.$$

Check the validity of this result by considering the examples of (i) an ideal-gas particle, (ii) a phonon and (iii) a roton.

Notes

[1] Actually it is the ratio $n/T^{3/2}$, rather than the quantities n and T separately, that determines the degree of degeneracy in a given system. For instance, white dwarf stars, even at temperatures of order 10^7 K, constitute statistically degenerate systems; see Sec. 8.4.

[2] The theory of this section is restricted to a system of *nonrelativistic* particles. For the general case, see Kothari and Singh (1941) and Landsberg and Dunning-Davis (1965).

[3] Remember that the largest value z can have *in principle* is unity. In fact, as $T \to 0$, $z = N_0/(N_0 + 1) \to N/(N+1)$, which is very nearly unity (but certainly on the right side of it).

[4] Of course, the repercussions of this phenomenon in the coordinate space are no less curious. It prepares the stage for the onset of *superfluidity*, a quantum manifestation discussed in Sec. 7.5.

[5] For a rigorous discussion of the onset of Bose–Einstein condensation, see Landsberg (1954b), where an attempt has also been made to coordinate much of the previously published work on this topic. For a more recent study, see Greenspoon and Pathria (1974).

[6] An equivalent relationship is: $g_{3/2}(z)/g_{3/2}(1) = (T_c/T)^{3/2} < 1$.

[7] Equation (6.2.12) gives, for an ideal classical gas, $\ln \mathcal{Q} = zV/\lambda^3$. Accordingly, $N \equiv z(\partial \ln \mathcal{Q}/\partial z) = z(V/\lambda^3)$, with the result that $z = (\lambda^3/v)$.

[8] Actually, for all $T \leq T_c$, we can write

$$P(T) = P(T_c) \cdot (T/T_c)^{5/2} \simeq 0.5134(N_e kT/V).$$

We infer that, while particles in the condensed phase do not exert any pressure at all, particles in the excited states are about half as effective as in the Boltzmannian case.

[9] As is well known, the longitudinal modes do not appear in the case of radiation.

[10] The factor 2 in this expression arises essentially from the same cause as in the Rayleigh expression (2). However, in the present context, it would be more appropriate to regard it as representing the two states of polarization of the photon spin.

[11] Compared with the standard Bose–Einstein result, say eqn. (7.1.2), formula (7) suggests that we are dealing here with a case for which z is precisely equal to unity. It is not difficult to see that this is due to the fact that the total number of particles in the present case is *indefinite*. For then, their equilibrium number $\overline{N}$ has to be determined by the condition that the free energy of the system is a minimum, i.e. $\{(\partial A/\partial N)_{N=\overline{N}}\}_{V,T} = 0$, which, by definition, implies that $\mu = 0$ and hence $z = 1$.

[12] The Rayleigh–Jeans formula follows directly if we use for $\langle \varepsilon_s \rangle$ the equipartition value kT rather than the quantum-theoretical value (1).

[13] Here, use has been made of the fact that the value of the definite integral is $6\zeta(4) = \pi^4/15$; see Appendix D.

[14] Of course, the number of phonons themselves is indefinite. As a result, the chemical potential of the phonon gas is also zero.

[15] The thermal energy of the solid may as well be written as $\sum_i \langle n_i \rangle \hbar\omega_i$, where $\langle n_i \rangle \{= (e^{\hbar\omega_i/kT} - 1)^{-1}\}$ is the *mean occupation number* of the phonon level ε_i. Clearly, the phonons, like photons, obey Bose–Einstein statistics, with $\mu = 0$.

[16] Actually, when the temperature is high enough, so that *all* $(\hbar\omega_i/kT) \ll 1$, the general formula (8) itself reduces to the classical one. This corresponds to the situation when each of the $3N$ modes of vibration possesses a thermal energy kT.

[17] It can be shown that, according to this theory, deviations from the T^3-law should not exceed 2 per cent so long as $T < \Theta_D/10$. However, in the case of *metals*, one cannot expect to reach a true T^3-region because, well before that, the specific heat of the electron gas might become a dominant contribution (see Sec. 8.3); unless the two contributions are separated out, one is likely to obtain a somewhat suppressed value of Θ_D from these observations.

[18] See, for example, Fowler and Guggenheim (1960), p. 144.

[19] The term "roton" for these excitations was coined by Landau who had originally thought that these excitations might, in some way, represent local disturbances of a *rotational* character in the liquid. However, subsequent theoretical work, especially that of Feynman (1953, 1954) and of Brueckner and Sawada (1957), did not support this contention. Nevertheless, the name "roton" has stayed on.

[20] This seems to confirm a prediction by Pitaevskii (1959) that an end point in the spectrum might occur at a "critical" value p_c of the excitation momentum, where ε_c is equal to 2Δ and $(\partial\varepsilon/\partial p)_c$ is zero.

[21] Looking back at integral (9), what we have done here amounts to replacing p^2 in the integrand by its mean value p_0^2 and then carrying out integration over the "complete" range of the variable $(p - p_0)$.

[22] This result is highly suggestive of the fact that for rotons there is only *one* true degree of freedom, namely the *magnitude* of the roton momentum, that is thermally effective!

[23] The drift velocity $\boldsymbol{v}$ must satisfy the condition $(\boldsymbol{v} \cdot \boldsymbol{p}) \leq \varepsilon$, for otherwise some of the occupation numbers will become negative! This leads to the existence of a *critical velocity* v_c for these excitations, such that for v exceeding v_c the formalism developed here breaks down. It is not difficult to see that this (critical) velocity is given by the relation: $v_c = (\varepsilon/p)_{\mathrm{min}}$, as in eqn. (24).

CHAPTER 8

IDEAL FERMI SYSTEMS

8.1. Thermodynamic behavior of an ideal Fermi gas

According to Secs 6.1 and 6.2, we obtain for an ideal Fermi gas

$$\frac{PV}{kT} \equiv \ln \mathcal{Q} = \sum_{\varepsilon} \ln(1 + ze^{-\beta\varepsilon}) \tag{1}$$

and

$$N \equiv \sum_{\varepsilon} \langle n_{\varepsilon} \rangle = \sum_{\varepsilon} \frac{1}{z^{-1}e^{\beta\varepsilon} + 1}, \tag{2}$$

where $\beta = 1/kT$ and $z = \exp(\mu/kT)$. Unlike the Bose case, the parameter z in the Fermi case can take on *unrestricted* values: $0 \leq z < \infty$. Moreover, in view of the Pauli exclusion principle, the question of a large number of particles occupying a single energy state does not even arise in this case; hence, there is no phenomenon like Bose–Einstein condensation here. Nevertheless, at sufficiently low temperatures, Fermi gas displays its own brand of quantal behavior, a detailed study of which is of great physical interest.

If we replace summations over ε by corresponding integrations, eqns (1) and (2) become

$$\frac{P}{kT} = \frac{g}{\lambda^3} f_{5/2}(z) \tag{3}$$

and

$$\frac{N}{V} = \frac{g}{\lambda^3} f_{3/2}(z), \tag{4}$$

where g is a weight factor arising from the "internal structure" of the particles (such as *spin*), λ is the mean thermal wavelength of the particles

$$\lambda = h/(2\pi mkT)^{1/2}, \tag{5}$$

while $f_{\nu}(z)$ are Fermi–Dirac functions defined by, see Appendix E,

$$f_{\nu}(z) = \frac{1}{\Gamma(\nu)} \int_0^{\infty} \frac{x^{\nu-1}\,dx}{z^{-1}e^{x} + 1} = z - \frac{z^2}{2^{\nu}} + \frac{z^3}{3^{\nu}} - \cdots. \tag{6}$$

Eliminating z between eqns (3) and (4), we obtain the *equation of state* of the Fermi gas.

The internal energy U of the Fermi gas is given by

$$U \equiv -\left(\frac{\partial}{\partial\beta}\ln\mathcal{Q}\right)_{z,V} = kT^2\left(\frac{\partial}{\partial T}\ln\mathcal{Q}\right)_{z,V}$$

$$= \frac{3}{2}kT\frac{gV}{\lambda^3}f_{5/2}(z) = \frac{3}{2}NkT\frac{f_{5/2}(z)}{f_{3/2}(z)}; \tag{7}$$

thus, quite generally, this system satisfies the relationship

$$P = \tfrac{2}{3}(U/V). \tag{8}$$

The specific heat C_V of the gas can be obtained by differentiating (7) with respect to T, keeping N and V constant, and making use of the relationship

$$\frac{1}{z}\left(\frac{\partial z}{\partial T}\right)_V = -\frac{3}{2T}\frac{f_{3/2}(z)}{f_{1/2}(z)}, \tag{9}$$

which follows from eqn. (4) and the recurrence formula (E.4). The final result is

$$\frac{C_V}{Nk} = \frac{15}{4}\frac{f_{5/2}(z)}{f_{3/2}(z)} - \frac{9}{4}\frac{f_{3/2}(z)}{f_{1/2}(z)}. \tag{10}$$

For the Helmholtz free energy of the gas, we get

$$A \equiv N\mu - PV = NkT\left\{\ln z - \frac{f_{5/2}(z)}{f_{3/2}(z)}\right\}, \tag{11}$$

and for the entropy

$$S \equiv \frac{U-A}{T} = Nk\left\{\frac{5}{2}\frac{f_{5/2}(z)}{f_{3/2}(z)} - \ln z\right\}. \tag{12}$$

In order to determine the various properties of the Fermi gas in terms of the particle density $n(=N/V)$ and the temperature T, we need to know the functional dependence of the parameter z on n and T; this information is formally contained in the implicit relationship (4). For detailed studies, one is sometimes obliged to make use of numerical tables of the functions $f_\nu(z)$ (see Appendix E); for physical understanding, however, the various limiting forms of these functions serve the purpose well.

Now, if the density of the gas is very low and/or its temperature very high, then the situation might correspond to

$$f_{3/2}(z) = \frac{nh^3}{g(2\pi mkT)^{3/2}} \ll 1; \tag{13}$$

we then speak of the gas as being highly *nondegenerate*. In view of the expansion (6), this implies that z itself is much smaller than unity and, in consequence, $f_\nu(z) \simeq z$. Expressions for the various thermodynamic properties of the gas then become

$$P = NkT/V; \quad U = \tfrac{3}{2}NkT; \quad C_V = \tfrac{3}{2}Nk, \tag{14}$$

$$A = NkT\left\{\ln\left(\frac{n\lambda^3}{g}\right) - 1\right\} \tag{15}$$

and

$$S = Nk\left\{\frac{5}{2} - \ln\left(\frac{n\lambda^3}{g}\right)\right\}. \tag{16}$$

If the parameter z is small in comparison with unity but not very small, then we should make a fuller use of the series (6) in order to eliminate z between eqns (3) and (4). The procedure is just the same as in the corresponding Bose case, i.e. we first invert the series appearing in (4) to obtain an expansion for z in powers of $(n\lambda^3/g)$ and then substitute this expansion into the series appearing in (3). The equation of state then takes the form of the *virial expansion*

$$\frac{PV}{NkT} = \sum_{l=1}^{\infty}(-1)^{l-1}a_l\left(\frac{\lambda^3}{gv}\right)^{l-1}, \tag{17}$$

where $v = 1/n$, while the coefficients a_l are the same as quoted in (7.1.14). For the specific heat, in particular, we obtain

$$\begin{aligned} C_V &= \frac{3}{2}Nk\sum_{l=1}^{\infty}(-1)^{l-1}\frac{5-3l}{2}a_l\left(\frac{\lambda^3}{gv}\right)^{l-1} \\ &= \frac{3}{2}Nk\left[1 - 0.0884\left(\frac{\lambda^3}{gv}\right) + 0.0066\left(\frac{\lambda^3}{gv}\right)^2 - 0.0004\left(\frac{\lambda^3}{gv}\right)^3 + \cdots\right]. \end{aligned} \tag{18}$$

Thus, at finite temperatures, the specific heat of the gas is smaller than its limiting value $\frac{3}{2}Nk$. As will be seen in the sequel, the specific heat of the ideal Fermi gas decreases *monotonically* as the temperature of the gas falls; see Fig. 8.2 and compare it with the corresponding Fig. 7.4 for the ideal Bose gas.

If the density n and the temperature T are such that the parameter $(n\lambda^3/g)$ is of order unity, the foregoing expansions cannot be of much use. In that case, one may have to make recourse to numerical calculation. However, if $(n\lambda^3/g) \gg 1$, the functions involved can be expressed as *asymptotic* expansions in powers of $(\ln z)^{-1}$; we then speak of the gas as being degenerate. As $(n\lambda^3/g) \to \infty$, our functions assume a closed form, with the result that the expressions for the various thermodynamic quantities pertaining to the system become highly simplified; we then speak of the gas as being *completely degenerate*. For simplicity, we shall first discuss the main features of the system in a state of complete degeneracy.

In the limit $T \to 0$, which implies $(n\lambda^3/g) \to \infty$, the mean occupation numbers of the single-particle state $\varepsilon(p)$ become

$$\langle n_\varepsilon\rangle \equiv \frac{1}{e^{(\varepsilon-\mu)/kT}+1} = \begin{cases} 1 & \text{for} \quad \varepsilon < \mu_0 \\ 0 & \text{for} \quad \varepsilon > \mu_0, \end{cases} \tag{19}$$

where μ_0 is the chemical potential of the system at $T = 0$. The function $\langle n_\varepsilon\rangle$ is thus a *step function* which stays constant at the (highest) value 1 right from $\varepsilon = 0$ to $\varepsilon = \mu_0$ and then suddenly drops to the (lowest) value 0; see the dotted line in

Fig. 8.1. Thus, at $T = 0$, all single-particle states up to $\varepsilon = \mu_0$ are "completely" filled, with one particle per state (in accordance with the Pauli principle), while all single-particle states with $\varepsilon > \mu_0$ are empty. The limiting energy μ_0 is generally referred to as the *Fermi energy* of the system and is denoted by the symbol ε_F; the corresponding value of the single-particle momentum is referred to as the *Fermi momentum* and is denoted by the symbol p_F. The defining equation for these parameters is

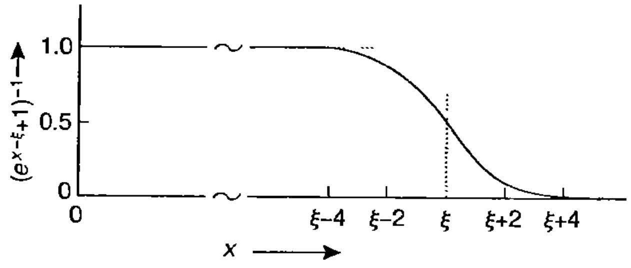


FIG. 8.1. Fermi distribution at low temperatures, with $x = \varepsilon/kT$ and $\xi = \mu/kT$. The rectangle denotes the limiting distribution as $T \to 0$, in that case, the Fermi function is unity for $\varepsilon < \mu_0$ and zero for $\varepsilon > \mu_0$.

$$\int_0^{\varepsilon_F} a(\varepsilon)\, d\varepsilon = N, \tag{20}$$

where $a(\varepsilon)$ denotes the *density of states* of the system and is given by the general expression

$$a(\varepsilon) = \frac{gV}{h^3} 4\pi p^2 \frac{dp}{d\varepsilon}. \tag{21}$$

We readily obtain

$$N = \frac{4\pi g V}{3h^3} p_F^3, \tag{22}$$

whence

$$p_F = \left(\frac{3N}{4\pi g V}\right)^{1/3} h; \tag{23}$$

accordingly, in the *nonrelativistic* case,

$$\varepsilon_F = \left(\frac{3N}{4\pi g V}\right)^{2/3} \frac{h^2}{2m} = \left(\frac{6\pi^2 n}{g}\right)^{2/3} \frac{\hbar^2}{2m}. \tag{24}$$

The ground-state, or zero-point, energy of the system is then given by

$$\begin{aligned} E_0 &= \frac{4\pi g V}{h^3} \int_0^{p_F} \left(\frac{p^2}{2m}\right) p^2\, dp \\ &= \frac{2\pi g V}{5mh^3} p_F^5, \end{aligned} \tag{25}$$

whence

$$\frac{E_0}{N} = \frac{3p_F^2}{10m} = \frac{3}{5}\varepsilon_F. \tag{26}$$

The ground-state pressure of the system is in turn given by

$$P_0 = \tfrac{2}{3}(E_0/V) = \tfrac{2}{5}n\varepsilon_F. \tag{27}$$

Substituting for ε_F, the foregoing expression takes the form

$$P_0 = \left(\frac{6\pi^2}{g}\right)^{2/3} \frac{\hbar^2}{5m} n^{5/3} \propto n^{5/3}. \tag{28}$$

The zero-point motion seen here is clearly a quantum effect arising because of the Pauli principle, according to which, even at $T = 0$ K, the particles constituting the system cannot settle down into a *single* energy state (as we had in the Bose case) and are therefore spread over a requisite number of lowest available energy states. As a result, the Fermi system, even at absolute zero, is quite live!

For a discussion of properties such as the specific heat and the entropy of the system, we must extend our study to *finite* temperatures. If we decide to restrict ourselves to low temperatures, then deviations from the ground-state results will not be too large; accordingly, an analysis based on the asymptotic expansions of the functions $f_\nu(z)$ would be quite appropriate. However, before we do that it seems useful to carry out a physical assessment of the situation with the help of the expression

$$\langle n_\varepsilon \rangle = \frac{1}{e^{(\varepsilon-\mu)/kT} + 1}. \tag{29}$$

The situation corresponding to $T = 0$ is summarized in eqn. (19) and is shown as a step function in Fig. 8.1. Deviations from this, when T is finite (but still much smaller than the *characteristic* temperature μ_0/k), will be significant *only* for those values of ε for which the magnitude of the quantity $(\varepsilon - \mu)/kT$ is of order unity (for otherwise the exponential term in (29) will not be much different from its ground-state value, viz. $e^{\pm\infty}$); see the solid curve in Fig. 8.1. We therefore conclude that the thermal excitation of the particles occurs only in a narrow energy range which is located around the energy value $\varepsilon = \mu_0$ and has a width $O(kT)$. The fraction of the particles that are thermally excited is, therefore, $O(kT/\varepsilon_F)$–the bulk of the system remaining uninfluenced by the rise in temperature.[1] This is the most characteristic feature of a degenerate Fermi system and is essentially responsible for both qualitative and quantitative differences between the physical behavior of this system and that of a corresponding classical system.

To conclude the argument, we observe that since the thermal energy per "excited" particle is $O(kT)$, the thermal energy of the whole system will be $O(Nk^2T^2/\varepsilon_F)$; accordingly, the specific heat of the system will be $O(Nk \cdot kT/\varepsilon_F)$. Thus, the low-temperature specific heat of a Fermi system differs from the classical value $\frac{3}{2}Nk$ by a factor which not only reduces it considerably in magnitude but also makes it temperature-dependent (varying as T^1). It will be seen repeatedly that the first-power dependence of C_V on T is a typical feature of Fermi systems at low temperatures.

For an analytical study of the Fermi gas at finite, but low, temperatures, we observe that the value of z, which was infinitely large at absolute zero, is now finite, though still large in comparison with unity. The functions $f_\nu(z)$ can, therefore, be expressed as asymptotic expansions in powers of $(\ln z)^{-1}$; see Sommerfeld's lemma (E. 15). For the values of ν we are presently interested in, namely $\frac{5}{2}$, $\frac{3}{2}$ and $\frac{1}{2}$, we have *to the first approximation*

$$f_{5/2}(z) = \frac{8}{15\pi^{1/2}}(\ln z)^{5/2}\left[1 + \frac{5\pi^2}{8}(\ln z)^{-2} + \cdots\right] \tag{30}$$

$$f_{3/2}(z) = \frac{4}{3\pi^{1/2}}(\ln z)^{3/2}\left[1 + \frac{\pi^2}{8}(\ln z)^{-2} + \cdots\right] \tag{31}$$

and

$$f_{1/2}(z) = \frac{2}{\pi^{1/2}}(\ln z)^{1/2}\left[1 - \frac{\pi^2}{24}(\ln z)^{-2} + \cdots\right]. \tag{32}$$

Substituting (31) into (4), we obtain

$$\frac{N}{V} = \frac{4\pi g}{3}\left(\frac{2m}{h^2}\right)^{3/2}(kT\ln z)^{3/2}\left[1 + \frac{\pi^2}{8}(\ln z)^{-2} + \cdots\right]. \tag{33}$$

In the zeroth approximation, this gives

$$kT\ln z \equiv \mu \simeq \left(\frac{3N}{4\pi g V}\right)^{2/3}\frac{h^2}{2m} \tag{34}$$

which is identical with the ground-state result $\mu_0 = \varepsilon_F$; see (24). In the next approximation, we obtain

$$kT\ln z \equiv \mu \simeq \varepsilon_F\left[1 - \frac{\pi^2}{12}\left(\frac{kT}{\varepsilon_F}\right)^2\right]. \tag{35}$$

Substituting (30) and (31) into (7), we obtain

$$\frac{U}{N} = \frac{3}{5}(kT\ln z)\left[1 + \frac{\pi^2}{2}(\ln z)^{-2} + \cdots\right]; \tag{36}$$

with the help of (35), this becomes

$$\frac{U}{N} = \frac{3}{5}\varepsilon_F\left[1 + \frac{5\pi^2}{12}\left(\frac{kT}{\varepsilon_F}\right)^2 + \cdots\right]. \tag{37}$$

The pressure of the gas is then given by

$$P = \frac{2}{3}\frac{U}{V} = \frac{2}{5}n\varepsilon_F\left[1 + \frac{5\pi^2}{12}\left(\frac{kT}{\varepsilon_F}\right)^2 + \cdots\right]. \tag{38}$$

As expected, the main terms of eqns (37) and (38) are identical with the ground-state results (26) and (27). From the temperature-dependent part of (37), we obtain for the low-temperature specific heat of the gas

$$\frac{C_V}{Nk} = \frac{\pi^2}{2}\frac{kT}{\varepsilon_F} + \cdots \tag{39}$$

Thus, for $T \ll T_F$, where $T_F (= \varepsilon_F/k)$ is the *Fermi temperature* of the system, the specific heat varies as the first power of temperature; moreover, in magnitude, it

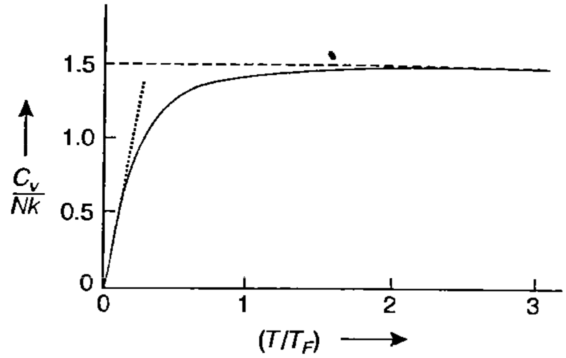


FIG. 8.2. The specific heat of an ideal Fermi gas; the dotted line depicts the *linear* behavior at low temperatures.

is considerably smaller than the classical value $\frac{3}{2}Nk$. The overall variation of C_V with T is shown in Fig. 8.2.

The Helmholtz free energy of the system follows directly from eqns (35) and (38):

$$\frac{A}{N} = \mu - \frac{PV}{N}$$
$$= \frac{3}{5}\varepsilon_F\left[1 - \frac{5\pi^2}{12}\left(\frac{kT}{\varepsilon_F}\right)^2 + \cdots\right], \tag{40}$$

whence

$$\frac{S}{Nk} = \frac{\pi^2}{2}\frac{kT}{\varepsilon_F} + \cdots. \tag{41}$$

Thus, as $T \to 0$, $S \to 0$, in accordance with the third law of thermodynamics.

8.2. Magnetic behavior of an ideal Fermi gas

We now turn our attention to studying the equilibrium state of a gas of non-interacting fermions in the presence of an external magnetic field B. The main problem here is to determine the net magnetic moment M acquired by the gas (as a function of B and T) and then calculate the susceptibility $\chi(T)$. The answer naturally depends upon the intrinsic magnetic moment μ^* of the particles and the corresponding multiplicity factor $(2J + 1)$; see, for instance, the treatment given in Sec. 3.9. According to the Boltzmannian treatment, one obtains a (positive) susceptibility $\chi(T)$ which, at high temperatures, obeys the *Curie law*: $\chi \propto T^{-1}$; at low temperatures, one obtains a state of magnetic saturation. However, if we treat the problem on the basis of Fermi statistics we obtain significantly different results, especially at low temperatures. In particular, since the Fermi gas is pretty live even at absolute zero, no magnetic saturation ever results; we rather obtain a limiting susceptibility χ_0 which is independent of temperature but is dependent on the density of the gas. Studies along these lines were first made by Pauli, in 1927, when he suggested that the conduction electrons in alkali metals be regarded as a "highly degenerate Fermi gas"; these studies enabled him to explain the physics behind the *feeble* and *temperature-independent* character of the paramagnetism of metals. Accordingly, this phenomenon is referred to as *Pauli paramagnetism*— contrast to the classical Langevin paramagnetism.

In quantum statistics, we encounter yet another effect which is totally absent classical statistics. This is diamagnetic in character and arises from the quantization of the orbits of *charged* particles in the presence of an external magnetic field or one may say, from the quantization of the (kinetic) energy of charged particles associated with their motion perpendicular to the direction of the field. The existence of this effect was first established by Landau (1930); so, we refer to it as *Landau diamagnetism.* This leads to an additional susceptibility $\chi(T)$ which, though negative in sign, is somewhat similar to the paramagnetic susceptibility in that it obeys Curie's law at high temperatures and tends to a temperature-independent but density-dependent limiting value as $T \to 0$. In general, the magnetic behavior of a Fermi gas is determined jointly by the intrinsic magnetic moment of the particles and the quantization of their orbits. If the spin–orbit interaction is negligible, the resultant behavior is obtained by a simple addition of the two effects.

A. Pauli paramagnetism

The energy of a particle, in the presence of an external magnetic field $\boldsymbol{B}$, is given by

$$\varepsilon = \frac{p^2}{2m} - \boldsymbol{\mu}^* \cdot \mathbf{B}, \tag{1}$$

where $\boldsymbol{\mu}^*$ is the intrinsic magnetic moment of the particle and m its mass. For simplicity, we assume that the particle spin is $\frac{1}{2}$; the vector $\boldsymbol{\mu}^*$ will then be either parallel to the vector $\boldsymbol{B}$ or antiparallel. We thus have two groups of particles in the gas:

(i) those having $\boldsymbol{\mu}^*$ parallel to $\boldsymbol{B}$, with $\varepsilon = p^2/2m - \mu^* B$, and
(ii) those having $\boldsymbol{\mu}^*$ antiparallel to $\boldsymbol{B}$, with $\varepsilon = p^2/2m + \mu^* B$.

At absolute zero, all energy levels up to the Fermi level ε_F will be filled, while all levels beyond ε_F will be empty. Accordingly, the kinetic energy of the particles in the first group will range between 0 and $(\varepsilon_F + \mu^* B)$, while the kinetic energy of the particles in the second group will range between 0 and $(\varepsilon_F - \mu^* B)$. The respective numbers of occupied energy levels (and hence of particles) in the two groups will, therefore, be

$$N^+ = \frac{4\pi V}{3h^3}\{2m(\varepsilon_F + \mu^* B)\}^{3/2} \tag{2}$$

and

$$N^- = \frac{4\pi V}{3h^3}\{2m(\varepsilon_F - \mu^* B)\}^{3/2}. \tag{3}$$

The net magnetic moment acquired by the gas is then given by

$$M = \mu^*(N^+ - N^-) = \frac{4\pi\mu^* V(2m)^{3/2}}{3h^3}\{(\varepsilon_F + \mu^* B)^{3/2} - (\varepsilon_F - \mu^* B)^{3/2}\}. \tag{4}$$

We thus obtain for the low-field susceptibility (per unit volume) of the gas

$$\chi_0 = \operatorname*{Lim}_{B\to 0}\left(\frac{M}{VB}\right) = \frac{4\pi\mu^{*2}(2m)^{3/2}\varepsilon_F^{1/2}}{h^3} \tag{5}$$

Making use of formula (8.1.24), with $g = 2$, the foregoing result may be written as

$$\chi_0 = \tfrac{3}{2}n\mu^{*2}/\varepsilon_F. \tag{6}$$

For comparison, the corresponding high-temperature result is given by eqn. (3.9.26), with $g = 2$ and $J = \frac{1}{2}$:

$$\chi_\infty = n\mu^{*2}/kT. \tag{7}$$

We note that $\chi_0/\chi_\infty = O(kT/\varepsilon_F)$.

To obtain an expression for χ that holds at all T, we proceed as follows. Denoting the number of particles with momentum $\boldsymbol{p}$ and magnetic moment parallel (or antiparallel) to the field by the symbol n_p^+ (or n_p^-), the total energy of the gas can be written as

$$\begin{aligned} E_n &= \sum_p \left[\left(\frac{p^2}{2m} - \mu^* B\right) n_p^+ + \left(\frac{p^2}{2m} + \mu^* B\right) n_p^-\right] \\ &= \sum_p (n_p^+ + n_p^-)\frac{p^2}{2m} - \mu^* B(N^+ - N^-), \end{aligned} \tag{8}$$

where N^+ and N^- denote the total number of particles in the two groups, respectively. The partition function of the system is then given by

$$Q(N) = \sum_{\{n_p^+\},\{n_p^-\}}{}' \exp(-\beta E_n), \tag{9}$$

where the primed summation is subject to the conditions

$$n_p^+, n_p^- = 0 \text{ or } 1, \tag{10}$$

and

$$\sum_p n_p^+ + \sum_p n_p^- = N^+ + N^- = N. \tag{11}$$

To evaluate the sum in (9), we first fix an arbitrary value of the number N^+ (which automatically fixes the value of N^- as well) and sum over *all* n_p^+ and n_p^- that conform to the fixed values of the numbers N^+ and N^- as well as to condition (10). Next, we sum over all possible values of N^+, viz. from $N^+ = 0$ to $N^+ = N$. We thus have

$$Q(N) = \sum_{N\;=0}^{N}\left[e^{\beta\mu^* B(2N^+ - N)}\left\{\sum_{\{n_p^+\}}{}'' \exp\left(-\beta\sum_p \frac{p^2}{2m}n_p^+\right)\sum_{\{n_p^-\}}{}''' \exp\left(-\beta\sum_p \frac{p^2}{2m}n_p^-\right)\right\}\right]; \tag{12}$$

here, the summation $\sum''$ is subject to the restriction $\sum_p n_p^+ = N^+$, while $\sum'''$ is subject to the restriction $\sum_p n_p^- = N - N^+$.

Now, let $Q_0(\mathcal{N})$ denote the partition function of an ideal Fermi gas of $\mathcal{N}$ "spinless" particles of mass m; then, obviously,

$$Q_0(\mathcal{N}) = \sum_{\{n_p\}}{}' \exp\left(-\beta \sum_p \frac{p^2}{2m} n_p\right) \equiv \exp\{-\beta A_0(\mathcal{N})\}, \tag{13}$$

where $A_0(\mathcal{N})$ is the free energy of this fictitious system. Equation (12) can then be written as

$$Q(N) = e^{-\beta\mu^* BN} \sum_{N^+=0}^{N} [e^{2\beta\mu^* BN^+} Q_0(N^+) Q_0(N - N^+)], \tag{14}$$

whence

$$\frac{1}{N} \ln Q(N) = -\beta\mu^* B + \frac{1}{N} \ln \sum_{N^+=0}^{N} [\exp\{2\beta\mu^* BN^+ - \beta A_0(N^+) - \beta A_0(N - N^+)\}] \tag{15}$$

As before, the logarithm of the sum $\sum_{N^+}$ may be replaced by the logarithm of the largest term in the sum; the error committed in doing so would be negligible in comparison with the term retained. Now, the value $\overline{N^+}$, of N^+, which corresponds to the largest term in the sum, can be ascertained by setting the differential coefficient of the general term, with respect to N^+, equal to zero; this gives

$$2\mu^* B - \left[\frac{\partial A_0(N^+)}{\partial N^+}\right]_{N^+=\overline{N^+}} - \left[\frac{\partial A_0(N - N^+)}{\partial N^+}\right]_{N^+=\overline{N^+}} = 0.$$

that is

$$\mu_0(\overline{N^+}) - \mu_0(N - \overline{N^+}) = 2\mu^* B. \tag{16}$$

where $\mu_0(\mathcal{N})$ is the chemical potential of the fictitious system of $\mathcal{N}$ "spinless" fermions.

The foregoing equation contains the general solution being sought for. To obtain an explicit expression for χ, we introduce a dimensionless parameter r, defined by

$$M = \mu^*(\overline{N^+} - \overline{N^-}) = \mu^*(2\overline{N^+} - N) = \mu^* Nr \quad (0 \le r \le 1); \tag{17}$$

eqn. (16) then becomes

$$\mu_0\left(\frac{1+r}{2}N\right) - \mu_0\left(\frac{1-r}{2}N\right) = 2\mu^* B. \tag{18}$$

If $B = 0$, $r = 0$ (which corresponds to a *completely random* orientation of the elementary moments). For small B, r would also be small; so, we may carry out a Taylor expansion of the left-hand side of (18) about $r = 0$. Retaining only the first term of the expansion, we obtain

$$r \simeq \frac{2\mu^* B}{\left.\frac{\partial \mu_0(xN)}{\partial x}\right|_{x=1/2}}. \tag{19}$$

The low-field susceptibility (per unit volume) of the system is then given by

$$\chi = \frac{M}{VB} = \frac{\mu^* N r}{VB} = \frac{2n\mu^{*2}}{\left.\dfrac{\partial \mu_0(xN)}{\partial x}\right|_{x=1/2}}, \tag{20}$$

which is the desired result.

For $T \to 0$, the chemical potential of the fictitious system can be obtained from eqn. (8.1.34), with $g = 1$:

$$\mu_0(xN) = \left(\frac{3xN}{4\pi V}\right)^{2/3} \frac{h^2}{2m},$$

whence

$$\left.\frac{\partial \mu_0(xN)}{\partial x}\right|_{x=1/2} = \frac{2^{4/3}}{3}\left(\frac{3N}{4\pi V}\right)^{2/3} \frac{h^2}{2m}. \tag{21}$$

On the other hand, the Fermi energy of the actual system is given by eqn. (8.1.24), with $g = 2$:

$$\varepsilon_F = \left(\frac{3N}{8\pi V}\right)^{2/3} \frac{h^2}{2m}. \tag{22}$$

Making use of eqns (21) and (22), we obtain from (20)

$$\chi_0 = \frac{2n\mu^{*2}}{\frac{4}{3}\varepsilon_F} = \frac{3}{2} n\mu^{*2}/\varepsilon_F, \tag{23}$$

in complete agreement with our earlier result (6). For finite but low temperatures, one has to use eqn. (8.1.35) instead of (8.1.34). The final result turns out to be

$$\chi \simeq \chi_0 \left[1 - \frac{\pi^2}{12}\left(\frac{kT}{\varepsilon_F}\right)^2\right]. \tag{24}$$

On the other hand, for $T \to \infty$, the chemical potential of the fictitious system follows directly from eqn. (8.1.4), with $g = 1$ and $f_{3/2}(z) \simeq z$:

$$\mu_0(xN) = kT \ln (xN\lambda^3/V),$$

whence

$$\left.\frac{\partial \mu_0(xN)}{\partial x}\right|_{x=1/2} = 2kT. \tag{25}$$

Equation (20) then gives

$$\chi_\infty = n\mu^{*2}/kT \tag{26}$$

in complete agreement with our earlier result (7). For large but finite temperatures, one has to take $f_{3/2}(z) \simeq z - (z^2/2^{3/2})$. The final result then turns out to be

$$\chi \simeq \chi_\infty \left(1 - \frac{n\lambda^3}{2^{5/2}}\right); \tag{27}$$

the correction term here is proportional to $(T_F/T)^{3/2}$ and tends to zero as $T \to \infty$.

B. Landau diamagnetism

We now study the magnetism arising from the quantization of the orbital motion of (charged) particles in the presence of an external magnetic field. In a uniform field of intensity B, directed along the z-axis, a charged particle would follow a helical path whose axis is parallel to the z-axis and whose projection on the (x, y)-plane is a circle. Motion along the z-direction has a constant linear velocity u_z, while that in the (x, y)-plane has a constant angular velocity eB/mc; the latter arises from the Lorentz force, $e(u \times B)/c$, experienced by the particle. Quantum-mechanically, the energy associated with the circular motion is *quantized*, in units of $e\hbar B/mc$. The energy associated with the linear motion along the z-axis is also quantized but, in view of the smallness of the energy intervals, this may be taken as a continuous variable. We thus have for the total energy of the particle[2]

$$\varepsilon = \frac{e\hbar B}{mc}\left(j + \frac{1}{2}\right) + \frac{p_z^2}{2m} \quad (j = 0, 1, 2, \ldots). \tag{28}$$

Now, these quantized energy levels are degenerate because they result from a "coalescing together" of an almost continuous set of the zero-field levels. A little reflection shows that all those levels for which the value of the quantity $(p_x^2 + p_y^2)/2m$ lay between $e\hbar Bj/mc$ and $e\hbar B(j + 1)/mc$ now "coalesce together" into a single level characterized by the quantum number j. The number of these levels is given by

$$\begin{aligned}\frac{1}{h^2}\int dx\,dy\,dp_x\,dp_y &= \frac{L_xL_y}{h^2}\pi\left[2m\frac{e\hbar B}{mc}\{(j+1) - j\}\right]\\ &= L_xL_y\frac{eB}{hc},\end{aligned} \tag{29}$$

which is independent of j. The *multiplicity factor* (29) is a quantum-mechanical measure of the freedom available to the particle for the centre of its orbit to be "located" anywhere in the total area L_xL_y of the physical space. Figure 8.3 depicts

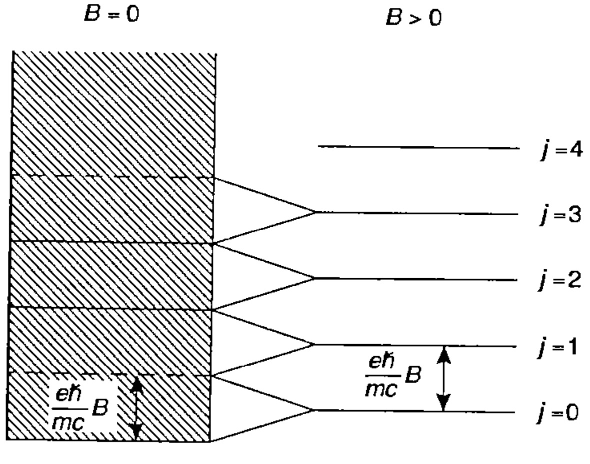


FIG. 8.3. The single-particle energy levels, for a two-dim ional motion, (i) in the absence of an external magnetic field ($B = 0$) and (ii) in resence of an external magnetic field ($B > 0$).

the manner in which the zero-field energy levels of the particle group themselves into a spectrum of oscillator-like levels on the application of the external magnetic field.

The grand partition function of the gas is given by the standard formula

$$\ln \mathcal{Q} = \sum_{\varepsilon} \ln\left(1 + z e^{-\beta\varepsilon}\right), \tag{30}$$

where the summation has to be carried over all single-particle states in the system. Substituting (28) for ε, making use of the multiplicity factor (29) and replacing the summation over p_z by an integration, we get

$$\ln \mathcal{Q} = \int_{-\infty}^{\infty} \frac{L_z\, dp_z}{h} \left[\sum_{j=0}^{\infty} \left(L_x L_y \frac{eB}{hc} \right) \ln \left\{ 1 + z e^{-\beta e\hbar B[j+(1/2)]/mc - \beta p_z^2/2m} \right\} \right]. \tag{31}$$

At high temperatures, $z \ll 1$; so, the system is effectively *Boltzmannian*. The grand partition function then reduces to

$$\begin{aligned} \ln \mathcal{Q} &= \frac{zVeB}{h^2 c} \int_{-\infty}^{\infty} e^{-\beta p_z^2/2m}\, dp_z \sum_{j=0}^{\infty} e^{-\beta e\hbar B[j+(1/2)]/mc} \\ &= \frac{zVeB}{h^2 c} \left(\frac{2\pi m}{\beta} \right)^{1/2} \left\{ 2 \sinh \left(\frac{\beta e\hbar B}{2mc} \right) \right\}^{-1} \end{aligned} \tag{32}$$

The equilibrium number of particles $\overline{N}$ and the magnetic moment M of the gas are then given by

$$\overline{N} = \left(z \frac{\partial}{\partial z} \ln \mathcal{Q} \right)_{B,V,T} \tag{33}$$

and

$$M = \left\langle -\frac{\partial H}{\partial B} \right\rangle = \frac{1}{\beta} \left(\frac{\partial}{\partial B} \ln \mathcal{Q} \right)_{z,V,T}, \tag{34}$$

where H is the Hamiltonian of the system; cf. eqn. (3.9.4). We thus obtain

$$\overline{N} = \frac{zV}{\lambda^3} \frac{x}{\sinh x} \tag{35}$$

and

$$M = \frac{zV}{\lambda^3} \mu_{\text{eff}} \left\{ \frac{1}{\sinh x} - \frac{x \cosh x}{\sinh^2 x} \right\}, \tag{36}$$

where $\lambda\{= h/(2\pi mkT)^{1/2}\}$ is the mean thermal wavelength of the particles, while

$$x = \beta \mu_{\text{eff}} B \quad (\mu_{\text{eff}} = eh/4\pi mc). \tag{37}$$

Clearly, if e and m are the electronic charge and the electronic mass, then μ_{eff} is the familiar Bohr magneton μ_B. Combining (35) and (36), we get

$$M = -\overline{N} \mu_{\text{eff}} L(x), \tag{38}$$

where $L(x)$ is the *Langevin function*:

$$L(x) = \coth x - \frac{1}{x}. \tag{39}$$

This result is very similar to the one obtained in the Langevin theory of paramagnetism; see Sec. 3.9. The presence of the negative sign, however, means that the effect obtained in the present case is *diamagnetic* in character. We also note that this effect is a direct consequence of quantization; it vanishes if we let $h \to 0$. This is in complete accord with the *Bohr–van Leeuwen theorem*, according to which the phenomenon of diamagnetism does not arise in classical physics; see Problem 3.43.

If the field intensity B and the temperature T are such that $\mu_{\text{eff}}B \ll kT$, then the foregoing results become

$$\overline{N} \simeq \frac{zV}{\lambda^3} \tag{40}$$

and

$$M \simeq -\overline{N}\mu_{\text{eff}}^2 B/3kT. \tag{41}$$

Equation (40) is in agreement with the zero-field formula $z \simeq n\lambda^3$, while (41) leads to the diamagnetic counterpart of the Curie law:

$$\chi_\infty = \frac{M}{VB} = -\bar{n}\mu_{\text{eff}}^2/3kT; \tag{42}$$

cf. eqn. (3.9.12). It may be noted here that the diamagnetic character of this phenomenon is independent of the sign of the charge on the particle. For an electron gas, in particular, the net susceptibility at high temperatures is given by the sum of expression (7), with μ^* replaced by μ_B, and expression (42):

$$\chi_\infty = \frac{n\left(\mu_B^2 - \frac{1}{3}\mu_B'^2\right)}{kT}, \tag{43}$$

where $\mu'_B = eh/4\pi m'c$, m' being the *effective mass* of the electron in the given system.

We shall now look at this problem at *all* temperatures, though we will continue to assume the magnetic field to be weak, so that $\mu_{\text{eff}}B \ll kT$. In view of the latter, the summation in (31) may be handled with the help of the Euler summation formula,

$$\sum_{j=0}^{\infty} f\left(j + \tfrac{1}{2}\right) \simeq \int_0^\infty f(x)\,dx + \tfrac{1}{24}f'(0), \tag{44}$$

with the result

$$\ln \mathcal{Q} = \frac{VeB}{h^2c}\left[\int_0^\infty dx \int_{-\infty}^{\infty} dp_z \ln\left\{1 + ze^{-\beta(2\mu_{\text{eff}}Bx + p_z^2/2m)}\right\} - \frac{1}{12}\beta\mu_{\text{eff}}B\int_{-\infty}^{\infty}\frac{dp_z}{z^{-1}e^{\beta(p_z^2/2m)} + 1}\right] \tag{45}$$

The first part here is independent of B, which can be seen by changing the variable from x to $x' = Bx$. The second part, with the substitution $\beta p_z^2/2m = y$, becomes

$$-\frac{\pi V(2m)^{3/2}}{6h^3}(\mu_{\text{eff}}B)^2\beta^{1/2}\int_0^\infty \frac{y^{-1/2}\,dy}{z^{-1}e^y + 1}. \tag{46}$$

The low-field susceptibility (per unit volume) of the gas is then given by

$$\chi = \frac{M}{VB} = \frac{1}{\beta VB}\left(\frac{\partial}{\partial B}\ln \mathcal{Q}\right)_{z,V,T}$$

$$= -\frac{(2\pi m)^{3/2}\mu_{\text{eff}}^2}{3h^3\beta^{1/2}} f_{1/2}(z), \tag{47}$$

which is the desired result. Note that, as before, the effect is diamagnetic in character–irrespective of the sign of the charge on the particle.

For $z \ll 1$, $f_{1/2}(z) \simeq z \simeq n\lambda^3$; we then recover our previous result (42). For $z \gg 1$ (which corresponds to $T \ll T_F$), $f_{1/2}(z) \approx (2/\pi^{1/2})(\ln z)^{1/2}$; we then get

$$\chi_0 \approx -\frac{2\pi(2m)^{3/2}\mu_{\text{eff}}^2\varepsilon_F^{1/2}}{3h^3} = -\frac{1}{2}n\mu_{\text{eff}}^2/\varepsilon_F; \tag{48}$$

here, use has also been made of the fact that $(\beta^{-1}\ln z) \simeq \varepsilon_F$. Note that, in magnitude, this result is precisely one-third of the corresponding paramagnetic result, see eqn. (6), provided that we take the μ^* of that expression to be equal to the μ_{eff} of this one.

8.3. The electron gas in metals

One physical system where the application of Fermi–Dirac statistics helped remove a number of inconsistencies and discrepancies is that of conduction electrons in metals. Historically, the electron theory of metals was developed by Drude (1900) and Lorentz (1904–5), who applied the statistical mechanics of Maxwell and Boltzmann to the electron gas and derived theoretical results for the various properties of metals. The Drude–Lorentz model did provide a reasonable theoretical basis for a partial understanding of the physical behavior of metals; however, it encountered a number of serious discrepancies of a qualitative as well as quantitative nature. For instance, the observed specific heat of metals appeared to be almost completely accountable by the lattice vibrations alone and practically no contribution seemed to be coming from the electron gas. The theory, however, insisted that, on the basis of the equipartition theorem, each electron in the gas should possess a mean thermal energy $\frac{3}{2}kT$ and hence make a contribution of $\frac{3}{2}k$ to the specific heat of the metal. Similarly, one expected the electron gas to exhibit the phenomenon of paramagnetism arising from the intrinsic magnetic moment μ_B of the electrons. According to the classical theory, the paramagnetic susceptibility would be given by (8.2.7), with μ^* replaced by μ_B. Instead, one found that the susceptibility of a normal nonferromagnetic metal was not only independent of temperature but had a magnitude which, at room temperatures, was hardly 1 per cent of the expected value.

The Drude–Lorentz theory was also applied to study transport properties of metals, such as the thermal conductivity K and the electrical conductivity σ. While the results for the individual conductivities were not very encouraging, their ratio did conform to the empirical law of Wiedemann and Franz (1853), as formulated by Lorenz (1872), namely that the quantity $K/\sigma T$ was a (universal) constant. The theoretical value of this quantity, which is generally known as the *Lorenz number*, turned out to be $3(k/e)^2 \simeq 2.48 \times 10^{-13}$ e.s.u./deg^2; the corresponding experimental values for most alkali and alkaline-earth metals were, however, found to be scattered around a mean value of 2.72×10^{-13} e.s.u./deg^2. A still more uncomfortable feature of the classical theory was the uncertainty in assigning an appropriate value to the mean free path of the electrons in a given metal and in ascribing to it an appropriate temperature dependence. For these reasons, the problem of the transport properties of metals also remained in a rather unsatisfactory state until the correct lead was provided by Sommerfeld (1928).

The most significant change introduced by Sommerfeld was the replacement of Maxwell–Boltzmann statistics by Fermi–Dirac statistics for describing the electron gas in a metal. And, with a single stroke of genius, he was able to set most of the things right. To see how it worked, let us first of all estimate the Fermi energy ε_F of the electron gas in a typical metal, say sodium. Referring to eqn. (8.1.24), with $g = 2$, we obtain

$$\varepsilon_F = \left(\frac{3N}{8\pi V}\right)^{2/3} \frac{h^2}{2m'}, \tag{1}$$

where m' is the *effective mass* of an electron in the gas.[3] The electron density N/V, in the case of a cubic lattice, may be written as

$$\frac{N}{V} = \frac{n_e n_a}{a^3}, \tag{2}$$

where n_e is the number of conduction electrons per atom, n_a the number of atoms per unit cell and a the lattice constant (or the cell length).[4] For sodium, $n_e = 1$, $n_a = 2$ and $a = 4.29$ Å. Substituting these numbers into (2) and writing $m' = 0.98 m_e$, we obtain from (1)

$$(\varepsilon_F)_{\mathrm{Na}} = 5.03 \times 10^{-12} \text{ erg} = 3.14 \text{ eV}. \tag{3}$$

Accordingly, we obtain for the Fermi temperature of the gas

$$\begin{aligned}(T_F)_{\mathrm{Na}} &= (1.16 \times 10^4) \times \varepsilon_F \text{ (in eV)} \\ &= 3.64 \times 10^4 K,\end{aligned} \tag{4}$$

which is considerably larger than the room temperature $T(\sim 3 \times 10^2$ K). The ratio T/T_F being of the order of 1 per cent, the conduction electrons in sodium constitute a *highly degenerate* Fermi system. This statement, in fact, applies to *all* metals because their Fermi temperatures are generally of order 10^4–10^5 K.

Now, the very fact that the electron gas in metals is a highly degenerate Fermi system is sufficient to explain away some of the basic difficulties of the Drude–Lorentz theory. For instance, the specific heat of this gas would no longer be given by the classical formula, $C_V = \frac{3}{2}Nk$, but rather by eqn. (8.1.39), viz.

$$C_V = \frac{\pi^2}{2} Nk(kT/\varepsilon_F); \tag{5}$$

obviously, the new result is much smaller in value because, at ordinary temperatures, the ratio $(kT/\varepsilon_F) \equiv (T/T_F) = O(10^{-2})$. It is t[illegible] hardly surprising that, at ordinary temperatures, the specific heat of metals is almost completely determined by the vibrational modes of the lattice and very little contribution comes from the conduction electrons. Of course, as temperature decreases, the specific heat due to lattice vibrations also decreases and finally becomes considerably smaller than the classical value; see Sec. 7.3, especially Fig. 7.8. A stage comes when the two contributions, both nonclassical, become comparable in value. Ultimately, at very low temperatures, the specific heat due to lattice vibrations, being proportional to T^3, becomes *even smaller* than the electronic specific heat, which is proportional to T^1. In general, we may write, for the low-temperature specific heat of a metal,

$$C_V = \gamma T + \delta T^3, \tag{6}$$

where the coefficient γ is given by eqn. (5), or better by the general formula stated in Problem 8.13, while the coefficient δ is given by eqn. (7.3.23). An experimental determination of the specific heat of metals at low temperatures is, therefore, expected not only to verify the theoretical result based on quantum statistics but also to evaluate some of the parameters of the problem. Such determinations have been made, among others, by Corak *et al.* (1955) who worked with copper, silver and gold in the temperature range 1–5 K. Their results for copper are shown in Fig. 8.4. The very fact that the (C_V/T) vs. T^2 plot is well approximated by a straight line vindicates the theoretical formula (6). Furthermore, the slope of this line gives the value of the coefficient δ, from which one can extract the Debye temperature Θ_D of the metal. One gets for copper: $\Theta_D = (343.8 \pm 0.5)$ K, which compares favorably with Leighton's theoretical estimate of 345 K (based on the elastic constants of the metal). The intercept on the (C_V/T)-axis yields the value of the coefficient γ, viz. (0.688 ± 0.002) millijoule mole^{-1} deg^{-2}, which agrees favorably with Jones' estimate of 0.69 millijoule mole^{-1} deg^{-2} (based on a density-of-states calculation).

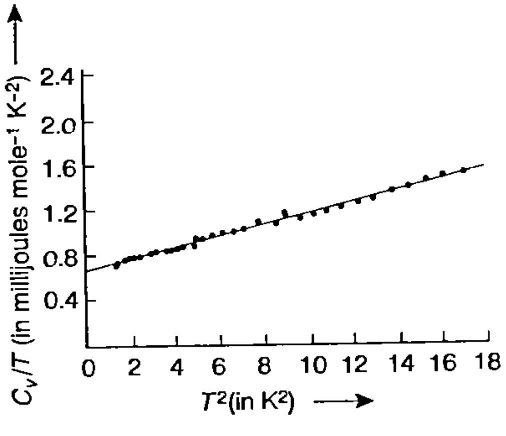


FIG. 8.4. The observed specific heat of copper in the temperature range 1–4 K (after Corak *et al.*, 1955).

The general pattern of the magnetic behavior of the electron gas in nonferromagnetic metals can be understood likewise. In view of the highly degenerate nature of the gas, the magnetic susceptibility χ is given by the Pauli result (8.2.6) plus the Landau result (8.2.48), and not by the classical result (8.2.7). In complete agreement with observation, the new result is (i) independent of temperature and (ii) considerably smaller in magnitude than the classical one.

As regards transport properties K and σ, the new theory again led to the *Wiedemann–Franz law*; the Lorenz number, however, became $(\pi^2/3)(k/e)^2$, instead of the classical $3(k/e)^2$. The resulting theoretical value, viz. 2.71×10^{-13} e.s.u./deg^2, turned out to be much closer to the experimental mean value quoted earlier. Of course, the situation regarding individual conductivities and the mean free path of the electrons did not improve until Bloch (1928) developed a theory that took into account interactions among the electron gas and the ion system in the metal. The theory of metals has steadily become more and more sophisticated; the important point, however, is that its development has all along been governed by the new statistics!

Before leaving this topic, we would like to give a brief account of the phenomena of *thermionic* and *photoelectric* emission (of electrons) from metals. In view of the fact that electronic emission does not take place spontaneously, we infer that the electrons inside a metal find themselves caught in some sort of a "potential well" created by the ions. The detailed features of the potential energy of an electron in this well must depend upon the structure of the given metal. For a study of electronic emission, however, we need not worry about these details and may assume instead that the potential energy of an electron stays constant (at a negative value, $-W$, say) throughout the interior of the metal and changes discontinuously to zero at the surface. Thus, while inside the metal, the electrons move about freely and independently of one another; however, as soon as any one of them approaches the surface of the metal and tries to escape, it encounters a potential barrier of height W. Accordingly, only those electrons whose kinetic energy (associated with the motion *perpendicular* to the surface) is greater than W can expect to escape through the surface barrier. At ordinary temperatures, especially in the absence of any external stimulus, such electrons are too few in any given metal, with the result that there is practically no spontaneous emission from metals. At high temperatures, and more so if there is an external stimulus present, the population of such electrons in a given metal could become large enough to yield a sizeable emission. We then speak of phenomena such as *thermionic effect* and *photoelectric effect*.

Strictly speaking, these phenomena are not equilibrium phenomena because electrons are flowing out *steadily* through the surface of the metal. However, if the number of electrons lost in a given interval of time is small in comparison with their total population in the metal, the magnitude of the emission current may be calculated on the assumption that the gas inside continues to be in a state of *quasi-static* thermal equilibrium. The mathematical procedure for this calculation would be very much the same as the one followed in Sec. 6.4 (for determining the *rate of effusion* R of the particles of a gas through an opening in the walls of the container). There is one difference, however: whereas in the case of effusion any particle that reached the opening with $u_z > 0$ could escape unquestioned, here we must have: $u_z > (2W/m)^{1/2}$, so that the particle in question could successfully cross over the potential barrier at the surface. Moreover, even if this condition is satisfied, there is no guarantee that the particle *will* really escape because the possibility of an inward reflection cannot be ruled out. In the following discussion, we shall disregard this possibility; however, if one is looking for a numerical comparison of theory with experiment, the results derived here must be multiplied by a factor $(1 - r)$, where r is the *reflection coefficient* of the surface.

Thermionic emission (the Richardson effect)

The number of electrons emitted per unit area of the metal surface per unit time is given by

$$R = \int\limits_{p_z=(2mW)^{1/2}}^{\infty} \int\limits_{p_x=-\infty}^{\infty} \int\limits_{p_y=-\infty}^{\infty} \left\{ \frac{2\,dp_x\,dp_y\,dp_z}{h^3} \frac{1}{e^{(\varepsilon-\mu)/kT}+1} \right\} u_z; \tag{7}$$

cf. the corresponding expression in Sec. 6.4. Integration over the variables p_x and p_y may be carried out by changing over to polar coordinates (p', ϕ), with the result

$$\begin{aligned} R &= \frac{2}{h^3} \int\limits_{p_z=(2mW)^{1/2}}^{\infty} \frac{p_z}{m} dp_z \int\limits_{p'=0}^{\infty} \frac{2\pi p'\,dp'}{\exp\{[(p'^2/2m)+(p_z^2/2m)-\mu]/kT\}+1} \\ &= \frac{4\pi kT}{h^3} \int\limits_{p_z=(2mW)^{1/2}}^{\infty} p_z\,dp_z \ln\left[1+\exp\{(\mu-p_z^2/2m)/kT\}\right] \\ &= \frac{4\pi mkT}{h^3} \int\limits_{\varepsilon_z=W}^{\infty} d\varepsilon_z \ln\left[1+e^{(\mu-\varepsilon_z)/kT}\right]. \end{aligned} \tag{8}$$

It so happens that the exponential term inside the logarithm, at all temperatures of interest, is much smaller than unity; see Note 5. We may, therefore, write $\ln(1+x) \simeq x$, with the result

$$\begin{aligned} R &= \frac{4\pi mkT}{h^3} \int\limits_{\varepsilon_z=W}^{\infty} d\varepsilon_z e^{(\mu-\varepsilon_z)/kT} \\ &= \frac{4\pi mk^2T^2}{h^3} e^{(\mu-W)/kT}. \end{aligned} \tag{9}$$

The thermionic current density is then given by

$$J = eR = \frac{4\pi mek^2}{h^3} T^2 e^{(\mu-W)/kT}. \tag{10}$$

It is only now that the difference between the classical statistics and the Fermi statistics really shows up. In the case of classical statistics, the fugacity of the gas is given by (see eqn. (8.1.4), with $f_{3/2}(z) \simeq z$)

$$z \equiv e^{\mu/kT} = \frac{n\lambda^3}{g} = \frac{nh^3}{2(2\pi mkT)^{3/2}}; \tag{11}$$

accordingly,

$$J_{\text{class}} = ne\left(\frac{k}{2\pi m}\right)^{1/2} T^{1/2} e^{-\phi/kT} \quad (\phi = W). \tag{12}$$

In the case of Fermi statistics, the chemical potential of the (highly degenerate) electron gas is practically independent of temperature and is very nearly equal to the Fermi energy of the gas ($\mu \simeq \mu_0 \equiv \varepsilon_F$); accordingly,

$$J_{\text{F.D.}} = \frac{4\pi mek^2}{h^3} T^2 e^{-\phi/kT} \quad (\phi = W - \varepsilon_F). \tag{13}$$

The quantity ϕ is generally referred to as the *work function* of the metal. According to (12), ϕ is exactly equal to the height of the surface barrier; according to (13), it is the height of the barrier *over and above the Fermi level* (see Fig. 8.5).

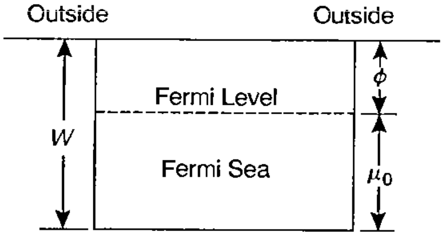


FIG. 8.5. The work function ϕ of a metal represents the height of the surface barrier *over and above the Fermi level.*

The theoretical results embodied in eqns (12) and (13) differ in certain important respects. The most striking difference seems to be in regard to the temperature dependence of the thermionic current density J. Actually, the major dependence here comes through the factor $\exp(-\phi/kT)$–so much so that whether we plot $\ln(J/T^{1/2})$ against $(1/T)$ or $\ln(J/T^2)$ against $(1/T)$ we obtain, in each case, a fairly good straight-line fit. Thus, from the point of view of the temperature dependence of J, a choice between formulae (12) and (13) is rather hard to make. However, the slope of the experimental line should give us *directly* the value of W if formula (12) applies or of $(W - \varepsilon_F)$ if formula (13) applies! Now, the value of W can be determined independently, for instance, by studying the "refractive index of a given metal for de Broglie waves associated with an electron beam impinging upon the metal". For a beam of electrons whose initial kinetic energy is E, we have

$$\lambda_{\text{out}} = \frac{h}{\sqrt{(2mE)}} \quad \text{and} \quad \lambda_{\text{in}} = \frac{h}{\sqrt{[2m(E+W)]}}, \tag{14}$$

so that the refractive index of the metal is given by

$$n = \frac{\lambda_{\text{out}}}{\lambda_{\text{in}}} = \left(\frac{E+W}{E}\right)^{1/2} \tag{15}$$

By studying electron diffraction for different values of E, one can derive the relevant value of W. In this manner, Davisson and Germer (1927) derived the value of W for a number of metals. For instance, they obtained for tungsten: $W \simeq 13.5$ eV. The experimental results on thermionic emission from tungsten are shown in Fig. 8.6. The value of ϕ resulting from the slope of the experimental line was about 4.5 eV. The large difference between these two values clearly shows that the classical formula (12) does not apply. That the quantum-statistical formula (13)

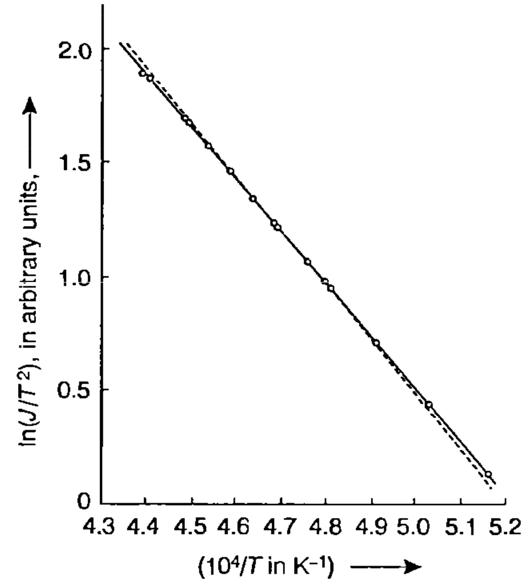


FIG. 8.6. Thermionic current from tungsten as a function of the temperature of the metal. The continuous line corresponds to $r = \frac{1}{2}$ while the broken line corresponds to $r = 0$, r being the reflection coefficient of the surface.

applies is supported is shown by the fact that the Fermi energy of tungsten is very nearly 9 eV; so, the value 4.5 eV for the work function of tungsten is correctly given by the difference between the depth W of the potential well and the Fermi energy ε_F. To quote another example, the experimental value of the work function for nickel was found to be about 5.0 eV, while the theoretical estimate for its Fermi energy turns out to be about 11.8 eV. Accordingly, the depth of the potential well in the case of nickel should be about 16.8 eV. The experimental value of W, obtained by studying electron diffraction in nickel, is indeed (17 ± 1) eV.[5]

The second point of difference between formulae (12) and (13) relates to the actual value of the current obtained. In this respect, the classical formula turns out to be a complete failure while the quantum-statistical formula fares reasonably well. The value of the constant factor in the latter formula is

$$\frac{4\pi mek^2}{h^3} = 120.4 \text{ amp cm}^{-2} \text{ deg}^{-2}; \tag{16}$$

of course, this has yet to be multiplied by the *transmission coefficient* $(1 - r)$. The corresponding experimental number, for most metals with clean surfaces, turns out to be in the range 60–120 amp $\text{cm}^{-2}\,\text{deg}^{-2}$.

Finally, we examine the influence of a moderately strong electric field on the thermionic emission from a metal–the so-called *Schottky effect*. Denoting the strength of the electric field by F and assuming the field to be uniform and directed perpendicular to the metal surface, the difference Δ between the potential energy of an electron at a distance x above the surface and of one inside the metal is given by

$$\Delta(x) = W - eFx - \frac{e^2}{4x} \quad (x > 0), \tag{17}$$

where the first arises from the potential well of the metal, the second from the (attractive) field present and the third from the attraction between the departing electron and the "image" induced in the metal; see Fig. 8.7. The largest value of the function $\Delta(x)$ occurs at $x = (e/4F)^{1/2}$, so that

$$\Delta_{\max} = W - e^{3/2}F^{1/2}; \qquad (18)$$

thus, the field has effectively reduced the height of the potential barrier by an amount $e^{3/2}F^{1/2}$. A corresponding reduction should take place in the work function as well.

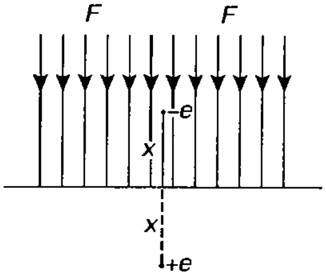


FIG. 8.7. A schematic diagram to illustrate the Schottky effect.

Accordingly, the thermionic current density in the presence of the field F would be higher than the one in the absence of the field:

$$J_F = J_0 \exp(e^{3/2}F^{1/2}/kT). \qquad (19)$$

A plot of $\ln(J_F/J_0)$ against $(F^{1/2}/T)$ should, therefore, be a straight line, with slope $e^{3/2}/k$. Working along these lines, de Bruyne (1928) obtained for the electronic charge a value of 4.84×10^{-10} e.s.u., which is remarkably close to the actual value of e.

The theory of the Schottky effect, as outlined here, holds for field strengths up to about 10^6 volts/cm. For fields stronger than that, one obtains the so-called *cold emission*, which means that the electric field is now strong enough to make the potential barrier practically ineffective; for details, see Fowler and Nordheim (1928).

B. Photoelectric emission (the Hallwachs effect)

The physical situation in the case of photoelectric emission is different from that in the case of thermionic emission, in that there exists now an external agency, the photon in the incoming beam of light, that helps an electron inside the metal in overcoming the potential barrier at the surface. The condition to be satisfied by the momentum component p_z of an electron in order that it could escape from the metal now becomes

$$(p_z^2/2m) + h\nu > W, \qquad (20)$$[6]

where ν is the frequency of the incoming light (assumed monochromatic). Proceeding in the same manner as in the case of thermionic emission, we obtain, instead of (8),

$$R = \frac{4\pi mkT}{h^3} \int_{\varepsilon_z = W - h\nu}^{\infty} d\varepsilon_z \ln \left[1 + e^{(\mu - \varepsilon_z)/kT}\right]. \tag{21}$$

We cannot, in general, approximate this integral the way we did there; the integrand here stays as it is. It is advisable, however, to change over to a new variable x, defined by

$$x = (\varepsilon_z - W + h\nu)/kT, \tag{22}$$

whereby eqn. (21) becomes

$$R = \frac{4\pi m(kT)^2}{h^3} \int_0^{\infty} dx \ln \left[1 + \exp\left\{\frac{h(\nu - \nu_0)}{kT} - x\right\}\right], \tag{23}$$

where

$$h\nu_0 = W - \mu \simeq W - \varepsilon_F = \phi. \tag{24}$$

The quantity ϕ will be recognized as the (thermionic) work function of the metal; accordingly, the characteristic frequency $\nu_0 (= \phi/h)$ may be referred to as the *threshold frequency* for photoelectric emission from the metal concerned. The current density of photoelectric emission is thus given by

$$J = \frac{4\pi mek^2}{h^3} T^2 \int_0^{\infty} dx \ln (1 + e^{\delta - x}), \tag{25}$$

where

$$\delta = h(\nu - \nu_0)/kT. \tag{26}$$

Integrating by parts, we find that

$$\int_0^{\infty} dx \ln (1 + e^{\delta - x}) = \int_0^{\infty} \frac{x\, dx}{e^{x - \delta} + 1} \equiv f_2(e^{\delta}); \tag{27}$$

see eqn. (8.1.6). Accordingly,

$$J = \frac{4\pi mek^2}{h^3} T^2 f_2(e^{\delta}). \tag{28}$$

For $h(\nu - \nu_0) \gg kT$, $e^{\delta} \gg 1$ and the function $f_2(e^{\delta}) \approx \delta^2/2$; see Sommerfeld's lemma (E.15). Equation (28) then becomes

$$J \approx \frac{2\pi me}{h} (\nu - \nu_0)^2, \tag{29}$$

which is completely independent of T; thus, when the energy of the light quantum is much greater than the work function of the metal, the temperature of the electron gas becomes a "dead" parameter of the problem. At the other extreme, when $\nu < \nu_0$ and $h|\nu - \nu_0| \gg kT$, then $e^{\delta} \ll 1$ and the function $f_2(e^{\delta}) \approx e^{\delta}$. Equation (28) then becomes

$$J \approx \frac{4\pi mek^2}{h^3} T^2 e^{(h\nu - \phi)/kT}, \tag{30}$$

which is just the thermionic current density (13), enhanced by the photon factor $\exp(h\nu/kT)$; in other words, the situation now is very much the same as in the case of thermionic emission, except for a diminished work function $\phi'(=\phi - h\nu)$. At the threshold frequency ($\nu = \nu_0$), $\delta = 0$ and the function $f_2(e^\delta) = f_2(1) = \pi^2/12$; see eqn. (E.14), with $j = 1$. Equation (28) then gives

$$J_0 = \frac{\pi^3 mek^2}{3h^3} T^2. \tag{31}$$

Figure 8.8 shows a plot of the experimental results for photoelectric emission from palladium ($\phi = 4.97$ eV). The agreement with theory is excellent. It will be noted that the plot includes some observations with $\nu < \nu_0$. The fact that we obtain a *finite* photo-current even for frequencies less than the so-called threshold frequency is fully consistent with the model considered here. The reason for this lies in the fact that, at any *finite* temperature T, there exists in the system a reasonable fraction of electrons whose energy ε exceeds the Fermi energy ε_F by amounts $O(kT)$. Therefore, if the light quantum $h\nu$ gets absorbed by one of these electrons, then condition (20) for photo-emission can be satisfied even if $h\nu < (W - \varepsilon_F) = h\nu_0$. Of course, the energy difference $h(\nu_0 - \nu)$ must not be much more than a few times kT, for otherwise the availability of the right kind of electrons will be extremely low. We therefore do expect a finite photo-current for radiation with frequencies less than the threshold frequency, provided that $h(\nu_0 - \nu) = O(kT)$.

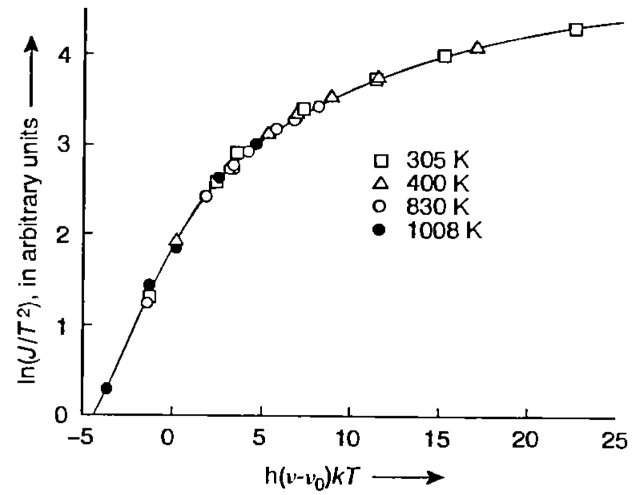


FIG. 8.8. Photoelectric current from palladium as a function of the quantity $h(\nu - \nu_0)/kT$. The plot includes data taken at several temperatures T for different frequencies ν.

The plot shown in Fig. 8.8, viz. $\ln(J/T^2)$ vs. δ, is generally known as the "Fowler plot". Fitting the observed photoelectric data to this plot, one can obtain the characteristic frequency ν_0 and hence the work function ϕ of the given metal. We have previously seen that the work function of a metal can be derived from thermionic data as well. It is gratifying to note that there is complete agreement between the two sets of results obtained for the work function of the various metals.

8.4. Statistical equilibrium of white dwarf stars

Historically, the first application of Fermi statistics appeared in the field of astrophysics (Fowler, 1926). It related to the study of thermodynamic equilibrium of *white dwarf stars*–the small-sized stars which are abnormally faint for their (white) color. The general pattern of color–brightness relationship among stars is such that, by and large, a star with red color is expected to be a "dull" star, while one with white color is expected to be a "brilliant" star. However, white dwarf stars constitute a glaring exception to this rule. The reason for this lies in the fact that these stars are relatively old ones whose hydrogen content is more or less used up, with the result that the thermonuclear reactions in them are now proceeding at a rather low pace, thus making these stars a lot less bright than one would expect on the basis of their color. The material content of white dwarf stars, at the present stage of their career, is mostly helium. And whatever little brightness they presently have derives mostly from the gravitational energy released as a result of a slow contraction of these stars–a mechanism first proposed by Kelvin, in 1861, as a "possible" source of energy for *all* stars!

A typical, though somewhat idealized, model of a white dwarf star consists of a mass $M(\sim 10^{33}$ g) of helium, packed into a ball of mass density $\rho(\sim 10^{7}$ g cm^{-3}), at a central temperature $T(\sim 10^{7}$ K). Now, a temperature of the order of 10^{7} K corresponds to a mean thermal energy (per particle) of the order of 10^{3} eV, which is much greater than the energy required for ionizing a helium atom. Thus, practically the whole of the helium in the star exists in a state of complete ionization. The microscopic constituents of the star may, therefore, be taken as N electrons (each of mass m) and $\frac{1}{2}N$ helium nuclei (each of mass $\simeq 4m_p$). The mass of the star is then given by

$$M \simeq N(m + 2m_p) \simeq 2Nm_p \tag{1}$$

and, hence, the electron density by

$$n = \frac{N}{V} \simeq \frac{M/2m_p}{M/\rho} = \frac{\rho}{2m_p}. \tag{2}$$

A typical value of the electron density in white dwarfs would, therefore, be $O(10^{30})$ electrons per cm^3. We thus obtain for the Fermi momentum of the electron gas (see eqn. (8.1.23), with $g = 2$)

$$p_F = \left(\frac{3n}{8\pi}\right)^{1/3} h = O(10^{-17}) \text{ g cm sec}^{-1}. \tag{3}$$

which is rather comparable with the characteristic momentum mc of an electron. The Fermi energy ε_F of the electron gas will, accordingly, be comparable with the rest energy mc^2 of an electron, i.e. $\varepsilon_F = O(10^6)$ eV and hence the Fermi temperature $T_F = O(10^{10})$ K. In view of these estimates, we conclude that (i) the dynamics of the electrons in this problem is *relativistic* and (ii) the electron gas, though at a temperature large in comparison with terrestrial standards, is, statistically speaking, in a state of (*almost*) *complete degeneracy*: $(T/T_F) = O(10^{-3})$. The second point was fully appreciated, and duly taken into account, by Fowler himself; the first one was taken care of later, by Anderson (1929) and by Stoner (1929–30). The problem, in full generality, was attacked by Chandrasekhar

(1931–35) to whom the final picture of the theo[illegible] white dwarf stars is chiefly due; for details, see Chandrasekhar (1939), where a complete bibliography of the subject is also given.

Now, the helium nuclei do not contribute as significantly to the dynamics of the problem as do the electrons; in the first approximation, therefore, we may neglect the presence of the nuclei in the system. For a similar reason, we may neglect the effect of the radiation as well. We may thus consider the electron gas alone. Further, for simplicity, we may assume that the electron gas is *uniformly* distributed over the body of the star; we are thus ignoring the spatial variation of the various parameters of the problem–a variation that is physically essential for the very stability of the star! The contention here is that, in spite of neglecting the spatial variation of the parameters involved, we expect that the results obtained will be correct, *at least* in a qualitative sense.

Let us then study the *ground-state* properties of a Fermi gas composed of N *relativistic* electrons ($g = 2$). First of all, we have

$$N = \frac{8\pi V}{h^3} \int_0^{p_F} p^2 \, dp = \frac{8\pi V}{3h^3} p_F^3, \tag{4}$$

whence

$$p_F = \left(\frac{3n}{8\pi}\right)^{1/3} h. \tag{5}$$

The energy–momentum relationship for a relativistic particle is

$$\varepsilon = mc^2[\{1 + (p/mc)^2\}^{1/2} - 1], \tag{6}$$

the speed of the particle being

$$u \equiv \frac{d\varepsilon}{dp} = \frac{(p/m)}{\{1 + (p/mc)^2\}^{1/2}}; \tag{7}$$

here, m denotes the rest mass of the electron. The pressure P_0 of the gas is then given by, see eqn. (6.4.3),

$$P_0 = \frac{1}{3}\frac{N}{V}\langle pu \rangle_0 = \frac{8\pi}{3h^3} \int_0^{p_F} \frac{(p^2/m)}{\{1 + (p/mc)^2\}^{1/2}} p^2 \, dp. \tag{8}$$

We now introduce a dimensionless variable θ, defined by

$$p = mc \sinh \theta, \tag{9}$$

which makes

$$u = c \tanh \theta. \tag{10}$$

Equations (4) and (8) then become

$$N = \frac{8\pi V m^3 c^3}{3h^3} \sinh^3 \theta_F = \frac{8\pi V m^3 c^3}{3h^3} x^3 \tag{11}$$

and

$$P_0 = \frac{8\pi m^4 c^5}{3h^3} \int_0^{\theta_F} \sinh^4 \theta \, d\theta = \frac{\pi m^4 c^5}{3h^3} A(x), \tag{12}$$

where

$$A(x) = x(x^2+1)^{1/2}(2x^2-3) + 3\sinh^{-1} x, \tag{13}$$

with

$$x = \sinh \theta_F = p_F/mc = (3n/8\pi)^{1/3}(h/mc). \tag{14}$$

The function $A(x)$ can be computed for any desired value of x. However, asymptotic results (for $x \ll 1$ and $x \gg 1$) are often useful; these are given by (see Kothari and Singh, 1942)

$$\left.\begin{aligned} A(x) &= \tfrac{8}{5}x^5 - \tfrac{4}{7}x^7 + \tfrac{1}{3}x^9 - \tfrac{5}{22}x^{11} + \qquad && \text{for} \quad x \ll 1 \\ &= 2x^4 - 2x^2 + 3\left(\ln 2x - \tfrac{7}{12}\right) + \tfrac{5}{4}x^{-2} + \cdots && \text{for} \quad x \gg 1 \end{aligned}\right\} \tag{15}$$

We shall now consider, somewhat crudely, the equilibrium configuration of this model. In the absence of gravitation, it will be necessary to have "external walls" for keeping the electron gas at a given density n. The gas will exert a pressure $P_0(n)$ on the walls and any compression or expansion (of the gas) will involve an expenditure of work. Assuming the configuration to be spherical, an adiabatic change in V will cause a change in the energy of the gas, as given by

$$dE_0 = -P_0(n)\,dV = -P_0(R) \cdot 4\pi R^2\, dR. \tag{16}$$

In the presence of gravitation, no external walls are needed, but the change in the kinetic energy of the gas, as a result of a change in the size of the sphere, will still be given by formula (16); of course, the expression for P_0, as a function of the "mean" density n, must now take into account the nonuniformity of the system–a fact being disregarded in the present simple-minded treatment. However, eqn. (16) alone no longer gives us the net change in the energy of the system; if that were the case, the system would expand indefinitely till both n and $P_0(n) \to 0$. Actually, there arises now a change in the *potential* energy as well: this is given by

$$dE_g = \left(\frac{dE_g}{dR}\right) dR = \alpha \frac{GM^2}{R^2}\, dR, \tag{17}$$

where M is the total mass of the gas, G the constant of gravitation, while α is a number (of the order of unity) whose exact value depends upon the nature of the spatial variation of n inside the sphere. If the system is in equilibrium, then the net change in its total energy $(E_0 + E_g)$, for an infinitesimal change in its size, should be zero; thus, for equilibrium,

$$P_0(R) = \frac{\alpha}{4\pi} \frac{GM^2}{R^4}. \tag{18}$$

For $P_0(R)$ we may substitute from eqn. (12), where the parameter x is now given by

$$x = \left(\frac{3n}{8\pi}\right)^{1/3} \frac{h}{mc} = \left(\frac{9N}{32\pi^2}\right)^{1/3} \frac{h/mc}{R}$$

or, in view of (1), by

$$x = \left(\frac{9M}{64\pi^2 m_p}\right)^{1/3}\frac{\hbar/mc}{R} = \left(\frac{9\pi M}{8m_p}\right)^{1/3}\frac{\hbar/mc}{R} \tag{19}$$

Equation (18) then takes the form

$$A\left(\left\{\frac{9\pi M}{8m_p}\right\}^{1/3}\frac{\hbar/mc}{R}\right) = \frac{3\alpha\hbar^3}{4\pi^2 m^4 c^5}\frac{GM^2}{R^4}$$

$$= 6\pi\alpha\left(\frac{\hbar/mc}{R}\right)^3\frac{GM^2/R}{mc^2}; \tag{20}$$

the function $A(x)$ is given by eqns (13) and (15).

Equation (20) establishes a one-to-one correspondence between the masses M and the radii R of white dwarf stars; it is, therefore, generally known as the *mass–radius relationship* for these stars. It is rather interesting to see the combinations of parameters that appear in this relationship; we have here (i) the mass of the star in terms of the proton mass, (ii) the radius of the star in terms of the Compton wavelength of the electron, and (iii) the gravitational energy of the star in terms of the rest energy of the electron. This relationship, therefore, exhibits a remarkable blending of quantum mechanics, special relatively and gravitation.

In view of the implicit character of relationship (20), we cannot express the radius of the star as an explicit function of its mass, except in two extreme cases. For this, we note that, since $M \sim 10^{33}$ g, $m_p \sim 10^{-24}$ g and $\hbar/mc \sim 10^{-11}$ cm, the argument of the function $A(x)$ will be of the order of unity when $R \sim 10^8$ cm. We may, therefore, define the two extreme cases as follows:

(i) $R \gg 10^8$ cm, which makes $x \ll 1$ and hence $A(x) \approx \frac{8}{5}x^5$, with the result

$$R \approx \frac{3(9\pi)^{2/3}}{40\alpha}\frac{\hbar^2 M^{-1/3}}{Gmm_p^{5/3}} \propto M^{-1/3} \tag{21}$$

(ii) $R \ll 10^8$ cm, which makes $x \gg 1$ and hence $A(x) \approx 2x^4 - 2x^2$, with the result

$$R \approx \frac{(9\pi)^{1/3}}{2}\frac{\hbar}{mc}\left(\frac{M}{m_p}\right)^{1/3}\left\{1 - \left(\frac{M}{M_0}\right)^{2/3}\right\}^{1/2} \tag{22}$$

where

$$M_0 = \frac{9}{64}\left(\frac{3\pi}{\alpha^3}\right)^{1/2}\frac{(\hbar c/G)^{3/2}}{m_p^2}. \tag{23}$$

We thus find that the greater the mass of a white dwarf star, the smaller its size. Not only that, there exists a limiting mass M_0, given by expression (23), that corresponds to a vanishing size of the star. Obviously, for $M > M_0$, our mass–radius relationship does not possess any real solution. We therefore conclude that all white dwarf stars in equilibrium must have a mass less than M_0–a conclusion fully upheld by observation.

The correct limiting mass of a white dwarf star i nerally referred to as the *Chandrasekhar limit*. The physical reason for the e ence of this limit is that for a mass exceeding this limit the ground-state pressure of the electron gas (that arises from the fact that the electrons obey the Pauli exclusion principle) would not be sufficient to support the system against its "tendency towards a gravitational collapse". The numerical value of the limiting mass, as given by expression (23), turns out to be $\sim 10^{33}$ g. Detailed investigations by Chandrasekhar led to the result:

$$M_0 = \frac{5.75}{\mu_e^2}\odot, \tag{24}$$

where $\odot$ denotes the mass of the sun, which is about 2×10^{33} g, while μ_e is a number that represents the degree of ionization in the gas. By definition, $\mu_e = M/Nm_H$; cf. eqn. (1). Thus, in most cases, $\mu_e \simeq 2$. Accordingly, $M_0 \simeq 1.44\odot$.

Figure 8.9 shows a plot of the theoretical relationship between the masses and the radii of white dwarf stars. One can see that the behavior in the two extreme regions, viz. for $R \gg l$ and $R \ll l$, is simulated quite well by formulae (21) and (22) of the treatment given here.

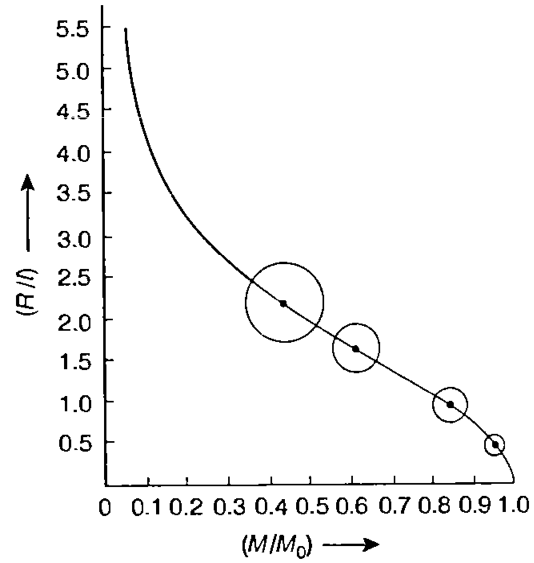


FIG. 8.9. The mass-radius relationship for white dwarfs (after Chandrasekhar, 1939). The masses are expressed in terms of the limiting mass M_0 and the radii in terms of a characteristic length l, which is given by $7.71\mu_e^{-1} \times 10^8$ cm $\simeq 3.86 \times 10^8$ cm.

8.5. Statistical model of the atom

Another application of the Fermi statistics was made by Thomas (1927) and Fermi (1928) for calculating the charge distribution and the electric field in the extra-nuclear space of a heavy atom. Their approach was based on the observation that the electrons in this system could be regarded as a completely degenerate Fermi gas of *non-uniform* density $n(\boldsymbol{r})$. By considering the equilibrium state of the configuration, one arrives at a differential equation whose solution gives directly the electric potential $\phi(\boldsymbol{r})$ and the electron density $n(\boldsymbol{r})$ at the point $\boldsymbol{r}$. By the very nature of the model, which is generally referred to as the *Thomas–Fermi model* of the atom, the resulting function $n(\boldsymbol{r})$ is a smoothly varying function of $\boldsymbol{r}$, devoid of the "peaks" that would remind one of the electron orbits of the Bohr theory. Nevertheless, the model has proved quite useful in deriving composite properties such as the binding energy of the atom. And, after suitable modifications, it has been successfully applied to molecules, solids and nuclei as well. Here, we propose to outline only the simplest treatment of the model, as applied to an atomic system; for further details and other applications, see Gombás (1949, 1952) and March (1957), where references to other contributions to the subject can also be found.

According to the statistics of a completely degenerate Fermi gas, we have exactly two electrons (with opposite spins) in each elementary cell, of the phase space, with $p \le p_F$; the Fermi momentum p_F of the electron gas is determined by the electron density n, according to the formula

$$p_F = (3\pi^2 n)^{1/3}\hbar. \tag{1}$$

In the system under study, the electron density varies from point to point; so would the value of p_F. We must, therefore, speak of the limiting momentum p_F as a function of $\boldsymbol{r}$, which is clearly a "quasi-classical" description of the situation. Such a description is justifiable if the de Broglie wavelength of the electrons in a given region of space is much smaller than the distance over which the functions $p_F(\boldsymbol{r})$, $n(\boldsymbol{r})$ and $\phi(\boldsymbol{r})$ undergo a significant variation; later on, it will be seen that this requirement is satisfied reasonably well by the heavier atoms.

Now, the total energy ε of an electron at the top of the Fermi sea *at the point* $\boldsymbol{r}$ is given by

$$\varepsilon(\boldsymbol{r}) = \frac{1}{2m} p_F^2(\boldsymbol{r}) - e\phi(\boldsymbol{r}), \tag{2}$$

where e denotes the magnitude of the electronic charge. When the system is in a *stationary* state, the value of $\varepsilon(\boldsymbol{r})$ should be the same throughout, so that electrons anywhere in the system do not have an overall tendency to "flow away" towards other parts of the system. Now, at the boundary of the system, p_F must be zero; by a suitable choice of the zero of energy, we can also have $\phi = 0$ there. Thus, the value of ε at the boundary of the system is zero; so must, then, be the value of ε throughout the system. We thus have, for all $\boldsymbol{r}$,

$$\frac{1}{2m} p_F^2(\boldsymbol{r}) - e\phi(\boldsymbol{r}) = 0. \tag{3}$$

Substituting from (1) and making use of the Poisson equation,

$$\nabla^2\phi(\boldsymbol{r}) = -4\pi\rho(\boldsymbol{r}) = 4\pi e n(\boldsymbol{r}), \tag{4}$$

we obtain

$$\nabla^2\phi(\boldsymbol{r}) = \frac{4e(2me)^{3/2}}{3\pi\hbar^3}\{\phi(\boldsymbol{r})\}^{3/2}. \tag{5}$$

Assuming spherical symmetry, eqn. (5) takes the form

$$\frac{1}{r^2}\frac{d}{dr}\left\{r^2\frac{d}{dr}\phi(r)\right\} = \frac{4e(2me)^{3/2}}{3\pi\hbar^3}\{\phi(r)\}^{3/2}, \tag{6}$$

which is known as the *Thomas–Fermi equation* of the system. Introducing dimensionless variables x and Φ, defined by

$$x = 2\left(\frac{4}{3\pi}\right)^{2/3} Z^{1/3}\frac{me^2}{\hbar^2}r = \frac{Z^{1/3}}{0.88534a_B}r \tag{7}$$

and

$$\Phi(x) = \frac{\phi(r)}{Ze/r}, \tag{8}$$

where Z is the atomic number of the system and a_B the first Bohr radius of the hydrogen atom, eqn. (6) reduces to

$$\frac{d^2\Phi}{dx^2} = \frac{\Phi^{3/2}}{x^{1/2}}. \tag{9}$$

Equation (9) is the *dimensionless Thomas–Fermi equation* of the system. The boundary conditions on the solution to this equation can be obtained as follows. As we approach the nucleus of the system ($r \to 0$), the potential $\phi(r)$ approaches the unscreened value Ze/r; accordingly, we must have: $\Phi(x \to 0) = 1$. On the other hand, as we approach the boundary of the system ($r \to r_0$), $\phi(r)$ in the case of a neutral atom must tend to zero; accordingly, we must have: $\Phi(x \to x_0) = 0$. In principle, these two conditions are sufficient to determine the function $\Phi(x)$ completely. However, it would be helpful if one knew the initial slope of the function as well, which in turn would depend upon the precise location of the boundary. Choosing the boundary to be at infinity ($r_0 = \infty$), the appropriate initial slope of the function $\Phi(x)$ turns out to be very nearly -1.5886; in fact, the nature of the solution near the origin is

$$\Phi(x) = 1 - 1.5886x + \tfrac{4}{3}x^{3/2} + \cdots \tag{10}$$

For $x > 10$, the approximate solution has been determined by Sommerfeld (1932).

$$\Phi(x) \approx \left\{1 + \left(\frac{x^3}{144}\right)^{\lambda}\right\}^{-1/\lambda} \tag{11}$$

where

$$\lambda = \frac{\sqrt{(73)} - 7}{6} \simeq 0.257. \tag{12}$$

As $x \to \infty$, the solution tends to the simple form: $\Phi(x) \approx 144/x^3$. The complete solution, which is a monotonically decreasing function of x, has been tabulated by Bush and Caldwell (1931). As a check on the numerical results, we note that the solution must satisfy the integral condition

$$\int_0^\infty \Phi^{3/2}x^{1/2}\,dx = 1, \tag{13}$$

which expresses the fact that the integral of the electron density $n(r)$ over the whole of the space available to the system must be equal to Z, the total number of electrons present.

From the function $\Phi(x)$, one readily obtains the electric potential $\phi(r)$ and the electron density $n(r)$:

$$\phi(r) = \frac{Ze}{r}\Phi\left(\frac{rZ^{1/3}}{0.88534a_B}\right) \propto Z^{4/3} \tag{14}$$

and

$$n(r) = \frac{(2me)^{3/2}}{3\pi^2\hbar^3}\{\phi(r)\}^{3/2} \propto Z^2. \tag{15}$$

Figure 8.10 shows a Thomas–Fermi plot of the electron distribution function $D(r)\{= n(r) \cdot 4\pi r^2\}$ for an atom of mercury; the actual "peaked" distribution, which conveys unmistakably the preference of the electrons to be in the vicinity of their semi-classical orbits, is also shown in the figure.

To calculate the *binding energy* of the atom, we should determine the total energy of the electron cloud. Now, the mean kinetic energy of an electron at the point $\boldsymbol{r}$ would be $\frac{3}{5}\varepsilon_F(r)$; by eqn. (3), this is equal to $\frac{3}{5}e\phi(r)$. The total kinetic energy of the electron cloud is, therefore, given by

$$\frac{3}{5}e\int_0^\infty \phi(r)n(r) \cdot 4\pi r^2\,dr. \tag{16}$$

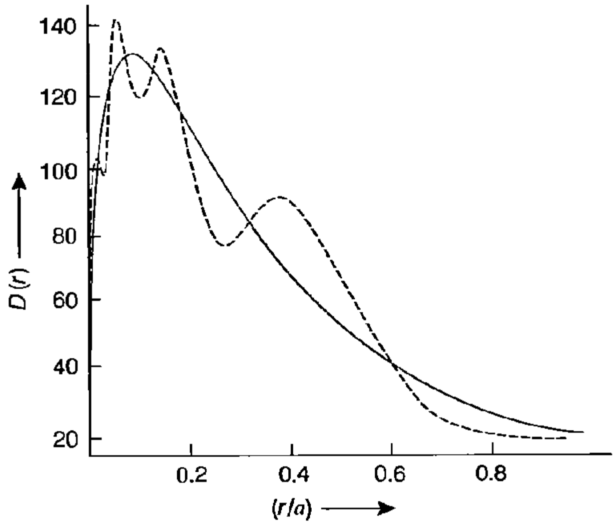


FIG. 8.10. The electron distribution function $D(r)$ for an atom of mercury. The distance r is expressed in terms of the atomic unit of length $a(= \hbar^2/me^2)$.

For the potential energy of the cloud, we note that a part of the potential $\phi(r)$ at the point $\boldsymbol{r}$ is due to the nucleus of the atom while the rest of it is due to the electron cloud itself; the former is clearly (Ze/r), so the latter must be $\{\phi(r) - Ze/r\}$. The total potential energy of the cloud is, therefore, given by

$$-e\int_0^\infty \left[\frac{Ze}{r} + \frac{1}{2}\left\{\phi(r) - \frac{Ze}{r}\right\}\right] n(r) \cdot 4\pi r^2\,dr. \tag{17}$$

We thus obtain for the total energy of the cloud

$$E_0 = \int_0^\infty \left\{ \frac{1}{10} e\phi(r) - \frac{1}{2}\frac{Ze^2}{r} \right\} n(r) \cdot 4\pi r^2 \, dr; \tag{18}$$

of course, the electron density $n(r)$, in terms of the potential function $\phi(r)$, is given by eqn. (15).

Now, Milne (1927) has shown that the integrals

$$\int_0^\infty \{\phi(r)\}^{5/2} r^2 \, dr \quad \text{and} \quad \int_0^\infty \{\phi(r)\}^{3/2} r \, dr, \tag{19}$$

that appear in the expression for E_0, can be expressed directly in terms of the *initial* slope of the function $\Phi(x)$, i.e. in terms of the number -1.5886 of eqn. (10). After a little calculus, one finds that

$$E_0 = \frac{1.5886}{0.88534} \left(\frac{e^2}{2a_B} \right) Z^{7/3} \left(\frac{1}{7} - 1 \right), \tag{20}$$

whence one obtains for the (Thomas–Fermi) *binding energy* of the atom:

$$E_B = -E_0 = 1.538 Z^{7/3} \chi, \tag{21}$$

where $\chi (= e^2/2a_B \simeq 13.6 \text{ eV})$ is the (actual) binding energy of the hydrogen atom. It is clear that our statistical result (21) cannot give us anything more than just the first term of an "asymptotic expansion" of the binding energy E_B in powers of the parameter $Z^{-1/3}$. For practical values of Z, other terms of the expansion are also important; however, they cannot be obtained from the simple-minded treatment given here. The interested reader may refer to the review article by March (1957).

In the end we observe that, since the total energy of the electron cloud is proportional to $Z^{7/3}$, the *mean* energy per electron would be proportional to $Z^{4/3}$; accordingly, the mean de Broglie wavelength of the electrons in the cloud would be proportional to $Z^{-2/3}$. At the same time, the overall linear dimensions of the cloud are proportional to $Z^{-1/3}$; see eqn. (7). We thus find that the quasi-classical description adopted in the Thomas–Fermi model is more appropriate for heavier atoms (so that $Z^{-2/3} \ll Z^{-1/3}$). Otherwise, too, the statistical nature of the approach demands that the number of particles in the system be large.

Problems

8.1. Let the Fermi distribution at low temperatures be represented by a *broken line*, as shown in Fig. 8.11, the line being tangential to the actual curve at $\varepsilon = \mu$. Show that this approximate representation yields a "correct" result for the low-temperature specific heat of the Fermi gas, except that the numerical factor turns out to be smaller by a factor of $4/\pi^2$. Discuss, in a qualitative manner, the origin of the numerical discrepancy.

8.2. For a Fermi–Dirac gas, we may define a temperature T_0 at which the chemical potential of the gas is zero ($z = 1$). Express T_0 in terms of the Fermi temperature T_F of the gas.

[*Hint*: Use eqn. (E. 14).]

8.3. Show that for an ideal Fermi gas

$$\frac{1}{z}\left(\frac{\partial z}{\partial T} \right)_P = -\frac{5}{2T} \frac{f_{5/2}(z)}{f_{3/2}(z)};$$

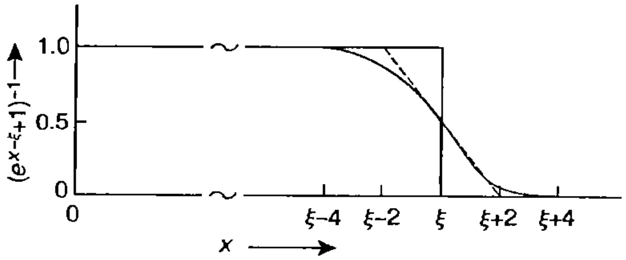


FIG. 8.11. An *approximate* representation of the Fermi distribution at low temperatures: here, $x = \varepsilon/kT$ and $\xi = \mu/kT$.

cf. eqn. (8.1.9). Hence show that

$$\gamma \equiv \frac{C_P}{C_V} = \frac{(\partial z/\partial T)_P}{(\partial z/\partial T)_v} = \frac{5}{3}\frac{f_{5/2}(z) f_{1/2}(z)}{\{f_{3/2}(z)\}^2}.$$

Check that at low temperatures

$$\gamma \simeq 1 + \frac{\pi^2}{3}\left(\frac{kT}{\varepsilon_F}\right)^2.$$

8.4. (a) Show that the isothermal compressibility κ_T and the adiabatic compressibility κ_S of an ideal Fermi gas are given by

$$\kappa_T = \frac{1}{nkT}\frac{f_{1/2}(z)}{f_{3/2}(z)}, \quad \kappa_S = \frac{3}{5nkT}\frac{f_{3/2}(z)}{f_{5/2}(z)},$$

where $n(= N/V)$ is the particle density in the gas. Check that at low temperatures

$$\kappa_T \simeq \frac{3}{2n\varepsilon_F}\left[1 - \frac{\pi^2}{12}\left(\frac{kT}{\varepsilon_F}\right)^2\right], \quad \kappa_S \simeq \frac{3}{2n\varepsilon_F}\left[1 - \frac{5\pi^2}{12}\left(\frac{kT}{\varepsilon_F}\right)^2\right]$$

(b) Making use of the thermodynamic relation

$$C_P - C_V = T\left(\frac{\partial P}{\partial T}\right)_V\left(\frac{\partial V}{\partial T}\right)_P = TV\kappa_T\left(\frac{\partial P}{\partial T}\right)_V^2$$

show that

$$\frac{C_P - C_V}{C_V} = \frac{4}{9}\frac{C_V}{Nk}\frac{f_{1/2}(z)}{f_{3/2}(z)}$$

$$\simeq \frac{\pi^2}{3}\left(\frac{kT}{\varepsilon_F}\right)^2 \quad (kT \ll \varepsilon_F).$$

(c) Finally, making use of the thermodynamic relation $\gamma = \kappa_T/\kappa_S$, verify the results of Problem 8.3.

8.5. Evaluate $(\partial^2 P/\partial T^2)_v$, $(\partial^2 \mu/\partial T^2)_v$ and $(\partial^2 \mu/\partial T^2)_P$ of an ideal Fermi gas and check that your results satisfy the thermodynamic relations

$$C_V = VT\left(\frac{\partial^2 P}{\partial T^2}\right)_v - NT\left(\frac{\partial^2 \mu}{\partial T^2}\right)_v$$

and

$$C_P = -NT\left(\frac{\partial^2 \mu}{\partial T^2}\right)_P.$$

Examine the low-temperature behavior of these quantities.

8.6. Show th velocity of sound w in an ideal Fermi gas is given by

$$w^2 = \frac{5kT}{3m}\frac{f_{5/2}(z)}{f_{3/2}(z)} = \frac{5}{9}\langle u^2\rangle,$$

where $\langle u^2\rangle$ is the mean square speed of the particles in the gas. Evaluate w in the limit $z \to \infty$ and compare it with the Fermi velocity u_F.

8.7. Show that for an ideal Fermi gas

$$\langle u\rangle\left\langle\frac{1}{u}\right\rangle = \frac{4}{\pi}\frac{f_1(z)f_2(z)}{\{f_{3/2}(z)\}^2},$$

u being the speed of a particle. Further show that at low temperatures

$$\langle u\rangle\left\langle\frac{1}{u}\right\rangle \simeq \frac{9}{8}\left[1+\frac{\pi^2}{12}\left(\frac{kT}{\varepsilon_F}\right)^2\right];$$

cf. Problem 6.6.

8.8. Obtain numerical estimates of the Fermi energy (in eV) and the Fermi temperature (in K) for the following systems:

(i) conduction electrons in silver, lead and aluminum,
(ii) nucleons in a heavy nucleus, such as $_{80}\mathrm{Hg}^{200}$, and
(iii) He^3 atoms in liquid helium three (atomic volume: 63 $Å^3$ per atom).

8.9. Making use of another term of the Sommerfeld lemma (E.15), show that *in the second approximation* the chemical potential of a Fermi gas at low temperatures is given by

$$\mu \simeq \varepsilon_F\left[1-\frac{\pi^2}{12}\left(\frac{kT}{\varepsilon_F}\right)^2-\frac{\pi^4}{80}\left(\frac{kT}{\varepsilon_F}\right)^4\right] \tag{8.1.35a}$$

and the mean energy per particle by

$$\frac{U}{N} \simeq \frac{3}{5}\varepsilon_F\left[1+\frac{5\pi^2}{12}\left(\frac{kT}{\varepsilon_F}\right)^2-\frac{\pi^4}{16}\left(\frac{kT}{\varepsilon_F}\right)^4\right]. \tag{8.1.37a}$$

Hence determine the T^3-correction to the customary T^1-result for the specific heat of an electron gas. Compare the magnitude of the T^3-term, in a typical metal such as copper, with the low-temperature specific heat arising from the Debye modes of the lattice. For further terms of these expansions, see Kiess (1987).

8.10. Consider an ideal Fermi gas, with energy spectrum $\varepsilon \propto p^s$, contained in a box of "volume" V in a space of n dimensions. Show that, for this system,

(i) $PV = \frac{s}{n}U$;

(ii) $\frac{C_V}{Nk} = \frac{n}{s}\left(\frac{n}{s}+1\right)\frac{f_{(n/s)+1}(z)}{f_{n/s}(z)} - \left(\frac{n}{s}\right)^2\frac{f_{n/s}(z)}{f_{(n/s)-1}(z)}$;

(iii) $\frac{C_P - C_V}{Nk} = \left(\frac{sC_V}{nNk}\right)^2\frac{f_{(n/s)-1}(z)}{f_{n/s}(z)}$;

(iv) the equation of an *adiabat* is: $PV^{1+(s/n)} = \text{const.}$,

(v) the index $(1+(s/n))$ in the foregoing equation agrees with the ratio (C_P/C_V) of the gas only when $T \gg T_F$. On the other hand, when $T \ll T_F$, the ratio $(C_P/C_V) \simeq 1 + (\pi^2/3)(kT/\varepsilon_F)^2$, irrespective of the values of s and n.

8.11. Examine results (ii) and (iii) of the preceding problem in the high-temperature limit $(T \gg T_F)$ as well as in the low-temperature limit $(T \ll T_F)$, and compare the resulting expressions with the ones pertaining to a nonrelativistic gas and an extreme relativistic gas in three dimensions.

8.12. Show that, in two dimensions, the specific heat $C_V(N, T)$ of an ideal Fermi gas is identical with the specific heat of an ideal Bose gas, for *all* N and T.

[*Hint*: It will suffice to show that, for given N and T, the thermal energies of the two systems differ at most by a constant. For this, first show that the fugacities, z_F and z_B, of the two systems are mutually related:

$$(1+z_F)(1-z_B)=1, \quad \text{i.e.} \quad z_B = z_F/(1+z_F).$$

Next, show that the functions $f_2(z_F)$ and $g_2(z_B)$ are also related:

$$f_2(z_F) = \int_0^{z_F} \frac{\ln(1+z)}{z}\,dz$$

$$= g_2\left(\frac{z_F}{1+z_F}\right) + \frac{1}{2}\ln^2(1+z_F).$$

It is now straightforward to show that

$$E_F(N,T) = E_B(N,T) + \text{const.},$$

the constant being $E_F(N,0)$.]

8.13. Show that, *quite generally*, the low-temperature behavior of the chemical potential, the specific heat and the entropy of an ideal Fermi gas is given by

$$\mu \simeq \varepsilon_F\left[1 - \frac{\pi^2}{6}\left(\frac{\partial \ln a(\varepsilon)}{\partial \ln \varepsilon}\right)_{\varepsilon=\varepsilon_F}\left(\frac{kT}{\varepsilon_F}\right)^2\right]$$

and

$$C_V \simeq S \simeq \frac{\pi^2}{3}k^2 T a(\varepsilon_F),$$

where $a(\varepsilon)$ is the *density of (the single-particle) states* in the system. Examine these results for a gas with energy spectrum $\varepsilon \propto p^s$, confined to a space of n dimensions, and discuss the special cases: $s = 1$ and 2, with $n = 2$ and 3.

[*Hint*: Use eqn. (E. 16).]

8.14. Investigate the Pauli paramagnetism of an ideal gas of fermions with intrinsic magnetic moment μ^* and spin $J\hbar(J = \frac{1}{2}, \frac{3}{2}, \ldots)$, and derive expressions for the low-temperature and high-temperature susceptibilities of the gas.

8.15. Show that expression (8.2.20) for the paramagnetic susceptibility of an ideal Fermi gas can be written in the form

$$\chi = \frac{n\mu^{*2}}{kT}\frac{f_{1/2}(z)}{f_{3/2}(z)}.$$

Using this result, verify eqns (8.2.24) and (8.2.27).

8.16. The observed value of γ, see eqn. (8.3.6), for sodium is 4.3×10^{-4} cal mole^{-1} K^{-2}. Evaluate the Fermi energy ε_F and the number density n of the conduction electrons in the sodium metal. Compare the latter result with the number density of atoms (given that, for sodium, $\rho = 0.954$ g cm^{-3} and $M = 23$).

8.17. Calculate the fraction of the conduction electrons in tungsten ($\varepsilon_F = 9.0$ eV) at 3000 K whose kinetic energy ε $\left(=\frac{1}{2}mu^2\right)$ is greater than $W(= 13.5$ eV$)$. Also calculate the fraction of the electrons whose kinetic energy associated with the z-component of their motion, namely $\left(\frac{1}{2}mu_z^2\right)$, is greater than 13.5 eV.

8.18. Show that the ground-state energy E_0 of a *relativistic* gas of electrons is given by

$$E_0 = \frac{\pi V m^4 c^5}{3h^3}B(x),$$

where

$$B(x) = 8x^3\{(x^2+1)^{1/2} - 1\} - A(x),$$

$A(x)$ and x being given by eqns (8.4.13) and (8.4.14). Check that the foregoing result for E_0 and eqn. (8.4.12) for P_0 satisfy the thermodynamic relations

$$E_0 + P_0 V = N\mu_0 \quad \text{and} \quad P_0 = -(\partial E_0/\partial V)_N.$$

8.19. Show that the low-temperature specific heat of the relativisitic Fermi gas, studied in Sec. 8.4, is given by

$$\frac{C_V}{Nk} = \pi^2 \frac{(x^2+1)^{1/2}}{x^2} \frac{kT}{mc^2} \quad \left(x = \frac{p_F}{mc}\right).$$

Check that this formula gives correct results for the nonrelativistic case as well as for the extreme relativistic case.

8.20. Express the integrals (8.5.19) in terms of the initial slope of the function $\Phi(x)$, and verify eqn. (8.5.20).

8.21. The total energy E of the electron cloud in an atom can be written as

$$E = K + V_{ne} + V_{ee},$$

where K is the kinetic energy of the electrons, V_{ne} the interaction energy between the electrons and the nucleus, and V_{ee} the mutual interaction energy of the electrons. Show that, according to the Thomas–Fermi model of a neutral atom,

$$K = -E, \quad V_{ne} = +\tfrac{7}{3}E \quad \text{and} \quad V_{ee} = -\tfrac{1}{3}E,$$

so that total $V = (V_{ne} + V_{ee}) = 2E$. Note that these results are consistent with the *virial theorem*; see Problem 3.20, with $n = -1$.

Notes

[1] We therefore speak of the totality of the energy levels filled at $T = 0$ as "the Fermi sea" and the small fraction of the particles that are excited near the top, when $T > 0$, as a "mist above the sea". Physically speaking, the origin of this behavior again lies in the Pauli exclusion principle, according to which a fermion of energy ε cannot absorb a quantum of thermal excitation ε_T if the energy level $\varepsilon + \varepsilon_T$ is already filled. Since $\varepsilon_T = O(kT)$, only those fermions which occupy energy levels near the top level ε_F, up to a depth $O(kT)$, can be thermally excited to go over to the available energy levels.

[2] See, for instance, Gol'dman *et al.* (1960); Problem 6.3.

[3] To justify the assumption that "the conduction electrons in a metal may be treated as 'free' electrons", it is necessary to ascribe to them an effective mass $m' \neq m$. This is an indirect way of accounting for the fact that the electrons in a metal are not really free; the ratio m'/m accordingly depends upon the structural details of the metal and, therefore, varies from metal to metal. In sodium, $m'/m \simeq 0.98$.

[4] Another way of expressing the electron density is to write: $N/V = f\rho/M$, where f is the valency of the metal, ρ its mass density and M the mass of an atom (ρ/M, thus, being the number density of the atoms).

[5] In the light of the numbers quoted here, one can readily see that the quantity $e^{(\mu-\varepsilon_z)/kT}$ in eqn. (8), being *at most* equal to $e^{(\mu_0-W)/kT} \equiv e^{-\phi/kT}$, is, at all temperatures of interest, much smaller than unity. This means that we are operating here in the (Maxwellian) tail of the Fermi–Dirac distribution and hence the approximation made in going from eqn. (8) to eqn. (9) is justified.

[6] In writing this condition, we have tacitly assumed that the momentum components p_x and p_y of the electron remain unchanged on the absorption of a photon.

CHAPTER 9

STATISTICAL MECHANICS OF INTERACTING SYSTEMS: THE METHOD OF CLUSTER EXPANSIONS

ALL the systems considered in the previous chapters were composed of, or could be regarded as composed of, *non-interacting* entities. Consequently, the results obtained, though of considerable intrinsic importance, may have limitations when applied to systems that actually exist in nature. For a real contact between the theory and experiment, one must take into account the interparticle interactions operating in the system. This can be done with the help of the formalism developed in Chapters 3–5 which, in principle, can be applied to an unlimited variety of physical systems and problems; in practice, however, one encounters in most cases serious difficulties of analysis. These difficulties are less stringent in the case of systems such as low-density gases, for which a corresponding non-interacting system can serve as an approximation. The mathematical expressions for the various physical quantities pertaining to such a system can be written in the form of *series expansions*, whose main terms denote the corresponding ideal-system results while the subsequent terms provide corrections arising from the interparticle interactions in the system. A systematic method of carrying out such expansions, in the case of real gases obeying classical statistics, was developed by Mayer and his collaborators (1937 onward) and is known as the method of *cluster expansions*. Its generalization, which applies to gases obeying quantum statistics, was initiated by Kahn and Uhlenbeck (1938) and was perfected by Lee and Yang (1959a, b; 1960a,b,c).

9.1. Cluster expansion for a classical gas

We start with a relatively simple physical system, namely a single-component, classical, monatomic gas whose potential energy is given by a sum of two-particle interactions u_{ij}. The Hamiltonian of the system is then given by

$$H = \sum_i \left(\frac{1}{2m}p_i^2\right) + \sum_{i<j} u_{ij} \quad (i, j = 1, 2, \ldots, N); \tag{1}$$

the summation in the second part goes over all the $N(N-1)/2$ pairs of particles in the system. In general, the potential u_{ij} is a function of the relative position vector $\boldsymbol{r}_{ij}(=\boldsymbol{r}_j-\boldsymbol{r}_i)$; however, if the two-body force is a central one, then the function u_{ij} depends only on the interparticle distance r_{ij}.

With the above Hamiltonian, the partition function of the system would be given by, see eqn. (3.5.5),

$$Q_N(V,T)=\frac{1}{N!h^{3N}}\int\exp\left\{-\beta\sum_i\left(\frac{1}{2m}p_i^2\right)-\beta\sum_{i<j}u_{ij}\right\}d^{3N}p\,d^{3N}r. \tag{2}$$

Integration over the momenta of the particles can be carried out straightforwardly, with the result

$$Q_N(V.T)=\frac{1}{N!\lambda^{3N}}\int\exp\left[-\beta\sum_{i<j}u_{ij}\right]d^{3N}r=\frac{1}{N!\lambda^{3N}}Z_N(V.T). \tag{3}$$

where $\lambda\{=h/(2\pi mkT)^{1/2}\}$ is the *mean thermal wavelength* of the particles, while the function $Z_N(V,T)$ stands for the integral over the space coordinates $\boldsymbol{r}_1,\boldsymbol{r}_2,\ldots,\boldsymbol{r}_N$:

$$Z_N(V,T)=\int\exp\left[-\beta\sum_{i<j}u_{ij}\right]d^{3N}r=\int\prod_{i<j}\left(e^{-\beta u_{ij}}\right)d^{3N}r. \tag{4}$$

The function $Z_N(V,T)$ is generally referred to as the *configuration integral* of the system. For a gas of non-interacting particles, the integrand in (4) is unity; we then have

$$Z_N^{(0)}(V,T)=V^N\quad\text{and}\quad Q_N^{(0)}(V,T)=\frac{V^N}{N!\lambda^{3N}}, \tag{5}$$

in agreement with our earlier result (3.5.9).

To treat the non-ideal case we introduce, after Mayer, the two-particle function f_{ij}, defined by the relationship

$$f_{ij}=e^{-\beta u_{ij}}-1. \tag{6}$$

In the absence of interactions, the function f_{ij} is identically zero; in the presence of interactions, it is nonzero but at sufficiently high temperatures it is quite small in comparison with unity. We therefore expect that the functions f_{ij} would be quite appropriate for carrying out a high-temperature expansion of the integrand in (4).

A typical plot of the functions u_{ij} and f_{ij} is shown in Fig. 9.1; we note that (i) the function f_{ij} is everywhere bounded and (ii) it becomes negligibly small as the interparticle distance r_{ij} becomes large in comparison with the "effective" range, r_0, of the potential.

Now, to evaluate the configuration integral (4), we expand its integrand in ascending powers of the functions f_{ij}:

$$\begin{aligned}Z_N(V,T)&=\int\prod_{i<j}\left(1+f_{ij}\right)d^3r_1\ldots d^3r_N\\&=\int\left[1+\sum f_{ij}+\sum f_{ij}f_{kl}+\cdots\right]d^3r_1\ldots d^3r_N.\end{aligned} \tag{7}$$

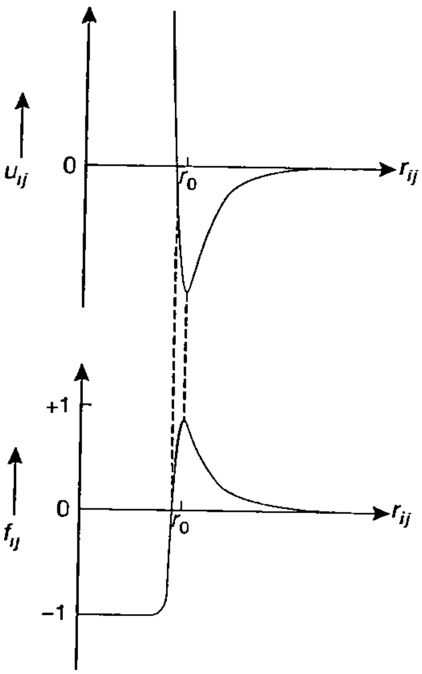
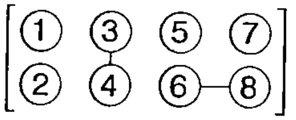
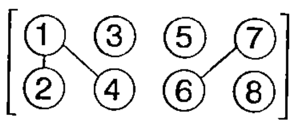


FIG. 9.1. A typical plot of the two-body potential function u_{ij} and the corresponding Mayer function f_{ij}.

A convenient way of enumerating the various terms in (7) is to associate each term with a corresponding N-particle graph. For instance, if N were 8, the terms

$$t_A = \int f_{34} f_{68}\, d^3r_1 \ldots d^3r_8 \quad \text{and} \quad t_B = \int f_{12} f_{14} f_{67}\, d^3r_1 \ldots d^3r_8 \tag{8}$$

in the expansion of the configuration integral Z_8 could be associated with the 8-particle graphs


and

, (9)

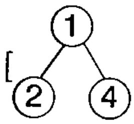
respectively. A closer look at the terms t_A and t_B (and at the corresponding graphs) suggests that we better regard these terms as suitably factorized (and the graphs correspondingly decomposed), that is,

$$t_A = \int d^3r_1 \int d^3r_2 \int d^3r_5 \int d^3r_7 \int f_{34}\, d^3r_3\, d^3r_4 \int f_{68}\, d^3r_6\, d^3r_8$$

$\equiv$ [①]·[②]·[⑤]·[⑦]·[③—④]·[⑥—⑧] (10)

and similarly

$$t_B = \int d^3r_3 \int d^3r_5 \int d^3r_8 \int f_{12} f_{14}\, d^3r_1\, d^3r_2\, d^3r_4 \int f_{67}\, d^3r_6\, d^3r_7$$

$\equiv$ [③]·[⑤]·[⑧]·[⑥—⑦]·[]. (11)

We may then say that the term t_A in the expansion of the integral Z_8 represents a “configuration” in which there are four “clusters’ one particle each and two “clusters” of two particles each, while the term t_B represents a “configuration” in which there are three "clusters" of one particle each, one "cluster" of two particles and one "cluster" of three particles.

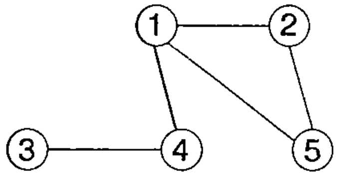
In view of this, we may introduce the notion of an *N-particle graph* which, by definition, is a "collection of N distinct circles, numbered $1, 2, \ldots, N$, with a number of lines linking some (or all) of the circles"; if the distinct pairs (of circles), which are linked through these lines, are denoted by the symbols $\alpha, \beta, \ldots, \lambda$ (each of these symbols denoting a distinct *pair* of indices out of the set $1, 2, \ldots, N$), then the graph represents the term

$$\int (f_\alpha f_\beta \cdots f_\lambda)\, d^3r_1 \ldots d^3r_N \tag{12}$$

of the expansion (7). A graph having the same number of linked pairs as this one but with the set $(\alpha', \beta', \ldots, \lambda')$ *distinct* from the set $(\alpha, \beta, \ldots, \lambda)$ will be counted as a distinct graph, for it represents a different term in the expansion; of course, these terms will belong to one and the same group in the expansion. Now, in view of the one-to-one correspondence between the various terms in the expansion (7) and the various N-particle graphs, we have

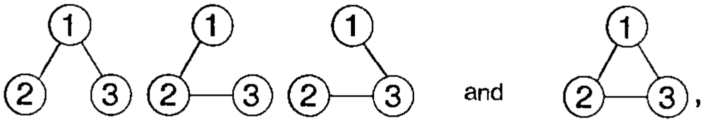
$$Z_N(V. T) = \text{sum of all distinct } N\text{-particle graphs.} \tag{13}$$

Further, in view of the possible factorization of the various terms (or the possible decomposition of the various graphs), we may introduce the notion of an *l-cluster* which, by definition, is an "l-particle graph in which each of the l circles, numbered $1, 2, \ldots, l$, is directly or indirectly linked with every other circle". As an example, we write here a 5-particle graph which is also a 5-cluster:



$$\equiv \int f_{12} f_{14} f_{15} f_{25} f_{34}\, d^3r_1 \ldots d^3r_5. \tag{14}$$

It is obvious that a cluster cannot be decomposed into simpler graphs inasmuch as the corresponding term cannot be factorized into simpler terms. Furthermore, a group of l particles (except when $l = 1$ or 2) can lead to a variety of l-clusters, some of which may be equal in value; for instance, a group of three particles leads to four different 3-clusters, namely



(15)

of which the first three are equal in value. In view of the variety of ways in which an l-cluster can appear, we may introduce the notion of a *cluster integral* b_l, defined by

$$b_l(V, T) = \frac{1}{l!\lambda^{3(l-1)}V} \times (\text{the sum of all possible } l\text{-clusters}). \tag{16}$$

So defined, the cluster integral $b_l(V, T)$ is dimensionless and, in the limit $V \to \infty$, approaches a finite value, $\bar{b}_l(T)$, which is independent of the size and the shape of the container less the latter is unduly queer). The first property is obvious. The second one follows by noting that if we hold one of the l particles fixed, at the point r_1 say, and carry out integration over the coordinates of the remaining $(l-1)$ particles, then, because of the fact that the functions f_{ij} extend only over a small finite range of distances, this integration would extend only over a limited region of the space available–a region whose *linear* dimensions are of the order of the range of the functions f_{ij};[1] the result of this integration will be practically independent of the volume of the container.[2] Finally, we integrate over the coordinates r_1 of the particle that was held fixed and obtain a straight factor of V; this cancels out the V in the denominator of the defining formula (16). Thus, the dependence of the cluster integral $b_l(V,T)$ on the size of the container is no more than a mere "surface effect"–an effect that disappears as $V \to \infty$, and we end up with a volume-independent number $\bar{b}_l(T)$.

Some of the simpler cluster integrals are

$$b_1 = \frac{1}{V}[\textcircled{1}] = \frac{1}{V}\int d^3r_1 \equiv 1, \tag{17}$$

$$\begin{aligned}
b_2 &= \frac{1}{2\lambda^3 V}[\textcircled{1}\!-\!\!-\!\textcircled{2}] = \frac{1}{2\lambda^3 V}\iint f_{12}\,d^3r_1\,d^3r_2 \\
&\approx \frac{1}{2\lambda^3}\int f_{12}\,d^3r_{12} = \frac{2\pi}{\lambda^3}\int\limits_0^\infty f(r)r^2\,dr \\
&= \frac{2\pi}{\lambda^3}\int\limits_0^\infty (e^{-u(r)/kT}-1)r^2\,dr,
\end{aligned} \tag{18}$$

$$\begin{aligned}
b_3 &= \frac{1}{6\lambda^6 V}\times[\text{sum of the clusters (15)}] \\
&= \frac{1}{6\lambda^6 V}\int(\underbrace{f_{12}f_{13}+f_{12}f_{23}+f_{13}f_{23}}_{}+f_{12}f_{13}f_{23})\,d^3r_1\,d^3r_2\,d^3r_3 \\
&\approx \frac{1}{6\lambda^6 V}\left[3V\iint f_{12}f_{13}\,d^3r_{12}\,d^3r_{13} + V\iint f_{12}f_{13}f_{23}\,d^3r_{12}\,d^3r_{13}\right] \\
&= 2b_2^2 + \frac{1}{6\lambda^6}\iint f_{12}f_{13}f_{23}\,d^3r_{12}\,d^3r_{13},
\end{aligned} \tag{19}$$

and so on.

We now proceed to evaluate the expression in (13). Obviously, an N-particle graph will consist of a number of clusters of which, say, m_1 are 1-clusters, m_2 are 2-clusters, m_3 are 3-clusters, and so on; the numbers $\{m_l\}$ must satisfy the condition

$$\sum_{l=1}^{N} l m_l = N, \qquad m_l = 0, 1, 2, \ldots, N. \tag{20}$$

However, a given set of numbers $\{m_l\}$ does not specify a unique, single graph; it represents a "collection of graphs" the sum total of which may be denoted by the symbol $S\{m_l\}$. We may then write

$$Z_N(V,T) = \sum_{\{m_l\}} S\{m_l\}, \tag{21}$$

where the primed summation $\sum'$ goes over all sets $\{m_l\}$ that conform to the restrictive condition (20). Equation (21) represents a systematic regrouping of the graphs, as opposed to the simple-minded grouping that first appeared in (7).

Our next task consists of evaluating the sum $S\{m_l\}$. To do this, we observe that the "collection of graphs" under the distribution set $\{m_l\}$ arises essentially from the following two causes:

(i) there are, in general, many different ways of *assigning* the N particles of the system to the $\sum_l m_l$ clusters, and

(ii) for any given assignment, there are, in general, many different ways of *forming* the various clusters, for even with a given group of l particles an l-cluster (if $l > 2$) can be formed in a number of different ways; see, for example, the four different ways of forming a 3-cluster with a given group of three particles, as listed in (15).

For cause (i), we obtain a straightforward factor of

$$\frac{N!}{(1!)^{m_1}(2!)^{m_2}\ldots} = \frac{N!}{\prod_l (l!)^{m_l}}. \tag{22}$$

Now, if cause (ii) were not there, i.e. if all l-clusters were unique in their formation, then the sum $S\{m_l\}$ would be given by the product of the combinatorial factor (22) with "the value of any one graph in the set-up, viz.

$$\prod_l (\text{the value of an } l\text{-cluster})^{m_l}\text{''}, \tag{23}$$

further corrected for the fact that any two arrangements which differ merely in the exchange of *all* the particles in one cluster with *all* the particles in another cluster of the same size, *must not* be counted as distinct, the corresponding correction factor being

$$\prod_l (1/m_l!). \tag{24}$$

A little reflection now shows that cause (ii) is completely and correctly taken care of if we replace the product of the expressions (23) and (24) by the expression[3]

$$\prod_l \left[(\text{the } \textit{sum} \text{ of the values of all possible } l\text{-clusters})^{m_l}/m_l!\right] \tag{25}$$

which, with the help of eqn. (16), may be written as

$$\prod_l \left[(b_l l!\, \lambda^{3(l-1)} V)^{m_l}/m_l!\right] \tag{26}$$

The sum $S\{m_l\}$ is now given by the product of the factor (22) and the expression (26). Substituting the result into (21), we obtain for the configuration integral

$$Z_N(V,T) = N!\,\lambda^{3N} \sum_{\{m_l\}}{}' \left[\prod_l \left\{\left(b_l \frac{V}{\lambda^3}\right)^{m_l} \frac{1}{m_l!}\right\}\right]. \tag{27}$$

Here, use has been made of the fact that

$$\prod_l (\lambda^{3l})^{m_l} = \lambda^{3\sum_l lm_l} = \lambda^{3N}; \tag{28}$$

see the restrictive condition (20). The partition function of the system now follows from eqns (3) and (27), with the result

$$Q_N(V,T) = {\sum_{\{m_l\}}}' \left[\prod_{l=1}^{N} \left\{ \left(b_l \frac{V}{\lambda^3} \right)^{m_l} \frac{1}{m_l!} \right\} \right]. \tag{29}$$

The evaluation of the primed sum in (29) is complicated by the restrictive condition (20) which must be obeyed by every set $\{m_l\}$. We therefore move over to the grand partition function of the system:

$$\mathcal{Q}(z, V, T) = \sum_{N=0}^{\infty} z^N Q_N(V, T). \tag{30}$$

Writing

$$z^N = z^{\sum_l lm_l} = \prod_l (z^l)^{m_l}, \tag{31}$$

substituting for $Q_N(V, T)$ from (29), and noting that a *restricted* summation over sets $\{m_l\}$, subject to the condition $\sum_l lm_l = N$, followed by a summation over all values of N (from $N = 0$ to $N = \infty$) is equivalent to an *unrestricted* summation over *all possible sets* $\{m_l\}$, we obtain

$$\begin{aligned}
\mathcal{Q}(z, V, T) &= \sum_{m_1, m_2, \ldots = 0}^{\infty} \left[\prod_{l=1}^{\infty} \left\{ \left(b_l z^l \frac{V}{\lambda^3} \right)^{m_l} \frac{1}{m_l!} \right\} \right] \\
&= \prod_{l=1}^{\infty} \left[\sum_{m_l=0}^{\infty} \left\{ \left(b_l z^l \frac{V}{\lambda^3} \right)^{m_l} \frac{1}{m_l!} \right\} \right] \\
&= \prod_{l=1}^{\infty} \left[\exp\left(b_l z^l \frac{V}{\lambda^3} \right) \right] = \exp\left[\sum_{l=1}^{\infty} b_l z^l \frac{V}{\lambda^3} \right]
\end{aligned} \tag{32}$$

and, hence,

$$\frac{1}{V} \ln \mathcal{Q} = \frac{1}{\lambda^3} \sum_{l=1}^{\infty} b_l z^l. \tag{33}$$

In the limit $V \to \infty$, we write

$$\frac{P}{kT} \equiv \operatorname*{Lim}_{V\to\infty} \left(\frac{1}{V} \ln \mathcal{Q} \right) = \frac{1}{\lambda^3} \sum_{l=1}^{\infty} \bar{b}_l z^l \tag{34}$$

and

$$\frac{N}{V} \equiv \operatorname*{Lim}_{V\to\infty} \left(\frac{z}{V} \frac{\partial \ln \mathcal{Q}}{\partial z} \right) = \frac{1}{\lambda^3} \sum_{l=1}^{\infty} l \bar{b}_l z^l. \tag{35}$$

Equations (34) and (35) constitute the famous *cluster expansions* of the Mayer–Ursell formalism. Eliminating the fugacity parameter z among these equations, we obtain the *equation of state* of the system.

9.2. Virial expansion of the equation of state

The approach developed in the preceding section leads to rigorous results only if we apply it to the gaseous phase alone. If we attempt to include in our study the phenomena of condensation, the critical point and the liquid phase, we encounter serious difficulties relating to (i) the limiting procedure involved in eqns (9.1.34) and (9.1.35), (ii) the convergence of the summations over l, and (iii) the volume dependence of the cluster integrals b_l. We therefore restrict our study to the gaseous phase alone. The equation of state may then be written in the form

$$\frac{Pv}{kT} = \sum_{l=1}^{\infty} a_l(T) \left(\frac{\lambda^3}{v}\right)^{l-1} \tag{1}$$

where $v(= V/N)$ denotes the volume per particle in the system. Expansion (1), which is supposed to have been obtained by eliminating z between eqns (9.1.34) and (9.1.35), is called the *virial expansion* of the system and the numbers $a_l(T)$ the *virial coefficients*.[4] To determine the relationship between the coefficients a_l and the cluster integrals $\bar{b}_l$, we invert eqn. (9.1.35) to obtain z as a power series in (λ^3/v) and substitute this into (9.1.34). This leads to eqn. (1), with

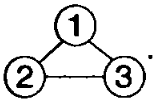
$$a_1 = \bar{b}_1 \equiv 1, \tag{2}$$

$$a_2 = -\bar{b}_2 = -\frac{2\pi}{\lambda^3} \int_0^{\infty} \left(e^{-u(r)/kT} - 1\right) r^2\, dr, \tag{3}$$

$$a_3 = 4b_2^2 - 2\bar{b}_3 = -\frac{1}{3\lambda^6} \int_0^{\infty}\int_0^{\infty} f_{12} f_{13} f_{23}\, d^3r_{12}\, d^3r_{13}, \tag{4}$$

$$a_4 = -20b_2^3 + 18\bar{b}_2\bar{b}_3 - 3\bar{b}_4 = \ldots \tag{5}$$

and so on; here, use has also been made of formulae (9.1.17–19). We note that the coefficient a_l is completely determined by the quantities $\bar{b}_1, \bar{b}_2, \ldots, \bar{b}_l$, i.e. by the sequence of configuration integrals $Z_1, Z_2, \ldots, Z_l$; see also eqns (9.4.5–8).

From eqn. (4) we observe that the third virial coefficient of the gas is determined solely by the 3-cluster . This suggests that the higher-order virial coefficients may also be determined solely by a special "subgroup" of the various l-clusters. This is indeed true, and the relevant result is that, in the limit of infinite volume[5],

$$a_l = -\frac{l-1}{l}\beta_{l-1} \quad (l \geq 2), \tag{6}$$

where β_{l-1} is the so-called *irreducible cluster integral*, defined as

$$\beta_{l-1} = \frac{1}{(l-1)!\,\lambda^{3(l-1)}V} \times \text{(the sum of all irreducible } l\text{-clusters)}; \tag{7}$$

by an *irreducible* l-cluster we mean an "l-particle graph which is multiply-connected (in the sense that there are at least two entirely independent, non-intersecting paths linking each pair of circles in the graph)". For instance, of the four possible 3-clusters, see (9.1.15), only the last one is irreducible. Indeed, if we express eqn. (4) in terms of this particular cluster and make use of the definition (7) for β_2, we obtain for the third virial coefficient

$$a_3 = -\tfrac{2}{3}\beta_2, \tag{8}$$

in agreement with the general result (6).[6]

The quantities β_{l-1}, like b_l, are dimensionless and, in the limit $V \to \infty$, approach finite values which are independent of the size and the shape of the container (unless the latter is unduly queer). Moreover, the two sets of quantities are mutually related; see eqns (9.4.27) and (9.4.29).

9.3. Evaluation of the virial coefficients

If a given system does not depart much from the ideal-gas behavior, its equation of state is given adequately by the first few virial coefficients. Now, since $a_1 \equiv 1$, the lowest-order virial coefficient that we need to consider here is a_2, which is given by eqn. (9.2.3):

$$a_2 = -b_2 = \frac{2\pi}{\lambda^3}\int_0^\infty \left(1 - e^{-u(r)/kT}\right) r^2\, dr, \tag{1}$$

$u(r)$ being the potential energy of interparticle interaction. A typical plot of the function $u(r)$ was shown in Fig. 9.1; a typical semi-empirical formula is given by (Lennard-Jones, 1924)

$$u(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{2}$$

It will be noted that the most significant features of an actual interparticle potential are well simulated by the Lennard-Jones formula (2). For instance, the function $u(r)$ given by (2) exhibits a "minimum", of value $-\varepsilon$, at a distance $r_0 (= 2^{1/6}\sigma)$ and rises to an infinitely large (positive) value for $r < \sigma$ and to a vanishingly small (negative) value for $r \gg \sigma$. The portion to the left of the "minimum" is dominated by *repulsive* interaction that comes into play when two particles come too close to one another, while the portion to the right is dominated by *attractive* interaction that operates between particles when they are separated by a respectable distance. For most practical purposes, the precise form of the repulsive part of the potential is not very important; it may as well be replaced by the crude approximation

$$u(r) = +\infty \quad (\text{for } r < r_0). \tag{3}$$

which amounts to attributing an *impenetrable* core, of diameter r_0, to each particle. The precise form of the attractive part is, however, important; in view of the fact that there exists good theoretical basis for the sixth-power attractive potential (see Problem 3.36), this part may simply be written as

$$u(r) = -u_0 (r_0/r)^6 \quad (r \geq r_0). \tag{4}$$

The potential given by expressions (3) and (4) may, therefore, be used if one is only interested in a qualitative assessment of the situation and not in a quantitative comparison between the theory and experiment.

Substituting (3) and (4) into (1), we obtain for the second virial coefficient

$$a_2 = \frac{2\pi}{\lambda^3}\left[\int\limits_0^{r_0} r^2\,dr + \int\limits_{r_0}^{\infty}\left[1 - \exp\left\{\frac{u_0}{kT}\left(\frac{r_0}{r}\right)^6\right\}\right] r^2\,dr\right] \tag{5}$$

The first integral is straightforward; the second one is considerably simplified if we assume that $(u_0/kT) \ll 1$, which makes the integrand very nearly equal to $-(u_0/kT)(r_0/r)^6$. Equation (5) then gives

$$a_2 \simeq \frac{2\pi r_0^3}{3\lambda^3}\left(1 - \frac{u_0}{kT}\right). \tag{6}$$

Substituting (6) into the expansion (9.2.1), we obtain a first-order improvement on the ideal-gas law, viz.

$$P \simeq \frac{kT}{v}\left\{1 + \frac{2\pi r_0^3}{3v}\left(1 - \frac{u_0}{kT}\right)\right\} \tag{7a}$$

$$= \frac{kT}{v}\left\{1 + \frac{B_2(T)}{v}\right\}, \text{ say.} \tag{7b}$$

The coefficient B_2, which is also sometimes referred to as the second virial coefficient of the system, is given by

$$B_2 \equiv a_2\lambda^3 \simeq \frac{2\pi r_0^3}{3}\left(1 - \frac{u_0}{kT}\right). \tag{8}$$

In our derivation it was explicitly assumed that (i) the potential function $u(r)$ is given by the simplified expressions (3) and (4), and (ii) $(u_0/kT) \ll 1$. We cannot, therefore, expect formula (8) to be a faithful representation of the second virial coefficient of a real gas. Nevertheless, it does correspond, almost exactly, to the *van der Waals approximation* to the equation of state of a real gas. This can be seen by rewriting (7a) in the form

$$\left(P + \frac{2\pi r_0^3 u_0}{3v^2}\right) \simeq \frac{kT}{v}\left(1 + \frac{2\pi r_0^3}{3v}\right) \simeq \frac{kT}{v}\left(1 - \frac{2\pi r_0^3}{3v}\right)^{-1}$$

which readily leads to the van der Waals equation of state

$$\left(P + \frac{a}{v^2}\right)(v - b) \simeq kT, \tag{9}$$

where

$$a = \frac{2\pi r_0^3 u_0}{3} \quad\text{and}\quad b = \frac{2\pi r_0^3}{3} \equiv 4v_0. \tag{10}$$

We note that the parameter b in the van der Waals equation of state is exactly four times the actual *molecular volume* v_0, the latter being the "volume of a sphere of diameter r_0"; cf. Problem 1.4. We also note that in this derivation we have assumed that $b \ll v$, which means that the gas is sufficiently dilute for the mean interparticle distance to be much larger than the effective range of the interparticle interaction. Finally, we observe that, according to this simple-minded calculation, the van der Waals parameters a and b are temperature-independent, which in reality is not true.

A realistic study of the second virial coefficient requires the use of a realistic potential, such as the one given by Lennard-Jones, for evaluating the integral in (1). This has indeed been done and the results obtained are shown in Fig. 9.2, where the reduced coefficient $B_2'(= B_2/r_0^3)$ is plotted against the reduced temperature $T'(= kT/\varepsilon)$:

$$B_2'(T') = 2\pi \int_0^\infty (1 - e^{-u'(r')/T'}) r'^2 \, dr', \tag{11}$$

with

$$u'(r') = \left\{ \left(\frac{1}{r'}\right)^{12} - 2\left(\frac{1}{r'}\right)^{6} \right\}, \tag{12}$$

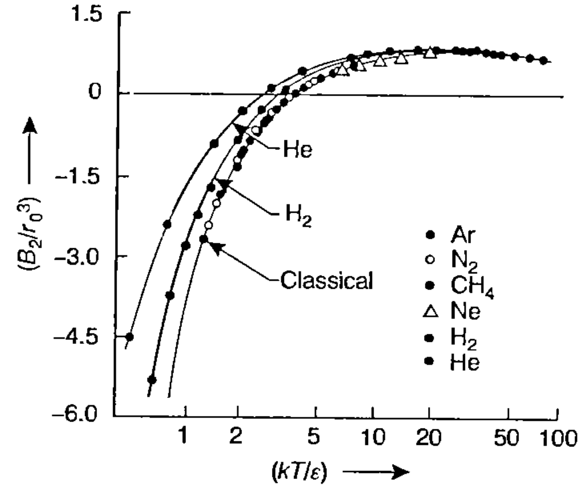


FIG. 9.2. A dimensionless plot showing the temperature dependence of the second virial coefficient of several gases (after Hirschfelder *et al.*, 1954).

r' being equal to (r/r_0); expressed in this form, the quantity B_2' is a *universal* function of T'. Included in the plot are experimental results for several gases. We note that in most cases the agreement is reasonably good; this is especially satisfying in view of the fact that in each case we had only two adjustable parameters, r_0 and ε, against a much larger number of experimental points available. In the first place, this agreement vindicates the adequacy of the Lennard-Jones potential for providing an analytical description of a typical interparticle potential. Secondly, it enables one to derive empirical values of the respective parameters of the potential; for instance, one obtains for argon: $r_0 = 3.82$ Å and $\varepsilon/k = 120$ K.[7] One cannot fail to observe that the lighter gases, hydrogen and helium, constitute exceptions to the rather general rule of agreement between theory and experiment. The reason for this lies in the fact that in the case of these gases quantum-mechanical effects assume considerable importance–more so at low temperatures. To substantiate this point, we have included in Fig. 9.2 theoretical curves for H_2 and He taking into account the quantum-mechanical effects as well; as a result, we find once again a fairly good agreement between the theory and experiment.

As regards higher-order virial coefficients ($l > 2$), we confine our discussion to a gas of hard spheres alone. We then have

$$\left.\begin{aligned} u(r) &= +\infty \quad \text{for} \quad r < r_0 \\ &= \;\; 0 \quad \text{for} \quad r > r_0 \end{aligned}\right\} \tag{13}$$

and, hence,

$$\left.\begin{aligned} f(r) &= -1 \quad \text{for} \quad r < r_0 \\ &= \;\; 0 \quad \text{for} \quad r > r_0 \end{aligned}\right\}. \tag{14}$$

The second virial coefficient of the gas is then given by

$$a_2 = \frac{2\pi r_0^3}{3\lambda^3} = 4\frac{v_0}{\lambda^3}; \tag{15}$$

cf. (6). The third virial coefficient can be determined with the help of eqn. (9.2.4), viz.

$$a_3 = -\frac{1}{3\lambda^6}\int_0^\infty\int_0^\infty f_{12}f_{13}f_{23}\,d^3r_{12}\,d^3r_{13}. \tag{16}$$

To evaluate this integral, we first fix the positions of particles 1 and 2 (such that $r_{12} < r_0$) and let particle 3 take all possible positions so that we can effect an integration over the variable r_{13}; see Fig. 9.3. Since our integrand is equal to -1 when each of the distances r_{13} and r_{23} (like r_{12}) is less than r_0 and is 0 otherwise, we have

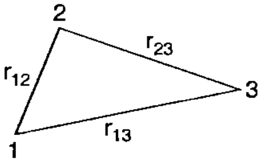


FIG. 9.3.

$$a_3 = \frac{1}{3\lambda^6}\int_{r_{12}=0}^{r_0}\left\{\int' d^3r_{13}\right\} d^3r_{12}, \tag{17}$$

where the primed integration arises from particle 3 taking all possible positions of interest. In view of the conditions $r_{13} < r_0$ and $r_{23} < r_0$, this integral is precisely equal to the "volume common to the spheres S_1 and S_2, each of radius r_0, centered at the *fixed* points 1 and 2"; see Fig. 9.4. This in turn can be obtained by calculating the "volume swept by the shaded area in the figure on going through a complete revolution about the line of centres". One gets:

$$\int' d^3r_{13} = \int_0^{\sqrt{[r_0^2-(r_{12}/2)^2]}} \{2(r_0^2 - y^2)^{1/2} - r_{12}\} 2\pi y\, dy. \tag{18}$$

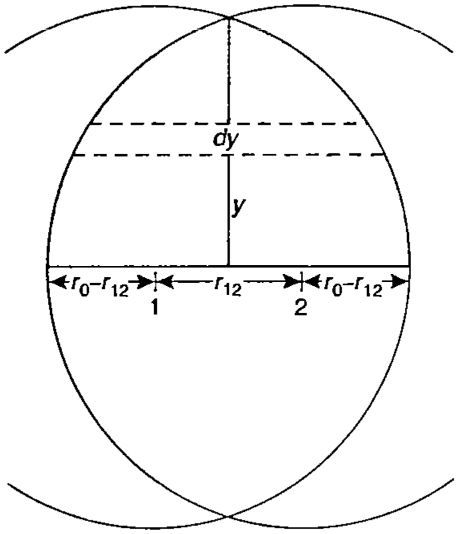


FIG. 9.4.

While the quantity within the curly brackets denotes the length of the strip shown in the figure, the element of area $2\pi y\, dy$ arises from the revolution; the limits of integration for y can be checked rather easily. The evaluation of the integral is straightforward; we obtain

$$\int' d^3r_{13} = \frac{4\pi}{3}\left\{r_0^3 - \frac{3r_0^2 r_{12}}{4} + \frac{r_{12}^3}{16}\right\}. \tag{19}$$

Substituting this into (17) and carrying out integration over r_{12}, we finally obtain

$$a_3 = \frac{5\pi^2 r_0^6}{18\lambda^6} = \frac{5}{8}a_2^2. \tag{20}$$

The fourth virial coefficient of the hard-sphere gas has also been evaluated exactly. It is given by (Boltzmann, 1899; Majumdar, 1929)[8]

$$\begin{aligned} a_4 &= \left\{\frac{1283}{8960} + \frac{3}{2}\cdot\frac{73\sqrt{(2)} + 1377\{\tan^{-1}\sqrt{(2)} - \pi/4\}}{1120\pi}\right\} a_2^3 \\ &= 0.28695 a_2^3. \end{aligned} \tag{21}$$

The fifth and sixth virial coefficients of this system have been computed numerically, with the results (Ree and Hoover, 1964)

$$a_5 = (0.1103 \pm 0.003) a_2^4 \tag{22}$$

and

$$a_6 = (0.0386 \pm 0.004) a_2^5. \tag{23}$$

Ree and Ho[illegible]r's estimate of the seventh virial coefficient is 0.0127 a_2^6, but the error in the evaluation of this coefficient has not been specified.

9.4. General remarks on cluster expansions

Shortly after the pioneering work of Mayer and his collaborators, Kahn and Uhlenbeck (1938) initiated the development of a similar treatment for quantum-mechanical systems. Of course, their treatment applied to the limiting case of classical systems as well but it faced certain inherent difficulties of analysis, some of which were later removed by the formal methods developed by Lee and Yang (1959a, b; 1960a,b,c). We propose to examine these developments in the remaining sections of this chapter. First, however, we would like to make a few general observations on the problem of cluster expansions. These observations, due primarily to Ono (1951) and Kilpatrick (1953), are of considerable interest insofar as they hold for a very large class of physical systems. For instance, the system may be quantum-mechanical or classical, it may be a multi-component one or single-component, its molecules may be polyatomic or monatomic, etc. All we have to assume is that (i) the system is gaseous in state and (ii) the partition functions $Q_N(V,T)$, for some low values of N, can somehow be obtained. We can then calculate the "cluster integrals" b_l, and the virial coefficients a_l, of the system in the following straightforward manner.

Quite generally, the *grand partition function* of the system can be written as

$$\mathcal{Q}(z,V,T) \equiv \sum_{N=0}^{\infty} Q_N(V,T) z^N = \sum_{N=0}^{\infty} \frac{Z_N(V,T)}{N!} \left(\frac{z}{\lambda^3}\right)^N \tag{1}$$

where we have introduced the "configuration integrals" $Z_N(V,T)$, defined in analogy with eqn. (9.1.3) of the classical treatment:

$$Z_N(V,T) \equiv N!\,\lambda^{3N} Q_N(V,T). \tag{2}$$

Dimensionally, the quantity Z_N is like (a volume)^N; moreover, the quantity Z_0 (like Q_0) is supposed to be identically equal to 1, while $Z_1(\equiv \lambda^3 Q_1)$ is identically equal to V. We then have, in the limit $V \to \infty$,

$$\frac{P}{kT} \equiv \frac{1}{V} \ln \mathcal{Q} = \frac{1}{V} \ln \left\{ 1 + \frac{Z_1}{1!}\left(\frac{z}{\lambda^3}\right)^1 + \frac{Z_2}{2!}\left(\frac{z}{\lambda^3}\right)^2 + \cdots \right\} \tag{3}$$

$$= \frac{1}{\lambda^3} \sum_{l=1}^{\infty} \mathfrak{b}_l z^l, \text{ say.} \tag{4}$$

Again, the last expression has been written in analogy with the classical expansion (9.1.34); the coefficients $\mathfrak{b}_l$ may, therefore, be looked upon as the *cluster integrals* of the given system. Expanding (3) as a power series in z and equating respective coefficients with the $\mathfrak{b}_l$ of (4), we obtain

$$\mathfrak{b}_1 = \frac{1}{V} Z_1 \equiv 1, \tag{5}$$

$$\mathfrak{b}_2 = \frac{1}{2!\,\lambda^3 V}(Z_2 - Z_1^2), \tag{6}$$

$$\bar{b}_3 = \frac{1}{3!\,\lambda^6 V}(Z_3 - 3Z_2Z_1 + 2Z_1^3), \tag{7}$$

$$\bar{b}_4 = \frac{1}{4!\,\lambda^9 V}(Z_4 - 4Z_3Z_1 - 3Z_2^2 + 12Z_2Z_1^2 - 6Z_1^4), \tag{8}$$

and so on. We note that, for all $l > 1$, the sum of the coefficients appearing within the parentheses is identically equal to zero. Consequently, in the case of an ideal classical gas, for which $Z_i \equiv V^i$, see eqn. (9.1.4), all cluster integrals with $l > 1$ vanish. This, in turn, implies the vanishing of all the virial coefficients of the gas (except, of course, a_1 which is identically equal to unity).

Comparing eqns (6)–(8) with eqn. (9.1.16), we find that the expressions involving the products of the various Z_i that appear within the parentheses play the same role here as "the sum of all possible l-clusters" does in the classical case. We therefore expect that, in the limit $V \to \infty$, the $\bar{b}_l$ here would also be independent of the size and the shape of the container (unless the latter is unduly queer). This, in turn, requires that the various combinations of the Z_i appearing within the parentheses here must all be proportional to the *first* power of V. This observation leads to the very interesting result, first noticed by Rushbrooke, namely

$$b_l = \frac{1}{l!\,\lambda^{3(l-1)}} \times (\text{the coefficient of } V^1 \text{ in the volume expansion of } Z_l). \tag{9}$$

At this stage, it seems worthwhile to point out that the expressions appearing within the parentheses of eqns (6)–(8) are well known in mathematical statistics as the *semi-invariants* of Thiele. The general formula for these expressions is

$$\begin{aligned}(\ldots)_l &\equiv \bar{b}_l\{l!\,\lambda^{3(l-1)}V\}\\ &= l!\sideset{}{'}\sum_{\{m_i\}}(-1)^{\sum_i m_i - 1}\left[\left(\sum_i m_i - 1\right)!\prod_i\left\{\frac{(Z_i/i!)^{m_i}}{m_i!}\right\}\right]\end{aligned} \tag{10}$$

where the primed summation goes over all sets $\{m_i\}$ that conform to the condition

$$\sum_{i=1}^{l} i m_i = l; \quad m_i = 0, 1, 2, \ldots \tag{11}$$

Relations inverse to (10) can be written down by referring to eqn. (9.1.29) of the classical treatment; thus

$$Z_M \equiv M!\,\lambda^{3M}Q_M = M!\,\lambda^{3M}\sideset{}{'}\sum_{\{m_l\}}\prod_{l=1}^{M}\left\{\frac{(V\bar{b}_l/\lambda^3)^{m_l}}{m_l!}\right\}, \tag{12}$$

where the primed summation goes over all sets $\{m_l\}$ that conform to the condition

$$\sum_{l=1}^{M} l m_l = M; \quad m_l = 0, 1, 2 \ldots \tag{13}$$

The calculation of the virial coefficients a_l now consists of a straightforward step that involves a use of formulae (5)–(8) in conjunction with formulae (9.2.2.–5). It appears, however, of interest to demonstrate here the manner in which the general relationship (9.2.6) between the virial coefficients a_l and the "irreducible cluster integrals" β_{l-1} arises mathematically. As a bonus, we will acquire yet another interpretation of the β_k.

Now, in view of the relations

$$\frac{P}{kT} \equiv \operatorname*{Lim}_{V\to\infty}\left(\frac{1}{V}\ln \mathscr{Q}\right) = \frac{1}{\lambda^3}\sum_{l=1}^{\infty} b_l z^l \tag{14}$$

and

$$\frac{1}{v} \equiv \operatorname*{Lim}_{V\to\infty}\left(\frac{z}{V}\frac{\partial \ln \mathscr{Q}}{\partial z}\right) = \frac{1}{\lambda^3}\sum_{l=1}^{\infty} l b_l z^l, \tag{15}$$

we can write

$$\frac{P(z)}{kT} = \int_0^z \frac{1}{v(z)}\frac{dz}{z}. \tag{16}$$

We introduce a new variable x, defined by

$$x = n\lambda^3 = \lambda^3/v. \tag{17}$$

In terms of this variable, eqn. (15) becomes

$$x(z) = \sum_{l=1}^{\infty} l b_l z^l, \tag{18}$$

the inverse of which may be written as (see Mayer and Harrison, 1938; also Kahn, 1938)

$$z(x) = x \exp\{-\phi(x)\}. \tag{19}$$

In view of the fact that, for $z \ll 1$, the variables z and x are practically the same, the function $\phi(x)$ must tend to zero as $x \to 0$; it may, therefore, be expressed as a power series in x:

$$\phi(x) = \sum_{k=1}^{\infty} \beta_k x^k. \tag{20}$$

It may be mentioned beforehand that the coefficients β_k of this expansion are ultimately going to be identified as the "irreducible cluster integrals" β_{l-1}. Substituting from eqns (17), (19) and (20) into eqn. (16), we get

$$\begin{aligned}\frac{P(x)}{kT} &= \int_0^x \frac{x}{\lambda^3}\left\{\frac{1}{x} - \phi'(x)\right\} dx = \frac{1}{\lambda^3}\left[x - \int_0^x \left\{\sum_{k=1}^{\infty} k\beta_k x^k\right\} dx\right] \\ &= \frac{x}{\lambda^3}\left[1 - \sum_{k=1}^{\infty}\left(\frac{k}{k+1}\beta_k x^k\right)\right].\end{aligned} \tag{21}$$

Combining (17) and (21), we obtain

$$\frac{Pv}{kT} = 1 - \sum_{k=1}^{\infty}\left(\frac{k}{k+1}\beta_k x^k\right). \tag{22}$$

Comparing this result with the virial expansion (9.2.1), we arrive at the desired relationship:

$$a_l = -\frac{l-1}{l}\beta_{l-1} \quad (l > 1). \tag{23}$$

In view of the generality of the present approach, the β_k appearing here may be regarded as a *generalization* of the irreducible cluster integrals of Mayer.

Finally, we would like to derive a relationship between the β_k and the $\bar{b}_l$. For this, we make use of a theorem due to Lagrange which, for the present purpose, states that "the solution $x(z)$ to the equation

$$z(x) = x/f(x) \tag{24}$$

is given by

$$x(z) = \sum_{j=1}^{\infty} \frac{z^j}{j!}\left[\frac{d^{j-1}}{d\xi^{j-1}}\{f(\xi)\}^j\right]_{\xi=0} \text{''} ; \tag{25}$$

it is obvious that the expression within the square brackets is $(j-1)!$ times "the coefficient of ξ^{j-1} in the Taylor expansion of the function $\{f(\xi)\}^j$ about the point $\xi = 0$". Applying this theorem to the function

$$f(x) = \exp\{\phi(x)\} = \exp\left\{\sum_{k=1}^{\infty}\beta_k x^k\right\} = \prod_{k=1}^{\infty}\exp(\beta_k x^k), \tag{26}$$

we obtain

$$x(z) = \sum_{j=1}^{\infty}\frac{z^j}{j!}(j-1)! \times \left\{\text{the coefficient of } \xi^{j-1} \text{ in the Taylor expansion} \right.$$
$$\left. \text{of } \prod_{k=1}^{\infty}\exp(j\beta_k\xi^k) \text{ about } \xi = 0\right\}.$$

Comparing this with eqn. (18), we get

$$\bar{b}_j = \frac{1}{j^2} \times \left\{\text{the coefficient of } \xi^{j-1} \text{ in } \prod_{k=1}^{\infty}\left[\sum_{m_k \geq 0}\frac{(j\beta_k)^{m_k}}{m_k!}\xi^{km_k}\right]\right\}$$
$$= \frac{1}{j^2}\sum_{\{m_k\}}{}' \prod_{k=1}^{j-1}\frac{(j\beta_k)^{m_k}}{m_k!}, \tag{27}$$

where the primed summation goes over all sets $\{m_k\}$ that conform to the condition

$$\sum_{k=1}^{j-1} km_k = j-1; \quad m_k = 0, 1, 2, \ldots \tag{28}$$

Formula (27) was first obtained by Mrs Mayer in 1937. Its inverse, however, was established much later (Mayer, 1942; Kilpatrick, 1953):

$$\beta_{l-1} = \sum_{\{m_i\}}{}' (-1)^{\sum_i m_i - 1}\frac{(l-2+\sum_i m_i)!}{(l-1)!}\prod_i \frac{(i\bar{b}_i)^{m_i}}{m_i!}, \tag{29}$$

where the primed summation goes over all sets $\{m_i\}$ that conform to the condition

$$\sum_{i=2}^{l}(i-1)m_i = l-1; \quad m_i = 0, 1, 2, \ldots \tag{30}$$

It is not difficult to see that the highest value of the index i in the set $\{m_i\}$ would be l (the corresponding set having all its m_i equal to 0, except m_l which would be equal to 1); accordingly, the highest order to which the quantities $\bar{b}_i$ would appear in the expression for β_{l-1} is that of $\bar{b}_l$. We thus see once again that the virial coefficient a_l is completely determined by the quantities $\bar{b}_1, \bar{b}_2, \ldots, \bar{b}_l$.

9.5. Exact treatment of the second virial coefficient

We now present a formulation, due originally to Beth and Uhlenbeck (1936, 1937), that enables us to make an exact calculation of the second virial coefficient of a quantum-mechanical system from a knowledge of the two-body potential $u(r)$.[9] In view of eqn. (9.4.6),

$$\bar{b}_2 = -a_2 = \frac{1}{2\lambda^3 V}(Z_2 - Z_1^2). \tag{1}$$

For the corresponding non-interacting system, one would have

$$\bar{b}_2^{(0)} = -a_2^{(0)} = \frac{1}{2\lambda^3 V}(Z_2^{(0)} - Z_1^{(0)2}); \tag{2}$$

the superscript (0) on the various symbols implies that they pertain to the non-interacting system. Combining (1) and (2), and remembering that $Z_1 = Z_1^{(0)} = V$, we obtain

$$\bar{b}_2 - \bar{b}_2^{(0)} = \frac{1}{2\lambda^3 V}(Z_2 - Z_2^{(0)}) \tag{3}$$

which, by virtue of the relation (9.4.2), becomes

$$\bar{b}_2 - \bar{b}_2^{(0)} = \frac{\lambda^3}{V}(Q_2 - Q_2^{(0)}) = \frac{\lambda^3}{V}\,\mathrm{Tr}\,(e^{-\beta\hat{H}_2} - e^{-\beta\hat{H}_2^{(0)}}). \tag{4}$$

For evaluating the trace in (4), we need to know the eigenvalues of the two-body Hamiltonian which, in turn, requires the solving of the Schrödinger equation[10]

$$\hat{H}_2\Psi_\alpha(r_1, r_2) = E_\alpha\Psi_\alpha(r_1, r_2), \tag{5}$$

where

$$\hat{H}_2 = -\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) + u(r_{12}). \tag{6}$$

Transforming to the *centre-of-mass* coordinates $R\{= \frac{1}{2}(r_1 + r_2)\}$ and the *relative* coordinates $r\{= (r_2 - r_1)\}$, we have

$$\Psi_\alpha(R, r) = \psi_j(R)\psi_n(r) = \left\{\frac{1}{V^{1/2}}e^{i(\mathbf{P}_j\cdot R)/\hbar}\right\}\psi_n(r), \tag{7}$$

with

$$E_\alpha = \frac{P_j^2}{2(2m)} + \varepsilon_n. \tag{8}$$

Here, $\boldsymbol{P}$ denotes the total momentum of the two particles and $2m$ their total mass, while ε denotes the energy associated with the relative motion of the particles; the symbol α refers to the set of quantum numbers j and n that determine the actual values of the variables $\boldsymbol{P}$ and ε. The wave equation for the relative motion will be

$$\left\{-\frac{\hbar^2}{2\left(\frac{1}{2}m\right)}\nabla_r^2 + u(r)\right\}\psi_n(\boldsymbol{r}) = \varepsilon_n\psi_n(\boldsymbol{r}), \tag{9}$$

$\frac{1}{2}m$ being the reduced mass of the particles; the normalization condition for the relative wave function will be

$$\int |\psi_n(\boldsymbol{r})|^2\, d^3r = 1. \tag{10}$$

Equation (4) thus becomes

$$\begin{aligned}\bar{b}_2 - \bar{b}_2^{(0)} &= \frac{\lambda^3}{V}\sum_\alpha\left\{e^{-\beta E_\alpha} - e^{-\beta E_\alpha^{(0)}}\right\}\\ &= \frac{\lambda^3}{V}\sum_j e^{-\beta P_j^2/4m}\sum_n\left\{e^{-\beta\varepsilon_n} - e^{-\beta\varepsilon_n^{(0)}}\right\}.\end{aligned} \tag{11}$$

For the first sum we obtain

$$\sum_j e^{-\beta P_j^2/4m} \approx \frac{4\pi V}{h^3}\int_0^\infty e^{-\beta P^2/4m}P^2\, dP = \frac{8^{1/2}V}{\lambda^3}, \tag{12}$$

so that eqn. (11) becomes

$$\bar{b}_2 - \bar{b}_2^{(0)} = 8^{1/2}\sum_n\left\{e^{-\beta\varepsilon_n} - e^{-\beta\varepsilon_n^{(0)}}\right\}. \tag{13}$$

The next step consists in examining the energy spectra, ε_n and $\varepsilon_n^{(0)}$, of the two systems. In the case of a non-interacting system, all we have is a "continuum"

$$\varepsilon_n^{(0)} = \frac{p^2}{2\left(\frac{1}{2}m\right)} = \frac{\hbar^2k^2}{m} \quad (k = p/\hbar), \tag{14}$$

with the standard density of states $g^{(0)}(k)$. In the case of an interacting system, we may have a set of *discrete eigenvalues* ε_B (that correspond to "bound" states), along with a "continuum"

$$\varepsilon_n = \frac{\hbar^2k^2}{m} \quad (k = p/\hbar), \tag{15}$$

with a characteristic density of states $g(k)$. Consequently, eqn. (13) can be written as

$$\bar{b}_2 - \bar{b}_2^{(0)} = 8^{1/2}\sum_B e^{-\beta\varepsilon_B} + 8^{1/2}\int_0^\infty e^{-\beta\hbar^2k^2/m}\{g(k) - g^{(0)}(k)\}\, dk, \tag{16}$$

where the summation in the first part goes over all bound states made possible by the two-body interaction.

The next thing to consider is the density of states $g(k)$. For this, we note that, since the two-body potential has been assumed to be central, the wave function $\psi_n(\boldsymbol{r})$ for the relative motion may be written as a product of a radial function $\chi(r)$ and a spherical harmonic $Y(\theta, \varphi)$:

$$\psi_{klm}(\boldsymbol{r}) = A_{klm}\frac{\chi_{kl}(r)}{r}Y_{l,m}(\theta, \varphi). \tag{17}$$

Moreover, the requirement of symmetry, namely $\psi(-\boldsymbol{r}) = \psi(\boldsymbol{r})$ for bosons and $\psi(-\boldsymbol{r}) = -\psi(\boldsymbol{r})$ for fermions, imposes the restriction that the quantum number l be even for bosons and odd for fermions. The (outer) boundary condition on the wave function may be written as

$$\chi_{kl}(R_0) = 0, \tag{18}$$

where R_0 is a fairly large value (of the variable r) that ultimately goes to infinity. The asymptotic form of the function $\chi_{kl}(r)$ is well known:

$$\chi_{kl}(r) \propto \sin\left\{kr - \frac{l\pi}{2} + \eta_l(k)\right\}; \tag{19}$$

accordingly, we must have

$$kR_0 - \frac{l\pi}{2} + \eta_l(k) = n\pi, \quad n = 0, 1, 2, \ldots \tag{20}$$

The symbol $\eta_l(k)$ here stands for the *scattering phase shift*, due to the two-body potential $u(r)$, for the lth partial wave of wave number k.

Equation (20) determines the full spectrum of the partial waves. To obtain an expression for the respective density of state $g_l(k)$, we observe that the wave number difference Δk between two consecutive states n and $n + 1$ is given by the formula

$$\left\{R_0 + \frac{d\eta_l(k)}{dk}\right\}\Delta k = \pi, \tag{21}$$

so that

$$g_l(k) = \frac{2l+1}{\Delta k} = \frac{2l+1}{\pi}\left\{R_0 + \frac{\partial\eta_l(k)}{\partial k}\right\}; \tag{22}$$

the factor $(2l + 1)$ has been included here to take account of the fact that each eigenvalue k pertaining to an lth partial wave is $(2l + 1)$-fold degenerate (because the magnetic quantum number m can take any of the values $l, (l-1), \ldots, -l$, without affecting the eigenvalue). The total density of states, $g(k)$, of *all* partial waves of wave numbers around the value k is then given by

$$g(k) = \sum_l{}' g_l(k) = \frac{1}{\pi}\sum_l{}'(2l+1)\left\{R_0 + \frac{\partial\eta_l(k)}{\partial k}\right\}; \tag{23}$$

note that the primed summation $\sum'$ goes over $l = 0, 2, 4, \ldots$ in the case of bosons and over $l = 1, 3, 5, \ldots$ in the case of fermions. For the corresponding non-interacting case, we have (since all $\eta_l(k) = 0$)

$$g^{(0)}(k) = \frac{R_0}{\pi}\sum_l{}'(2l+1). \tag{24}$$

Combining (23) and (24), we obtain

$$g(k) - g^{(0)}(k) = \frac{1}{\pi}\sum_{l}{}'(2l+1)\frac{\partial \eta_l(k)}{\partial k}. \tag{25}$$

Substituting (25) into (16), we obtain the desired result

$$\bar{b}_2 - \bar{b}_2^{(0)} = 8^{1/2}\sum_{B} e^{-\beta\varepsilon_B} + \frac{8^{1/2}}{\pi}\sum_{l}{}'(2l+1)\int_0^\infty e^{-\beta\hbar^2k^2/m}\frac{\partial \eta_l(k)}{\partial k}\,dk \tag{26}$$

which, in principle, is calculable for any given potential $u(r)$ through the respective phase shifts $\eta_l(k)$.

Equation (26) can be used for determining the quantity $\bar{b}_2 - \bar{b}_2^{(0)}$. In order to determine $\bar{b}_2$ itself, we must know the value of $\bar{b}_2^{(0)}$. This has already been obtained in Sec. 7.1 for bosons and in Sec. 8.1 for fermions; see eqns (7.1.13) and (8.1.17). Thus

$$\bar{b}_2^{(0)} = -a_2^{(0)} = \pm\frac{1}{2^{5/2}}, \tag{27}$$

where the upper sign holds for bosons and the lower sign for fermions. It may be noted that the foregoing result can be obtained directly from the relationship

$$\bar{b}_2^{(0)} = \frac{1}{2\lambda^3 V}(Z_2^{(0)} - Z_1^{(0)2}) = \frac{\lambda^3}{V}\left(Q_2^{(0)} - \frac{1}{2}Q_1^{(0)2}\right)$$

by substituting for $Q_2^{(0)}$ the exact expression (5.5.25):

$$\bar{b}_2^{(0)} = \frac{\lambda^3}{V}\left[\left\{\frac{1}{2}\left(\frac{V}{\lambda^3}\right)^2 \pm \frac{1}{2^{5/2}}\left(\frac{V}{\lambda^3}\right)^1\right\} - \frac{1}{2}\left(\frac{V}{\lambda^3}\right)^2\right] = \pm\frac{1}{2^{5/2}}. \tag*{(28)[11]}$$

It is of interest to note that this result can also be obtained by using the classical formula (9.1.18) and substituting for the two-body potential $u(r)$ the "statistical potential" (5.5.28); thus

$$\begin{aligned}\bar{b}_2^{(0)} &= \frac{2\pi}{\lambda^3}\int_0^\infty (e^{-u_s(r)/kT} - 1)r^2\,dr \\ &= \pm\frac{2\pi}{\lambda^3}\int_0^\infty e^{-2\pi r^2/\lambda^2}r^2\,dr = \pm\frac{1}{2^{5/2}}.\end{aligned} \tag{29}$$

As an illustration of this method, we shall calculate the second virial coefficient of a gas of hard spheres. The two-body potential in this case may be written as

$$\left.\begin{aligned}u(r) &= +\infty \quad \text{for} \quad r < r_0 \\ &= 0 \quad\quad \text{for} \quad r > r_0\end{aligned}\right\} \tag{30}$$

The scattering phase shifts $\eta_l(k)$ can now be determined by making use of the (inner) boundary condition, namely that $\chi(r) = 0$ for all $r < r_0$ and hence it vanishes as $r \to r_0$ from above. We thus obtain (see, for example, Schiff, 1968)

$$\eta_l(k) = \tan^{-1}\frac{j_l(kr_0)}{n_l(kr_0)}, \tag{31}$$

where $j_l(x)$ and $n_l(x)$ are, respectively, the "spherical Bessel functions" and the "spherical Neumann functions":

$$j_0(x) = \frac{\sin x}{x}, \quad j_1(x) = \frac{\sin x - x\cos x}{x^2}, \quad j_2(x) = \frac{(3-x^2)\sin x - 3x\cos x}{x^3}\ldots$$

and

$$n_0(x) = -\frac{\cos x}{x}, \quad n_1(x) = -\frac{\cos x + x\sin x}{x^2},$$

$$n_2(x) = -\frac{(3-x^2)\cos x + 3x\sin x}{x^3}, \ldots .$$

Accordingly,

$$\eta_0(k) = \tan^{-1}\{-\tan(kr_0)\} = -kr_0, \tag{32}$$

$$\eta_1(k) = \tan^{-1}\left\{-\frac{\tan(kr_0) - kr_0}{1 + kr_0\tan(kr_0)}\right\} = -\{kr_0 - \tan^{-1}(kr_0)\}$$

$$= -\frac{(kr_0)^3}{3} + \frac{(kr_0)^5}{5} - \cdots, \tag{33}$$

$$\eta_2(k) = \tan^{-1}\left\{-\frac{\tan(kr_0) - 3(kr_0)/[3-(kr_0)^2]}{1 + 3(kr_0)\tan(kr_0)/[3-(kr_0)^2]}\right\}$$

$$= -\left\{kr_0 - \tan^{-1}\frac{3(kr_0)}{3-(kr_0)^2}\right\} = -\frac{(kr_0)^5}{45} + \cdots, \tag{34}$$

and so on. We now have to substitute these results into formula (26). However, before doing that we should point out that, in the case of hard-sphere interaction, (i) we cannot have bound states at all and (ii) since, for all l, $\eta_l(0) = 0$, the integral in (26) can be simplified by a prior integration by parts. Thus, we have

$$\bar{b}_2 - \bar{b}_2^{(0)} = \frac{8^{1/2}\lambda^2}{\pi^2}\sum_l{}'(2l+1)\int_0^\infty e^{-\beta\hbar^2k^2/m}\eta_l(k)k\,dk. \tag{35}$$

Substituting for $l = 0$ and 2 in the case of bosons and for $l = 1$ in the case of fermions, we obtain (to fifth power in r_0/λ)

$$\bar{b}_2 - \bar{b}_2^{(0)} = -2\left(\frac{r_0}{\lambda}\right)^1 - \frac{10\pi^2}{3}\left(\frac{r_0}{\lambda}\right)^5 - \cdots \qquad \text{(Bose)} \tag{36}$$

$$= -6\pi\left(\frac{r_0}{\lambda}\right)^3 + 18\pi^2\left(\frac{r_0}{\lambda}\right)^5 - \cdots \qquad \text{(Fermi)}, \tag{37}$$

which may be compared with the classical result $-(2\pi/3)(r_0/\lambda)^3$.

9.6. Cluster expansion for a quantum-mechanical system

When it comes to calculating b_l for $l > 2$ we have no formula comparable in simplicity to formula (9.5.26) for b_2. This is due to the fact that we have no treatment of the l-body problem (for $l > 2$) which is as neat as the phase-shift analysis of the two-body problem. Nevertheless, a formal theory for the calculation of higher-order "cluster integrals" has been developed by Kahn and Uhlenbeck (1938); an elaboration by Lee and Yang (1959a, b; 1960a,b,c) has made this theory almost as good for treating a quantum-mechanical system as Mayer's theory has been for a classical gas. The basic approach in this theory is to evolve a scheme for expressing the grand partition function of the given system in essentially the same way as Mayer's cluster expansion does for a classical gas. However, because of the interplay of quantum-statistical effects and the effects arising from interparticle interactions, the mathematical structure of this theory is considerably involved.

We consider here a quantum-mechanical system of N identical particles enclosed in a box of volume V. The Hamiltonian of the system is assumed to be of the form

$$\hat{H}_N = -\frac{\hbar^2}{2m}\sum_{i=1}^{N}\nabla_i^2 + \sum_{i<j} u(r_{ij}). \tag{1}$$

Now, the partition function of the system is given by

$$\begin{aligned} Q_N(V,T) &\equiv \operatorname{Tr}(e^{-\beta\hat{H}_N}) = \sum_\alpha e^{-\beta E_\alpha} \\ &= \sum_\alpha \int_V \{\Psi_\alpha^*(1,\ldots,N)e^{-\beta\hat{H}_N}\Psi_\alpha(1,\ldots,N)\}\, d^{3N}r, \end{aligned} \tag{2}$$

where the functions Ψ_α are supposed to form a complete set of (properly symmetrized) orthonormal wave functions of the system, while the numbers $1,\ldots,N$ denote the position coordinates $r_1,\ldots,r_N$, respectively. We may as well introduce the *probability density operator* $\hat{W}_N$ of the system through the matrix elements

$$\begin{aligned} \langle 1',\ldots,N'|\hat{W}_N|1,\ldots,N\rangle &\equiv N!\,\lambda^{3N}\sum_\alpha\{\Psi_\alpha(1',\ldots,N')e^{-\beta\hat{H}_N}\Psi_\alpha^*(1,\ldots,N)\} \\ &= N!\,\lambda^{3N}\sum_\alpha\{\Psi_\alpha(1',\ldots,N')\Psi_\alpha^*(1,\ldots,N)\}e^{-\beta E_\alpha}. \end{aligned} \tag{3}$$

We denote the diagonal elements of the operator $\hat{W}_N$ by the symbols $W_N(1,\ldots,N)$; thus

$$W_N(1,\ldots,N) = N!\,\lambda^{3N}\sum_\alpha\{\Psi_\alpha(1,\ldots,N)\Psi_\alpha^*(1,\ldots,N)\}e^{-\beta E_\alpha}, \tag{4}$$

whence eqn. (2) takes the form

$$Q_N(V,T) = \frac{1}{N!\,\lambda^{3N}}\int_V W_N(1,\ldots,N)\,d^{3N}r = \frac{1}{N!\,\lambda^{3N}}\operatorname{Tr}(\hat{W}_N). \tag{5}$$

A comparison of eqn. (5) with eqns (9.1.3) and (9.4.2) shows that the "trace of the probability density operator $\hat{W}_N$" is the analogue of the "configuration integral" Z_N, and the quantity $W_N(1, \ldots, N)\, d^{3N}r$ is a measure of the probability that the "configuration" of the given system is found to be within the interval $[(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N), (\boldsymbol{r}_1 + d\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N + d\boldsymbol{r}_N)]$.

Before we proceed further, let us acquaint ourselves with some of the basic properties of the matrix elements (3):

(i)
$$\begin{aligned}\langle 1' | \hat{W}_1 | 1 \rangle &= \lambda^3 \sum_{p} \left\{ \frac{1}{\sqrt{V}} e^{i(\boldsymbol{p}\cdot\boldsymbol{r}_1')/\hbar} \frac{1}{\sqrt{V}} e^{-i(\boldsymbol{p}\cdot\boldsymbol{r}_1)/\hbar} \right\} e^{-\beta p^2/2m} \\ &\simeq \frac{\lambda^3}{V} \int\!\!\int\!\!\int_{\infty}^{+\infty} \frac{V d^3p}{h^3} e^{\{i\boldsymbol{p}\cdot(\boldsymbol{r}_1' - \boldsymbol{r}_1)/\hbar - \beta p^2/2m\}} \\ &= e^{-\pi|\boldsymbol{r}_1' - \boldsymbol{r}_1|^2/\lambda^2};\end{aligned} \tag{6}$$

cf. eqn. (5.3.14) for the density matrix of a single particle. The foregoing result is a manifestation of the quantum-mechanical, *not quantum-statistical*, correlation between the positions $\boldsymbol{r}$ and $\boldsymbol{r}'$ of a given particle (or, for that matter, any particle in the system). This correlation extends over distances of the order of λ which is, therefore, a measure of the linear dimensions of the wave packet representing the particle. As $T \to \infty$, and hence $\lambda \to 0$, the matrix element (6) tends to zero for all finite values of $|\boldsymbol{r}_1' - \boldsymbol{r}_1|$.

(ii)
$$\langle 1 | \hat{W}_1 | 1 \rangle = 1; \tag{7}$$

consequently, by eqn. (5),

$$Q_1(V, T) = \frac{1}{\lambda^3} \int_V 1\, d^3r = \frac{V}{\lambda^3}. \tag{8}$$

(iii) Whatever the symmetry character of the wave functions Ψ, the diagonal elements $W_N(1, \ldots, N)$ of the probability density operator $\hat{W}_N$ are *symmetric* in respect of a permutation among the arguments $(1, \ldots, N)$.

(iv) The elements $W_N(1, \ldots, N)$ are *invariant* under a unitary transformation of the set $\{\Psi_\alpha\}$.

(v) Suppose that the coordinates $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N$ are such that they can be divided into two groups, A and B, with the property that *any* two coordinates, say $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$, of which one belongs to group A and the other to group B, satisfy the conditions that

  (a) the separation r_{ij} is much larger than the mean thermal wavelength λ of the particles, and

  (b) it is also much larger than the effective range r_0 of the two-body potential, then

$$W_N(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \simeq W_A(\boldsymbol{r}_A) W_B(\boldsymbol{r}_B), \tag{9}$$

where $\boldsymbol{r}_A$ and $\boldsymbol{r}_B$ denote *collectively* the coordinates in group A and group B, respectively. It is not easy to furnish here a rigorous mathematical proof of this property, though physically it is quite understandable. One can see this by noting that, in view of conditions (a) and (b), there does not exist any spatial correlation between the particles of group A on one hand and the particles of group B on the other (either by virtue of statistics or by virtue of interparticle interactions). The two groups, therefore, behave towards each other like two *independent* entities. It is then natural that, to a very good approximation, the probability density W_N of the composite configuration be equal to the product of the probability densities W_A and W_B.

We now proceed with the formulation. First of all, to fix ideas about the approach to be followed, we may consider the case with $N = 2$. In that case, as $r_{12} \to \infty$, we expect, in view of property (v) above, that

$$W_2(1, 2) \to W_1(1)W_1(2) = 1. \tag{10}$$

In general, however, $W_2(1, 2)$ will be different from $W_1(1)W_1(2)$. Now, if we denote the difference between $W_2(1, 2)$ and $W_1(1)W_1(2)$ by the symbol $U_2(1, 2)$, then, as $r_{12} \to \infty$,

$$U_2(1, 2) \to 0. \tag{11}$$

It is not difficult to see that the quantity $U_2(1, 2)$ is the quantum-mechanical analogue of the Mayer function f_{ij}. With this in mind, we introduce a sequence of *cluster functions* $\hat{U}_l$ defined by the hierarchy[12]

$$\langle 1'|\hat{W}_1|1\rangle = \langle 1'|\hat{U}_1|1\rangle, \tag{12}$$

$$\langle 1', 2'|\hat{W}_2|1, 2\rangle = \langle 1'|\hat{U}_1|1\rangle\langle 2'|\hat{U}_1|2\rangle + \langle 1', 2'|\hat{U}_2|1, 2\rangle, \tag{13}$$

$$\begin{aligned}\langle 1', 2', 3', |\hat{W}_3|1, 2, 3\rangle = {} & \langle 1'|\hat{U}_1|1\rangle\langle 2'|\hat{U}_1|2\rangle\langle 3'|\hat{U}_1|3\rangle \\ & + \langle 1'|\hat{U}_1|1\rangle\langle 2', 3'|\hat{U}_2|2, 3\rangle \\ & + \langle 2'|\hat{U}_1|2\rangle\langle 1', 3'|\hat{U}_2|1, 3\rangle \\ & + \langle 3'|\hat{U}_1|3\rangle\langle 1', 2'|\hat{U}_2|1, 2\rangle \\ & + \langle 1', 2', 3'|\hat{U}_3|1, 2, 3\rangle,\end{aligned} \tag{14}$$

and so on. A particular function $\hat{U}_l$ is thus defined with the help of the first l equations of the hierarchy. The last equation in the hierarchy will be (writing only the diagonal elements)

$$W_N(1, \ldots, N) = \sum_{\{m_l\}}{}' \left\{ \sum_P \underbrace{[U_1() \ldots U_1()]}_{m_1 \text{ factors}} \underbrace{[U_2() \ldots U_2()]}_{m_2 \text{ factors}} \ldots \right\}, \tag{15}$$

where the primed summation goes over all sets $\{m_l\}$ that conform to the condition

$$\sum_{l=1}^{N} l m_l = N; \quad m_l = 0, 1, 2, \ldots \tag{16}$$

Moreover, in selecting the arguments of the various U_l appearing in (15), out of the numbers $1, \ldots, N$, one has to remember that a permutation of the arguments within the same bracket is *not* regarded as leading to anything distinctly different from what one h[illegible]efore the permutation; the symbol $\sum_P$ then denotes a summation over all *distinct* ways of selecting the arguments under the set $\{m_l\}$.

Relations inverse to the ones written above are easy to obtain. One gets

$$\langle 1'|\hat{U}_1|1\rangle = \langle 1'|\hat{W}_1|1\rangle, \tag{17}$$

$$\langle 1',2'|\hat{U}_2|1,2\rangle = \langle 1',2'|\hat{W}_2|1,2\rangle - \langle 1'|\hat{W}_1|1\rangle\langle 2'|\hat{W}_1|2\rangle, \tag{18}$$

$$\begin{aligned}\langle 1',2',3'|\hat{U}_3|1,2,3\rangle = {} & \langle 1',2',3'|\hat{W}_3|1,2,3\rangle \\ & - \langle 1'|\hat{W}_1|1\rangle\langle 2',3'|\hat{W}_2|2,3\rangle \\ & - \langle 2'|\hat{W}_1|2\rangle\langle 1',3'|\hat{W}_2|1,3\rangle \\ & - \langle 3'|\hat{W}_1|3\rangle\langle 1',2'|\hat{W}_2|1,2\rangle \\ & + 2\langle 1'|\hat{W}_1|1\rangle\langle 2'|\hat{W}_1|2\rangle\langle 3'|\hat{W}_1|3\rangle, \end{aligned} \tag{19}$$

and so on; compare the right-hand sides of these equations with the expressions appearing within the parentheses in eqns (9.4.5–7). We note that (i) the coefficient of a general term here is

$$(-1)^{\sum_l m_l - 1}\left(\sum_l m_l - 1\right)!, \tag{20}$$

where $\sum_l m_l$ is the number of the W_n in the term, and (ii) the sum of the coefficients of all the terms on the right-hand side of eqns (18), (19), ... is identically zero. Moreover, the diagonal elements $U_l(1,\ldots,l)$, just like the diagonal elements of the operators $\hat{W}_n$, are symmetric in respect of permutations among the arguments $(1,\ldots,l)$, and are determined by the sequence of the diagonal elements $W_1, W_2, \ldots, W_l$. Finally, in view of the property (v) of the W_n, as embodied in formula (9), the U_l possess the following property:

$$U_l(1,\ldots,l) \simeq 0 \quad \text{if} \quad r_{ij} \gg \lambda, r_0; \tag{21}$$

here, r_{ij} is the separation between any two of the coordinates $(1,\ldots,l)$.[13]

We now define the "cluster integral" b_l by the formula

$$b_l(V,T) = \frac{1}{l!\,\lambda^{3(l-1)}V}\int U_l(1,\ldots,l)\,d^{3l}r; \tag{22}$$

cf. eqn. (9.1.16). Clearly, the quantity $b_l(V,T)$ is dimensionless and, by virtue of property (21) of the elements $U_l(1,\ldots,l)$, is practically independent of V (so long as V is large). In the limit $V \to \infty$, $b_l(V,T)$ tends to a finite volume-independent value, which may be denoted by $\bar{b}_l(T)$. We then obtain for the partition function of the system, see eqns (5) and (15),

$$Q_N(V,T) = \frac{1}{N!\,\lambda^{3N}}\int d^{3N}r\left\{\sum_{\{m_l\}}{}^{'}\left[\sum_P [U_1\ldots U_1][U_2\ldots U_2]\ldots\right]\right\} \tag{23}$$

$$\begin{aligned} &= \frac{1}{N!\,\lambda^{3N}}\sum_{\{m_l\}}{}^{'}\frac{N!}{(1!)^{m_1}(2!)^{m_2}\ldots m_1!m_2!\ldots}\times \\ &\qquad \int d^{3N}r\{[U_1\ldots U_1][U_2\ldots U_2]\ldots\}. \end{aligned} \tag{24}$$

In writing the last result we have made use of the fact that, since a permutation among the arguments of the functions U_l does not affect the value of the integral concerned, the summation over P may be replaced by any one term of the summation, multiplied by the number of *distinct* permutations allowed by the set $\{m_l\}$; cf. the corresponding product of the numbers (9.1.22) and (9.1.24). Making use of the definition (22), eqn. (24) can be written as

$$Q_N(V,T) = \frac{1}{\lambda^{3N}} \sum_{\{m_l\}}{}' \left[\prod_{l=1}^{N} \{ b_l \lambda^{3(l-1)} V)^{m_l} / m_l! \} \right]$$

$$= \sum_{\{m_l\}}{}' \left[\prod_{l=1}^{N} \left\{ \left(b_l \frac{V}{\lambda^3} \right)^{m_l} \frac{1}{m_l!} \right\} \right]; \tag{25}$$

again, use has been made of the fact that

$$\prod_l (\lambda^{3l})^{m_l} = \lambda^{3 \sum_l l m_l} = \lambda^{3N}. \tag{26}$$

Equation (25) is formally identical with eqn. (9.1.29) of Mayer's theory. The subsequent development of the formalism, leading to the equation of state of the system, will also be formally identical with that theory. Thus, we finally obtain

$$\frac{P}{kT} = \frac{1}{\lambda^3} \sum_{l=1}^{\infty} b_l z^l \quad \text{and} \quad \frac{1}{v} = \frac{1}{\lambda^3} \sum_{l=1}^{\infty} l b_l z^l. \tag{27}$$

There are, however, important physical differences. We may recall that the calculation of the cluster integrals b_l in the classical case involved the evaluation of a number of finite, $3l$-dimensional integrals. The corresponding calculation in the quantum-mechanical case requires a knowledge of the functions U_l and hence of all W_n, with $n \leq l$; this in turn requires solutions of the n-body Schrödinger equation for all $n \leq l$. The case $l = 2$ can be handled neatly, as was done in Sec. 9.5. For $l > 2$, the mathematical procedure is rather cumbersome. Nevertheless, Lee and Yang (1959a, b, 1960a,b,c) have evolved a scheme that enables us to calculate the higher b_l in successive approximations. According to this scheme, the functions U_l of a given system may be evaluated by "separating out" the effects of statistics from those of interparticle interactions, i.e. we may first take care of the statistical aspect of the problem and then tackle the dynamical aspect of it. Thus, the whole feat is accomplished in two steps. First, the U-functions pertaining to the given system are expressed in terms of U-functions pertaining to a corresponding *quantum-mechanical system obeying Boltzmann statistics*, i.e. a (fictitious) system described by *unsymmetrized wave functions*. This step takes care of the statistics of the given system, i.e. of the symmetry properties of the wave functions describing the system. Next, the U-functions of the (fictitious) Boltzmannian system are expanded, loosely speaking, in powers of a *binary kernel* B (which is obtainable from a solution of the two-body problem with the given interaction). A commendable feature of the method is that it can be applied even if the given interaction contains a singular, repulsive cor[illegible]e. even if the potential energy for certain configurations of the system becomes infinitely large. Though the method is admirably systematic and fairly straightforward in principle, its application to real systems is quite complicated. We therefore turn to a more practical method–the method of quantized fields–which has been highly useful in the study of quantum-mechanical systems composed of interacting particles. For an exposition of the (binary collision) method of Lee and Yang, see Secs 9.7 and 9.8 of the first edition of this book.

In passing, we note yet another important difference between the quantum-mechanical case and the classical case. In the latter, if interparticle interactions are absent, then all $\bar{b}_l$, with $l \geq 2$, vanish. This is not true in the quantum-mechanical case; here, see Secs 7.1 and 8.1,

$$\bar{b}_l^{(0)} = (\pm 1)^{l-1} l^{-5/2}, \tag{28}$$

of which eqn. (9.5.27) was a special case.

Problems

9.1. For imperfect-gas calculations one sometimes employs the *Sutherland potential*

$$\begin{aligned} u(r) &= \infty && \text{for} \quad r < r_0 \\ &= -\varepsilon(\sigma/r)^6 && \text{for} \quad r > r_0. \end{aligned}$$

Using this potential, determine the second virial coefficient of a classical gas. Also determine first-order corrections to the gas law and to the various thermodynamic properties of the system.

9.2. According to Lennard-Jones, the physical behavior of most real gases can be well understood if the intermolecular potential is assumed to be of the form

$$u(r) = \frac{A}{r^m} - \frac{B}{r^n},$$

where n is very nearly equal to 6 while m ranges between 11 and 13. Determine the second virial coefficient of a Lennard-Jones gas and compare your result with that for a van der Waals gas; see eqn. (9.3.8).

9.3. (a) Show that for a gas obeying van der Waals' equation of state (9.3.9)

$$C_P - C_V = Nk\left\{1 - \frac{2a}{kTv^3}(v-b)^2\right\}^{-1}$$

(b) Also show that, for a van der Waals gas with *constant* specific heat C_V, an adiabatic process conforms to the equation

$$(v-b)T^{C_V/Nk} = \text{const};$$

cf. eqn. (1.4.30).

(c) Further show that the temperature change resulting from an expansion of the gas (into vacuum) from volume V_1 to volume V_2 is given by

$$T_2 - T_1 = \frac{N^2 a}{C_V}\left(\frac{1}{V_2} - \frac{1}{V_1}\right).$$

9.4. The coefficient of volume expansion α and the isothermal bulk modulus B of a gas are given by the empirical expressions

$$\alpha = \frac{1}{T}\left(1 + \frac{3a'}{vT^2}\right) \quad \text{and} \quad B = P\left(1 + \frac{a'}{vT^2}\right)^{-1}$$

where a' is a constant parameter. Show that these expressions are mutually compatible. Also derive the equation of state of this gas.

9.5. Show that the first-order Joule–Thomson coefficient of a [illegible] is given by the formula

$$\left(\frac{\partial T}{\partial P}\right)_H = \frac{N}{C_P}\left(T\frac{\partial(a_2\lambda^3)}{\partial T} - a_2\lambda^3\right),$$

where $a_2(T)$ is the second virial coefficient of the gas and H its enthalpy; see eqn. (9.2.1). Derive an explicit expression for the Joule–Thomson coefficient in the case of a gas with interparticle interaction

$$u(r) = \begin{cases} +\infty & \text{for} \quad 0 < r < r_0, \\ -u_0 & \text{for} \quad r_0 < r < r_1, \\ 0 & \text{for} \quad r_1 < r < \infty, \end{cases}$$

and discuss the temperature dependence of this coefficient.

9.6. Assume that the molecules of the nitrogen gas interact through the potential of the previous problem. Making use of the experimental data given below, determine the "best" empirical values for the parameters r_0, r_1 and u_0/k:

T (in K)	100	200	300	400	500
$a_2\lambda^3$ (in K per atm)	-1.80	-4.26×10^{-1}	-5.49×10^{-2}	$+1.12 \times 10^{-1}$	$+2.05 \times 10^{-1}$.

9.7. Determine the lowest-order corrections to the *ideal*-gas values of the Helmholtz free energy, the Gibbs free energy, the entropy, the internal energy, the enthalpy and the (constant-volume and constant-pressure) specific heats for a real gas. Discuss the temperature dependence of these corrections in the case of a gas whose molecules interact through the potential of Problem 9.5.

9.8. The molecules of a solid attract one another with a force $F(r) = \alpha(l/r)^5$. Two semi-infinite solids composed of n molecules per unit volume are separated by a distance d, i.e. the solids fill the whole of the space with $x \le 0$ and $x \ge d$. Calculate the force of attraction, per unit area of the surface, between the two solids.

9.9. Referring to eqn. (9.5.31) for the phase shifts $\eta_l(k)$ of a hard-sphere gas, show that for $kr_0 \ll 1$

$$\eta_l(k) \simeq -\frac{(kr_0)^{2l+1}}{(2l+1)\{1.3\ldots(2l-1)\}^2}.$$

9.10. Using the wave functions

$$u_p(r) = \frac{1}{\sqrt{V}}e^{i(p\cdot r)/\hbar}$$

to describe the motion of a free particle, write down the *symmetrized* wave functions for a pair of non-interacting bosons/fermions, and show that

$$\langle 1', 2'|\hat{U}_2^{S/A}|1, 2\rangle = \pm\langle 2'|\hat{W}_1|1\rangle\langle 1'|\hat{W}_1|2\rangle.$$

9.11. Show that for a gas composed of particles with spin J

$$\bar{b}_2^S(J) = (J+1)(2J+1)\bar{b}_2^S(0) + J(2J+1)\bar{b}_2^A(0)$$

and

$$\bar{b}_2^A(J) = J(2J+1)\bar{b}_2^S(0) + (J+1)(2J+1)\bar{b}_2^A(0).$$

9.12. Show that the coefficient $\bar{b}_2$ for a quantum-mechanical *Boltzmannian* gas composed of "spin-less" particles satisfies the following relations:

$$\begin{aligned} \bar{b}_2 &= \lim_{J\to\infty}\left\{\frac{1}{(2J+1)^2}\bar{b}_2^S(J)\right\} = \lim_{J\to\infty}\left\{\frac{1}{(2J+1)^2}\bar{b}_2^A(J)\right\} \\ &= \tfrac{1}{2}\{\bar{b}_2^S(0) + \bar{b}_2^A(0)\}. \end{aligned}$$

Obtain the value of $\bar{b}_2$, to fifth order in (r_0/λ), by using the Beth–Uhlenbeck expressions (9.5.36, 37), and compare your result with the classical value of $\bar{b}_2$, namely $-(2\pi/3)(r_0/\lambda)^3$.

Notes

[1] Hence the name "cluster".

[2] Of course, some dependence on the geometry of the container *will* arise if the fixed particle happens to be close to the walls of the container. This is, however, unimportant when $V \to \infty$.

[3] To appreciate the logic of this replacement, consider expression [ ] in (25) as a multinomial expansion and interpret the various terms of this expansion in terms of the variety of the l-clusters.

[4] For various manipulations with the virial equation of state, see Kilpatrick and Ford (1969).

[5] For proof, see Hill (1956), Secs 24 and 25; see also Sec. 9.4 of the text.

[6] It may be mentioned here that a 2-cluster is also regarded as an *irreducible* cluster. Accordingly, $\beta_1 = 2b_2$; see eqns (9.1.16) and (9.2.7). Equation (3) then gives: $a_2 = -b_2 = -\frac{1}{2}\beta_1$, again in agreement with the general result (6).

[7] Corresponding values for various other gases have been summarized in Hill (1960), p. 484.

[8] See also Katsura (1959).

[9] For a discussion of the third virial coefficient, see Pais and Uhlenbeck (1959).

[10] For simplicity, we assume the particles to be "spinless". For the influence of spin, see Problems 9.11 and 9.12.

[11] This calculation incidentally verifies the general formula (9.4.9) for the case $l = 2$. By that formula, the "cluster integral" $\bar{b}_2$ of a given system would be equal to $1/(2\lambda^3)$ times the coefficient of V^1 in the volume expansion of the "configuration integral" Z_2 of the system. In the case under study, this coefficient is $\pm\lambda^3/2^{3/2}$; hence the result.

[12] The functions U_l were first introduced by Ursell, in 1927, in order to simplify the classical configuration integral. Their introduction into the quantum-mechanical treatment is due to Kahn and Uhlenbeck (1938).

[13] This can be seen by examining the break-up of the structure on the right-hand side of any equation in the hierarchy (18, 19, ...) when one or more of the l coordinates in the "cluster" get sufficiently separated from the rest.

CHAPTER 10

STATISTICAL MECHANICS OF INTERACTING SYSTEMS: THE METHOD OF QUANTIZED FIELDS

IN THIS chapter we present another method of dealing with systems composed of interacting particles. This method is based on the concept of a *quantized field* which is characterized by the field operators $\psi(\boldsymbol{r})$, and their hermitian conjugates $\psi^{\dagger}(\boldsymbol{r})$, that satisfy a set of well-defined commutation rules. In terms of these operators, one defines a number operator $\hat{N}$ and a Hamiltonian operator $\hat{H}$ that provide a suitable representation for a system composed of any finite number of particles and possessing any finite amount of energy. In view of its formal similarity with the Schrödinger formulation, the formulation in terms of a quantized field is generally referred to as the *second quantization* of the system. For convenience of calculation, the field operators $\psi(\boldsymbol{r})$ and $\psi^{\dagger}(\boldsymbol{r})$ are often expressed as superpositions of a set of single-particle wave functions $\{u_{\alpha}(\boldsymbol{r})\}$, with coefficients a_{α} and $a_{\alpha}^{\dagger}$; the latter turn out to be the *annihilation* and *creation* operators which again satisfy a set of well-defined commutation rules. The operators $\hat{N}$ and $\hat{H}$ then find a convenient expression in terms of the operators a_{α} and $a_{\alpha}^{\dagger}$, and the final formulation is well suited for a treatment based on operator algebra; as a result, many calculations, which would otherwise be tedious, can be carried out in a more or less straightforward manner.

10.1. The formalism of second quantization

To represent a system of particles by a *quantized field*, we invoke the field operators $\psi(\boldsymbol{r})$ and $\psi^{\dagger}(\boldsymbol{r})$, which are defined for all values of the position coordinate $\boldsymbol{r}$ and which operate upon a *Hilbert space*; a vector in this space corresponds to a particular state of the quantized field. The values of the quantities ψ and $\psi^{\dagger}$, at all $\boldsymbol{r}$, represent the *degrees of freedom* of the field; since $\boldsymbol{r}$ is a continuous variable, the number of these degrees of freedom is innumerably infinite. Now, if the given system is composed of bosons, the field operators $\psi(\boldsymbol{r})$ and $\psi^{\dagger}(\boldsymbol{r})$ are supposed to satisfy the commutation rules

$$[\psi(\boldsymbol{r}), \psi^{\dagger}(\boldsymbol{r}')] = \delta(\boldsymbol{r} - \boldsymbol{r}') \tag{1a}$$

$$[\psi(\boldsymbol{r}), \psi(\boldsymbol{r}')] = [\psi^{\dagger}(\boldsymbol{r}), \psi^{\dagger}(\boldsymbol{r}')] = 0, \tag{1b}$$

where the symbol $[A, B]$ stands for the commutator $(AB - BA)$ of the given operators A and B. If, on the other hand, the given system is composed of fermions, then the field operators are supposed to satisfy the rules

$$\{\psi(\boldsymbol{r}), \psi^{\dagger}(\boldsymbol{r}')\} = \delta(\boldsymbol{r} - \boldsymbol{r}') \tag{2a}$$

$$\{\psi(\boldsymbol{r}), \psi(\boldsymbol{r}')\} = \{\psi^{\dagger}(\boldsymbol{r}), \psi^{\dagger}(\boldsymbol{r}')\} = 0, \tag{2b}$$

where the symbol $\{A, B\}$ stands for the anticommutator $(AB + BA)$ of the given operators A and B. In the case of fermions, the operators $\psi(\boldsymbol{r})$ and $\psi^{\dagger}(\boldsymbol{r})$ possess certain explicit properties which follow directly from (2b), namely

$$\psi(\boldsymbol{r})\psi(\boldsymbol{r}') = -\psi(\boldsymbol{r}')\psi(\boldsymbol{r}), \quad \therefore \psi(\boldsymbol{r})\psi(\boldsymbol{r}) = 0 \quad \text{for all } \boldsymbol{r}; \tag{2c}$$

similarly,

$$\psi^{\dagger}(\boldsymbol{r})\psi^{\dagger}(\boldsymbol{r}') = -\psi^{\dagger}(\boldsymbol{r}')\psi^{\dagger}(\boldsymbol{r}), \quad \therefore \psi^{\dagger}(\boldsymbol{r})\psi^{\dagger}(\boldsymbol{r}) = 0 \quad \text{for all } \boldsymbol{r}. \tag{2d}$$

Clearly, no such property holds for the field operators pertaining to bosons. In the sequel we shall see that the mathematical difference between the commutation rules (1) for the boson field operators and rules (2) for the fermion field operators is intimately related to the fundamental difference in the symmetry properties of the respective wave functions in the Schrödinger formulation. Of course, in their own place, both sets of rules, (1) and (2), are essentially axiomatic.

We now introduce two hermitian operators, the *particle-number operator* $\hat{N}$ and the *Hamiltonian operator* $\hat{H}$, through definitions that hold for bosons as well as fermions:

$$\hat{N} \equiv \int d^3r\, \psi^{\dagger}(\boldsymbol{r})\psi(\boldsymbol{r}) \tag{3}$$

and

$$\hat{H} \equiv -\frac{\hbar^2}{2m}\int d^3r\, \psi^{\dagger}(\boldsymbol{r})\nabla^2\psi(\boldsymbol{r}) + \frac{1}{2}\iint d^3r_1 d^3r_2\, \psi^{\dagger}(\boldsymbol{r}_1)\,\psi^{\dagger}(\boldsymbol{r}_2)\,u(\boldsymbol{r}_1, \boldsymbol{r}_2)\,\psi(\boldsymbol{r}_2)\psi(\boldsymbol{r}_1), \tag{4}$$

where $u(\boldsymbol{r}_1, \boldsymbol{r}_2)$ denotes the two-body interaction potential in the given system. It is quite natural to interpret the product $\psi^{\dagger}(\boldsymbol{r})\psi(\boldsymbol{r})$ as the *number density operator* of the field. The similarity between the foregoing definitions and the expressions for the expectation values of the corresponding physical quantities in the Schrödinger formulation is fairly obvious. However, the similarity is only "formal" because, while there we are concerned with the wave functions of the given system (which are *c-numbers*), here we are concerned with the operators of the corresponding matter field. We can easily verify that, irrespective of the commutation rules obeyed by the operators $\psi(\boldsymbol{r})$ and $\psi^{\dagger}(\boldsymbol{r})$, the operators $\hat{N}$ and $\hat{H}$ commute:

$$[\hat{N}, \hat{H}] = 0; \tag{5}$$

accordingly, the operators $\hat{N}$ and $\hat{H}$ can be diagonalized simultaneously.

We now choose a *complete orthonormal basis* of the Hilbert space, such that any vector $|\Psi_{n}\rangle$ among the basis is a simultaneous eigenstate of the operators

$\hat{N}$ and $\hat{H}$. We may, therefore, denote any particular member of the basis by the symbol $|\Psi_{NE}\rangle$, with the properties

$$\hat{N}|\Psi_{NE}\rangle = N|\Psi_{NE}\rangle, \quad \hat{H}|\Psi_{NE}\rangle = E|\Psi_{NE} > \tag{6}$$

and

$$\langle\Psi_{NE}|\Psi_{NE}\rangle = 1. \tag{7}$$

The vector $|\Psi_{00}\rangle$, which represents the *vacuum state* of the field and is generally denoted by the symbol $|0\rangle$, is assumed to be unique; it possesses the obvious properties

$$\hat{N}|0\rangle = \hat{H}|0\rangle = 0 \quad \text{and} \quad \langle 0|0\rangle = 1. \tag{8}$$

Next we observe that, regardless of whether we employ the boson commutation rules (1) or the fermion rules (2), the operator $\hat{N}$ and the operators $\psi(\boldsymbol{r})$ and $\psi^{\dagger}(\boldsymbol{r})$ satisfy the commutation properties

$$[\psi(\boldsymbol{r}), \hat{N}] = \psi(\boldsymbol{r}) \quad \text{and} \quad [\psi^{\dagger}(\boldsymbol{r}), \hat{N}] = -\psi^{\dagger}(\boldsymbol{r}), \tag{9}$$

whence it follows that

$$\hat{N}\psi(\boldsymbol{r})|\Psi_{NE}\rangle = \left(\psi(\boldsymbol{r})\hat{N} - \psi(\boldsymbol{r})\right)|\Psi_{NE}\rangle = (N-1)\psi(\boldsymbol{r})|\Psi_{NE}\rangle \tag{10}$$

and

$$\hat{N}\psi^{\dagger}(\boldsymbol{r})|\Psi_{NE}\rangle = \left(\psi^{\dagger}(\boldsymbol{r})\hat{N} + \psi^{\dagger}(\boldsymbol{r})\right)|\Psi_{NE}\rangle = (N+1)\psi^{\dagger}(\boldsymbol{r})|\Psi_{NE}\rangle. \tag{11}$$

Clearly, the state $\psi(\boldsymbol{r})|\Psi_{NE}\rangle$ is also an eigenstate of the operator $\hat{N}$, but with eigenvalue $(N-1)$; thus, the application of the operator $\psi(\boldsymbol{r})$ on to the state $|\Psi_{NE}\rangle$ of the field *annihilates* one particle from the field. Similarly, the state $\psi^{\dagger}(\boldsymbol{r})|\Psi_{NE}\rangle$ is an eigenstate of the operator $\hat{N}$, with eigenvalue $(N+1)$; thus, the application of the operator $\psi^{\dagger}(\boldsymbol{r})$ on to the state $|\Psi_{NE}\rangle$ of the field *creates* a particle in the field. In each case, the process (of annihilation or creation) is tied down to the point $\boldsymbol{r}$ of the field; however, the energy associated with the process, which also means the change in the energy of the field, remains undetermined; see eqns (18) and (19). By a repeated application of the operator $\psi^{\dagger}$ on to the vacuum state $|0\rangle$, we find that the eigenvalues of the operator $\hat{N}$ are $0, 1, 2, \ldots$. On the other hand, the application of the operator ψ on to the vacuum state $|0\rangle$ gives nothing but zero because, for obvious reasons, we cannot admit negative eigenvalues for the operator $\hat{N}$. Of course, if we apply the operator ψ on to the state $|\Psi_{NE}\rangle$ repeatedly N times, we end up with the vacuum state; we then have, by virtue of the orthonormality of the basis chosen,

$$\langle\Phi_n|\psi(\boldsymbol{r}_1)\psi(\boldsymbol{r}_2)\ldots\psi(\boldsymbol{r}_N)|\Psi_{NE}\rangle = 0 \tag{12}$$

unless the state $|\Phi_n\rangle$ is itself the vacuum state, in which case we would obtain a nonzero result instead. In terms of this latter result, we may define a function of the N coordinates $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N$, namely

$$\Psi_{NE}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = (N!)^{-1/2}\langle 0|\psi(\boldsymbol{r}_1)\ldots\psi(\boldsymbol{r}_N|\Psi_{NE}\rangle \tag{13}$$

Obviously, the function $\Psi_{NE}(r_1, \dots, r_N)$ has something to do with "an assemblage of N particles located at the points $r_1, \dots, r_N$ of the field" because their annihilation from those points of the field has led us to the vacuum state of the field. To obtain the precise meaning of this function, we first note that in the case of bosons (fermions) this function is symmetric (antisymmetric) with respect to an interchange of any two of the N coordinates; see eqns (1b) and (2b), respectively. Secondly, its norm is equal to unity, which can be seen as follows.

By the very definition of $\Psi_{NE}(r_1, \dots, r_N)$,

$$\begin{aligned}
&\int d^{3N}r\Psi^*_{NE}(r_1, \dots, r_N)\Psi_{NE}(r_1, \dots, r_N)\\
&\quad = (N!)^{-1} \int d^{3N}r\, \langle\Psi_{NE}|\psi^\dagger(r_N)\dots\psi^\dagger(r_1)|0\rangle\langle 0|\psi(r_1)\dots\psi(r_N)|\Psi_{NE}\rangle\\
&\quad = (N!)^{-1} \int d^{3N}r \textstyle\sum_n \langle\Psi_{NE}|\psi^\dagger(r_N)\dots\psi^\dagger(r_1)|\Phi_n\rangle\langle\Phi_n|\psi(r_1)\dots\psi(r_N)|\Psi_{NE}\rangle\\
&\quad = (N!)^{-1} \int d^{3N}r\, \langle\Psi_{NE}|\psi^\dagger(r_N)\dots\psi^\dagger(r_2)\psi^\dagger(r_1)\psi(r_1)\psi(r_2)\dots\psi(r_N)|\Psi_{NE}\rangle;
\end{aligned}$$

here, use has been made of eqn. (12) which holds for all $|\Phi_n\rangle$ except for the vacuum state, and of the fact that the summation of $|\Phi_n\rangle\langle\Phi_n|$ over the complete orthonormal set of the basis chosen is equivalent to a unit operator. We now carry out integration over r_1, yielding the factor

$$\int d^3r_1\psi^\dagger(r_1)\psi(r_1) = \hat{N}.$$

Next, we carry out integration over r_2, yielding the factor

$$\int d^3r_2\, \psi^\dagger(r_2)\hat{N}\psi(r_2) = \int d^3r_2\, \psi^\dagger(r_2)\psi(r_2)(\hat{N}-1) = \hat{N}(\hat{N}-1);$$

see eqn. (10). By iteration, we obtain

$$\begin{aligned}
&\int d^{3N}r\, \Psi^*_{NE}(r_1, \dots, r_N)\Psi_{NE}(r_1, \dots, r_N)\\
&\quad = (N!)^{-1}\langle\Psi_{NE}|\hat{N}(\hat{N}-1)(\hat{N}-2)\dots \text{ up to } N \text{ factors}|\Psi_{NE}\rangle\\
&\quad = (N!)^{-1}N!\langle\Psi_{NE}|\Psi_{NE}\rangle = 1.
\end{aligned} \tag{14}$$

Finally, we can show that, for bosons as well as fermions, the function $\Psi_{NE}(r_1, \dots, r_N)$ satisfies the differential equation, see Problem 10.1,

$$\left(-\frac{\hbar^2}{2m}\sum_{i=1}^{N}\nabla_i^2 + \sum_{i<j}u_{ij}\right)\Psi_{NE}(r_1, \dots, r_N) = E\Psi_{NE}(r_1, \dots, r_N), \tag{15}$$

which is simply the *Schrödinger equation* of an N-particle system. The function $\Psi_{NE}(r_1, \dots, r_N)$ is, therefore, the Schrödinger wave function of the system, with energy eigenvalue E; accordingly, the product $\Psi^*_{NE}\Psi_{NE}$ is the probability density for the particles of the system to be in the vicinity of the coordinates $(r_1, \dots, r_N)$, when the system happens to be in an eigenstate with energy E. This establishes the desired correspondence between the quantized field formulation and the Schrödinger formulation. In passing, we place on record the quantized-field expression for the function $\Psi^*_{NE}(r_1, \dots, r_N)$, which is the complex conjugate of the wave function $\Psi_{NE}(r_1, \dots, r_N)$, viz.

$$\Psi^*_{NE}(r_1, \dots, r_N) = (N!)^{-1/2}\langle\Psi_{NE}|\psi^\dagger(r_N)\dots\psi^\dagger(r_1)|0\rangle. \tag{16}$$

We now introduce a complete orthonormal set of single-particle wave functions $u_\alpha(\boldsymbol{r})$, where the suffix α provides a label for identifying the various single-particle states; it could, for instance, be the energy eigenvalue of the state (or the momentum $\boldsymbol{p}$, along with the spin component σ pertaining to the state). In view of the orthonormality of these wave functions,

$$\int d^3r\, u_\alpha^*(\boldsymbol{r}) u_\beta(\boldsymbol{r}) = \delta_{\alpha\beta}. \tag{17}$$

The field operators $\psi(\boldsymbol{r})$ and $\psi^\dagger(\boldsymbol{r})$ may now be expanded in terms of the functions $u_\alpha(\boldsymbol{r})$:

$$\psi(\boldsymbol{r}) = \sum_\alpha a_\alpha u_\alpha(\boldsymbol{r}) \tag{18}$$

and

$$\psi^\dagger(\boldsymbol{r}) = \sum_\alpha a_\alpha^\dagger u_\alpha^*(\boldsymbol{r}). \tag{19}$$

Relations inverse to (18) and (19) are

$$a_\alpha = \int d^3r\, \psi(\boldsymbol{r}) u_\alpha^*(\boldsymbol{r}) \tag{20}$$

and

$$a_\alpha^\dagger = \int d^3r\, \psi^\dagger(\boldsymbol{r}) u_\alpha(\boldsymbol{r}). \tag{21}$$

The coefficients a_α and $a_\alpha^\dagger$, like the field variables $\psi(\boldsymbol{r})$ and $\psi^\dagger(\boldsymbol{r})$, are operators which operate upon the elements of the relevant Hilbert space. Indeed, the operators a_α and $a_\alpha^\dagger$ now take over the role of the *degrees of freedom* of the field.

Substituting (18) and (19) into the set of rules (1) or (2), and making use of the closure property of the u_α, viz.

$$\sum_\alpha u_\alpha(\boldsymbol{r}) u_\alpha^*(\boldsymbol{r}') = \delta(\boldsymbol{r} - \boldsymbol{r}'), \tag{22}$$

we obtain[1] for the operators a_α and $a_\alpha^\dagger$

$$[a_\alpha, a_\beta^\dagger] = \delta_{\alpha\beta} \tag{23a}$$

$$[a_\alpha, a_\beta] = [a_\alpha^\dagger, a_\beta^\dagger] = 0 \tag{23b}$$

in the case of *bosons*, and

$$\{a_\alpha, a_\beta^\dagger\} = \delta_{\alpha\beta} \tag{24a}$$

$$\{a_\alpha, a_\beta\} = \{a_\alpha^\dagger, a_\beta^\dagger\} = 0 \tag{24b}$$

in the case of *fermions*. In the latter case, the operators a_α and $a_\alpha^\dagger$ possess certain explicit properties which follow directly from (24b), namely

$$a_\alpha a_\beta = -a_\beta a_\alpha, \quad \therefore a_\alpha a_\alpha = 0 \quad \text{for all } \alpha; \tag{24c}$$

similarly

$$a_\alpha^\dagger a_\beta^\dagger = -a_\beta^\dagger a_\alpha^\dagger, \quad \therefore a_\alpha^\dagger a_\alpha^\dagger = 0 \quad \text{for all } \alpha. \tag{24d}$$

No such property holds for operators pertaining to bosons. We will see very shortly that this vital difference between the commutation rules for the boson operators and those for the fermion operators is closely linked with the fact that while fermions have to conform to the restrictions imposed by the Pauli exclusion principle, there are no such restrictions for bosons.

We now proceed to express operators $\hat{N}$ and $\hat{H}$ in terms of a_α and $a_\alpha^\dagger$. Substituting (18) and (19) into (3), we obtain

$$\begin{aligned}\hat{N} &= \int d^3r \sum_{\alpha,\beta} a_\alpha^\dagger a_\beta u_\alpha^*(\boldsymbol{r}) u_\beta(\boldsymbol{r}) = \sum_{\alpha,\beta} a_\alpha^\dagger a_\beta \delta_{\alpha\beta} \\ &= \sum_\alpha a_\alpha^\dagger a_\alpha. \end{aligned} \tag{25}$$

It seems natural to speak of the operator $a_\alpha^\dagger a_\alpha$ as the *particle-number operator* pertaining to the single-particle state α. We denote this operator by the symbol $\hat{N}_\alpha$:

$$\hat{N}_\alpha = a_\alpha^\dagger a_\alpha. \tag{26}$$

It is easy to verify that, for bosons as well as fermions, the operators $\hat{N}_\alpha$ commute with one another; hence, they can be simultaneously diagonalized. Accordingly, we may choose a complete orthonormal basis of the Hilbert space in such a way that any vector belonging to the basis is a simultaneous eigenstate of all the operators $\hat{N}_\alpha$.[2] Let a particular member of the basis be denoted by the vector $|n_0, n_1, \ldots, n_\alpha, \ldots\rangle$, or by the shorter symbol $|\Phi_n\rangle$, with the properties

$$\hat{N}_\alpha |\Phi_n\rangle = n_\alpha |\Phi_n\rangle \tag{27}$$

and

$$\langle \Phi_n | \Phi_n \rangle = 1: \tag{28}$$

the number n_α, being the eigenvalue of the operator $\hat{N}_\alpha$ in the state $|\Phi_n\rangle$ of the field, denotes the number of particles in the single-particle state α of the given system. One particular member of the basis, for which $n_\alpha = 0$ for all α, will represent the *vacuum state* of the field; denoting the vacuum state by the symbol $|\Phi_0\rangle$, we have

$$\hat{N}_\alpha |\Phi_0\rangle = 0 \quad \text{for all } \alpha, \quad \text{and} \quad \langle \Phi_0 | \Phi_0 \rangle = 1. \tag{29}$$

Next we observe that, regardless of whether we employ the boson commutation rules (23) or the fermion rules (24), the operator $\hat{N}_\alpha$ and the operators a_α and $a_\alpha^\dagger$ satisfy the commutation properties

$$[a_\alpha, \hat{N}_\alpha] = a_\alpha \quad \text{and} \quad [a_\alpha^\dagger, \hat{N}_\alpha] = -a_\alpha^\dagger, \tag{30}$$

whence it follows that

$$\hat{N}_\alpha a_\alpha |\Phi_n\rangle = (a_\alpha \hat{N}_\alpha - a_\alpha)|\Phi_n\rangle = (n_\alpha - 1) a_\alpha |\Phi_n\rangle \tag{31}$$

and

$$\hat{N}_\alpha a_\alpha^\dagger |\Phi_n\rangle = (a_\alpha^\dagger \hat{N}_\alpha + a_\alpha^\dagger)|\Phi_n\rangle = (n_\alpha + 1) a_\alpha^\dagger |\Phi_n\rangle. \tag{32}$$

Clearly, the state $a_\alpha|\Phi_n\rangle$ is also an eigenstate of the operator $\hat{N}_\alpha$, but with eigenvalue $(n_\alpha - 1)$; thus, the application of the operator a_α on to the state $|\Phi_n\rangle$ of the field *annihilates* one particle from the field. Similarly, the state $a_\alpha^\dagger|\Phi_n\rangle$ is an eigenstate of the operator $\hat{N}_\alpha$, with eigenvalue $(n_\alpha + 1)$; thus, the application of the operator $a_\alpha^\dagger$ on to the state $|\Phi_n\rangle$ *creates* a particle in the field. The operators a_α and $a_\alpha^\dagger$ are, therefore, referred to as the *annihilation* and *creation operators*. Of course, in each case the process (of annihilation or creation) is tied down to the single-particle state α; however, the precise location of the happening (in the coordinate space) remains undetermined; see eqns (20) and (21). Now, since the application of the operator a_α or $a_\alpha^\dagger$ on to the state $|\Phi_n\rangle$ of the field does not affect the eigenvalues of the particle-number operators other than $\hat{N}_\alpha$, we may write

$$a_\alpha|n_0, n_1, \ldots, n_\alpha, \ldots\rangle = A(n_\alpha)|n_0, n_1, \ldots, n_\alpha - 1, \ldots\rangle \tag{33}$$

and

$$a_\alpha^\dagger|n_0, n_1, \ldots, n_\alpha, \ldots\rangle = B(n_\alpha)|n_0, n_1, \ldots, n_\alpha + 1, \ldots\rangle, \tag{34}$$

where the factors $A(n_\alpha)$ and $B(n_\alpha)$ can be determined with the help of the commutation rules governing the operators a_α and $a_\alpha^\dagger$. For bosons,

$$A(n_\alpha) = \sqrt{n_\alpha}, \quad B(n_\alpha) = \sqrt{(n_\alpha + 1)}; \tag{35}$$

consequently, if we regard the state $|\Phi_n\rangle$ to have arisen from the vacuum state $|\Phi_0\rangle$ by a repeated application of the creation operators, we can write

$$|\Phi_n\rangle = \frac{1}{\sqrt{(n_0!n_1!\ldots n_\alpha!\ldots)}}(a_0^\dagger)^{n_0}(a_1^\dagger)^{n_1}\ldots(a_\alpha^\dagger)^{n_\alpha}\ldots|\Phi_0\rangle. \tag{36}$$

In the case of fermions, the operators $a_\alpha^\dagger$ anticommute, with the result that $a_\alpha^\dagger a_\beta^\dagger = -a_\beta^\dagger a_\alpha^\dagger$; consequently, there would remain an uncertainty of a phase factor ± 1 *unless* the order in which the $a_\alpha^\dagger$ operate on the vacuum state is specified. To be definite, let us agree that, as indicated in eqn. (36), the $a_\alpha^\dagger$ are arranged in the order of increasing subscripts and the phase factor is then $+1$. Secondly, since the product $a_\alpha^\dagger a_\alpha^\dagger$ now vanishes, none of the n_α in (36) can exceed unity; the eigenvalues of the fermion operators $\hat{N}_\alpha$ are, therefore, restricted to 0 and 1, which is precisely the requirement of the Pauli exclusion principle.[3] Accordingly, the factor $[\Pi_\alpha(n_\alpha!)]^{-1/2}$ in (36) would be identically equal to unity. In passing, we note that in the case of fermions operation (33) has meaning only if $n_\alpha = 1$ and operation (34) has meaning only if $n_\alpha = 0$.

Finally, the substitution of expressions (18) and (19) into (4) gives for the Hamiltonian operator of the field

$$\hat{H} = -\frac{\hbar^2}{2m}\sum_{\alpha,\beta}\langle\alpha|\nabla^2|\beta\rangle a_\alpha^\dagger a_\beta + \frac{1}{2}\sum_{\alpha,\beta,\gamma,\lambda}\langle\alpha\beta|u|\gamma\lambda\rangle a_\alpha^\dagger a_\beta^\dagger a_\gamma a_\lambda, \tag{37}$$

where

$$\langle\alpha|\nabla^2|\beta\rangle = \int d^3r\, u_\alpha^*(\boldsymbol{r})\nabla^2 u_\beta(\boldsymbol{r}) \tag{38}$$

and

$$\langle u_{\nu}|u|\gamma\lambda\rangle = \iint d^3r_1\, d^3r_2\, u_\alpha^*(\boldsymbol{r}_1)u_\beta^*(\boldsymbol{r}_2)u_{12}u_\gamma(\boldsymbol{r}_2)u_\lambda(\boldsymbol{r}_1). \tag{39}$$

Now, if the single-particle wave functions are chosen to be

$$u_\alpha(\boldsymbol{r}) = \frac{1}{\sqrt{V}}e^{i\boldsymbol{p}_\alpha\cdot\boldsymbol{r}/\hbar}, \tag{40}$$

where $\boldsymbol{p}_\alpha$ denotes the momentum of the particle (assumed "spinless"), then the matrix elements (38) and (39) become

$$\langle\alpha|\nabla^2|\beta\rangle = \frac{1}{V}\int d^3r\, e^{-i\boldsymbol{p}_\alpha\cdot\boldsymbol{r}/\hbar}\left(-\frac{p_\beta^2}{\hbar^2}\right)e^{i\boldsymbol{p}_\beta\cdot\boldsymbol{r}/\hbar} = -\frac{p_\beta^2}{\hbar^2}\delta_{\alpha\beta} \tag{41}$$

and

$$\langle\alpha\beta|u|\gamma\lambda\rangle = \frac{1}{V^2}\iint d^3r_1\, d^3r_2\, e^{-i(\boldsymbol{p}_\alpha-\boldsymbol{p}_\lambda)\cdot\boldsymbol{r}_1/\hbar}u(\boldsymbol{r}_2-\boldsymbol{r}_1)e^{-i(\boldsymbol{p}_\beta-\boldsymbol{p}_\gamma)\cdot\boldsymbol{r}_2/\hbar}. \tag{42}$$

In view of the fact that the total momentum is conserved in each collision,

$$\boldsymbol{p}_\alpha+\boldsymbol{p}_\beta=\boldsymbol{p}_\gamma+\boldsymbol{p}_\lambda, \tag{43}$$

the matrix element (42) takes the form

$$\begin{aligned}\langle\alpha\beta|u|\gamma\lambda\rangle &= \frac{1}{V^2}\iint d^3r_1\, d^3r_2\, e^{i(\boldsymbol{p}_\gamma-\boldsymbol{p}_\beta)\cdot(\boldsymbol{r}_2-\boldsymbol{r}_1)/\hbar}u(\boldsymbol{r}_2-\boldsymbol{r}_1)\\ &= \frac{1}{V}\int d^3r\, e^{i\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}u(\boldsymbol{r}),\end{aligned} \tag{44}$$

where $\boldsymbol{p}$ denotes the *momentum transfer* during the collision:

$$\boldsymbol{p} = (\boldsymbol{p}_\gamma-\boldsymbol{p}_\beta) = -(\boldsymbol{p}_\lambda-\boldsymbol{p}_\alpha). \tag{45}$$

Substituting (41) and (44) into (37), we finally obtain

$$\hat{H} = \sum_{\boldsymbol{p}}\frac{p^2}{2m}a_{\boldsymbol{p}}^\dagger a_{\boldsymbol{p}} + \frac{1}{2}\sum{}' u_{\boldsymbol{p}_1,\boldsymbol{p}_2}^{\boldsymbol{p}_1',\boldsymbol{p}_2'}a_{\boldsymbol{p}_1'}^\dagger a_{\boldsymbol{p}_2'}^\dagger a_{\boldsymbol{p}_2}a_{\boldsymbol{p}_1}, \tag{46}$$

where $u_{\boldsymbol{p}_1,\boldsymbol{p}_2}^{\boldsymbol{p}_1',\boldsymbol{p}_2'}$ denotes the matrix element (44), with

$$\boldsymbol{p} = (\boldsymbol{p}_2-\boldsymbol{p}_2') = -(\boldsymbol{p}_1-\boldsymbol{p}_1'); \tag{47}$$

note that the primed summation in the second term of (46) goes only over those values of $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, $\boldsymbol{p}_1'$ and $\boldsymbol{p}_2'$ that conserve the total momentum of the particles: $\boldsymbol{p}_1'+\boldsymbol{p}_2'=\boldsymbol{p}_1+\boldsymbol{p}_2$. It is obvious that the main term in (46) represents the *kinetic energy* of the field ($a_{\boldsymbol{p}}^\dagger a_{\boldsymbol{p}}$ being the particle-number operator pertaining to the single-particle state $\boldsymbol{p}$), while the second term represents the *potential energy*.

In the case of spin-half fermions, the single-particle states have to be characterized not only by the value $\boldsymbol{p}$ of the particle momentum but also by the value σ of the z-component of its spin; accordingly, the creation and annihilation operators would carry double indices. The operator $\hat{H}$ then takes the form

$$\hat{H} = \sum_{\boldsymbol{p},\sigma}\frac{p^2}{2m}a_{\boldsymbol{p}\sigma}^\dagger a_{\boldsymbol{p}\sigma} + \frac{1}{2}\sum{}' u_{\boldsymbol{p}_1\sigma_1,\boldsymbol{p}_2\sigma_2}^{\boldsymbol{p}_1'\sigma_1',\boldsymbol{p}_2'\sigma_2'}a_{\boldsymbol{p}_1'\sigma_1'}^\dagger a_{\boldsymbol{p}_2'\sigma_2'}^\dagger a_{\boldsymbol{p}_2\sigma_2}a_{\boldsymbol{p}_1\sigma_1}; \tag{48}$$

the summation in the second term now goes only over those states (of the two particles) that conform to the conditions of both momentum conservation and spin conservation.

In the following sections we shall apply the formalism of second quantization to investigate low-temperature properties of systems composed of interacting particles. In most cases we shall study these systems under the approximating conditions $a/\lambda \ll 1$ and $na^3 \ll 1$, where a is the *scattering length* of the two-body interaction, λ the *mean thermal wavelength* of the particles and n the *particle density* in the system. Now, the effective scattering cross-section for the collision of two particles, each of mass m, is primarily determined by the "scattering amplitude" $a(p)$, where

$$a(p) = \frac{m}{4\pi\hbar^2} \int u(r) e^{i p \cdot r/\hbar} \, d^3 r, \tag{49}$$

p being the momentum transfer during the collision; if the potential is central, eqn. (49) takes the form

$$a(p) = \frac{m}{4\pi\hbar^2} \int\limits_0^\infty u(r) \frac{\sin(kr)}{kr} 4\pi r^2 \, dr \qquad \left(k = \frac{p}{\hbar}\right) \tag{50}$$

For low-energy scattering (which implies "slow" collisions), we have the limiting result

$$a = \frac{m u_0}{4\pi\hbar^2}, \qquad u_0 = \int u(r) d^3 r, \tag{51}$$

the quantity a being the *scattering length* of the given potential.[4] Alternatively, one may employ the S-wave scattering phase shift $\eta_0(k)$, see Sec. 9.5, and write on one hand

$$\tan \eta_0(k) \simeq -\frac{mk}{4\pi\hbar^2} \int\limits_0^\infty u(r) \frac{\sin^2(kr)}{(kr)^2} 4\pi r^2 \, dr \tag{52}$$

and on the other

$$\cot \eta_0(k) = -\frac{1}{ka} + \frac{1}{2} k r^* + \cdots, \tag{53}$$

where a is the "scattering length" and r^* the "effective range" of the potential. For low-energy scattering, eqns (52) and (53) once again lead to (51). In passing, we note that a is positive or negative according as the potential in question is predominantly repulsive or predominantly attractive; unless a statement is made to the contrary, we shall assume a to be positive.

10.2. Low-temperature behavior of an imperfect Bose gas

The Hamiltonian of the quantized field for spinless bosons is given by the expression (10.1.46), where the matrix element $u_{p_1, p_2}^{p_1', p_2'}$ is a function of the momentum p transferred during the collision and is given by formula (10.1.44). At low temperatures the particle momenta are small, we may insert for the matrix elements $u(p)$ their value at $p = 0$, namely u_0/V, where u_0 is given by eqn. (10.1.51). At the same time, we may retain only those terms in the sum $\sum'$ that pertain to a vanishing momentum transfer. We then have

$$\hat{H} = \sum_p \frac{p^2}{2m} a_p^\dagger a_p + \frac{2\pi a\hbar^2}{mV}\left[\sum_p a_p^\dagger a_p^\dagger a_p a_p \right.$$

$$\left. + \sum_{p_1,p_2}{}^{*} \left(a_{p_1}^\dagger a_{p_2}^\dagger a_{p_2} a_{p_1} + a_{p_2}^\dagger a_{p_1}^\dagger a_{p_2} a_{p_1}\right)\right], \tag{1}$$

where the summation $\sum^*$ goes over all p_1 and p_2 such that $p_1 \neq p_2$. Now

$$\sum_p a_p^\dagger a_p^\dagger a_p a_p = \sum_p a_p^\dagger (a_p a_p^\dagger - 1) a_p = \sum_p (n_p^2 - n_p) = \sum_p n_p^2 - N, \tag{2}$$

whereas

$$\sum_{p_1,p_2}{}^{*} a_{p_1}^\dagger a_{p_2}^\dagger a_{p_2}, a_{p_1} = \sum_{p_1,p_2}{}^{*} n_{p_1} n_{p_2} = \sum_{p_1} n_{p_1}(N - n_{p_1}) = N^2 - \sum_p n_p^2, \tag{3}$$

the same being true of the sum over the exchange terms $a_{p_2}^\dagger a_{p_1}^\dagger a_{p_2} a_{p_1}$. Collecting these results, the energy eigenvalues of the system turn out to be

$$E\{n_p\} = \sum_p n_p \frac{p^2}{2m} + \frac{2\pi a\hbar^2}{mV}\left[2N^2 - N - \sum_p n_p^2\right]$$

$$\simeq \sum_p n_p \frac{p^2}{2m} + \frac{2\pi a\hbar^2}{mV}(2N^2 - n_0^2). \tag{4}[5]$$

We first examine the *ground state* of the given system, which corresponds to the distribution set

$$n_p \simeq \begin{cases} N & \text{for} \quad p = 0 \\ 0 & \text{for} \quad p \neq 0, \end{cases} \tag{5}$$

with the result that

$$E_0 \simeq \frac{2\pi a\hbar^2 N^2}{mV}. \tag{6}$$

The ground-state pressure is then given by

$$P_0 = -\left(\frac{\partial E_0}{\partial V}\right)_N = \frac{2\pi a\hbar^2 N^2}{mV^2} = \frac{2\pi a\hbar^2 n^2}{m}. \tag{7}$$

where $n(= N/V)$ is the particle density in the system. This leads to the velocity of sound, c_0, given by

$$c_0^2 = \frac{1}{m}\frac{dP_0}{dn} = \frac{4\pi a\hbar^2 n}{m^2}. \tag{8}$$

Inserting numbers relevant to liquid He^4, viz. $a \simeq 2.2$ Å, $n = 1/v$ where $v \simeq 45$ Å^3 per particle and $m \simeq 6.65 \times 10^{-24}$ g, we obtain: $c_0 \simeq 125$ m/s. A comparison with the actual velocity of sound in the liquid, which is about 240 m/s, should not be too disheartening, for the theory developed here was never intended to be applicable to a liquid. Finally, the chemical potential of the system at $T = 0$ K turns out to be

$$\mu_0 = \left(\frac{\partial E_0}{\partial N}\right)_V = \frac{4\pi a\hbar^2 N}{mV} = \frac{4\pi a\hbar^2 n}{m}. \tag{9}$$

At finite temperatures the physical behavior of the system may be studied through its partition function

$$\begin{aligned} Q(N, V, T) &= \sum_{\{n_p\}} \exp\left(-\beta E\{n_p\}\right) \\ &= \sum_{\{n_p\}} \exp\left[-\beta\left\{\sum_p n_p \frac{p^2}{2m} + \frac{2\pi a\hbar^2 N^2}{mV}\left(2 - \frac{n_0^2}{N^2}\right)\right\}\right]. \end{aligned} \tag{10}$$

In the lowest approximation, the quantity (n_0/N) appearing here may be replaced by its ideal-gas value, as given in Sec. 7.1, viz.

$$\frac{n_0}{N} = 1 - \frac{\lambda_c^3}{\lambda^3} \qquad \left[\lambda = \frac{h}{(2\pi mkT)^{1/2}}, \quad \lambda_c = \{v\zeta(3/2)\}^{1/3}\right] \tag{11a}$$

$$= 1 - \frac{v}{v_c} \qquad \left[v = \frac{V}{N}, \quad v_c = \frac{\lambda^3}{\zeta(3/2)}\right]. \tag{11b}$$

We thus obtain, to *first* order in a,

$$\ln Q(N, V, T) \simeq \ln Q_{id}(N, V, T) - \beta\frac{2\pi a\hbar^2 N^2}{mV}\left(1 + \frac{2v}{v_c} - \frac{v^2}{v_c^2}\right). \tag{12}$$

The Helmholtz free energy, per particle, is then given by[6]

$$\frac{1}{N}A(N, V, T) \simeq \frac{1}{N}A_{id}(N, V, T) + \frac{2\pi a\hbar^2}{m}\left(\frac{1}{v} + \frac{2}{v_c} - \frac{v}{v_c^2}\right). \tag{13}$$

The pressure P and the chemical potential μ now follow straightforwardly:

$$P = -\left(\frac{\partial A}{\partial V}\right)_{N,T} = -\left(\frac{\partial(A/N)}{\partial v}\right)_T = P_{id} + \frac{2\pi a\hbar^2}{m}\left(\frac{1}{v^2} + \frac{1}{v_c^2}\right) \tag{14}$$

and

$$\mu = \frac{A}{N} + Pv = \mu_{id} + \frac{4\pi a\hbar^2}{m}\left(\frac{1}{v} + \frac{1}{v_c}\right) \tag{15}$$

which may be compared with the ground-state results (7) and (9) that pertain to $v_c = \infty$.

At the transition point (where $v = v_c$ and $\lambda = \lambda_c$), the pressure P_c and the chemical potential μ_c turn out to be

$$P_c = P_{id} + \frac{4\pi a\hbar^2}{m\lambda_c^6}\left\{\zeta\left(\frac{3}{2}\right)\right\}^2 = \frac{kT_c}{\lambda_c^3}\left[\zeta\left(\frac{5}{2}\right) + 2\left\{\zeta\left(\frac{3}{2}\right)\right\}^2\frac{a}{\lambda_c}\right] \tag{16}$$

and

$$\mu_c = \mu_{id} + \frac{8\pi a\hbar^2}{m\lambda_c^3}\zeta\left(\frac{3}{2}\right) = 4\zeta\left(\frac{3}{2}\right)kT_c\frac{a}{\lambda_c}; \tag{17}$$

the corresponding value of the fugacity, z_c, is given by

$$z_c = \exp(\mu_c/kT_c) \simeq 1 + 4\zeta(3/2)(a/\lambda_c). \tag{18}$$

For a slightly different approach, see Problem 10.2.

10.3. Low-lying states of an imperfect Bose gas

In the preceding section we examined first-order corrections to the low-temperature behavior of an imperfect Bose gas arising from interparticle interactions in the system. One important result emerging in that study was a *nonzero* velocity of sound, as given by eqn. (10.2.8). This raises the possibility that phonons, the quanta of sound field, might play an important role in determining the low-temperature behavior of this system–a role not seen in Sec. 10.2. To look into this question we explore the nature of the low-lying states of an imperfect Bose gas, in the hope that we discover an energy-momentum relation $\varepsilon(p)$ obeyed by the elementary excitations of the system of which phonons may be an integral part. For this, we have to go a step beyond the approximation adopted in Sec. 10.2 which, in turn, requires several improvements. To keep matters simple, we confine ourselves to situations in which the fraction of particles occupying the state with $p = 0$ is fairly close to 1 while the fraction of particles occupying states with $p \neq 0$ is much less than 1.

Going back to eqns (10.2.1)–(10.2.4), we first write

$$2N^2 - n_0^2 = N^2 + (N^2 - n_0^2) \simeq N^2 + 2N(N - n_0) = N^2 + 2N\sum_{p\neq 0} a_p^\dagger a_p. \tag{1}$$

Next, we retain another set of terms from the sum $\sum'$ in eqn. (10.1.46)–terms that involve a nonzero momentum transfer, viz.

$$\sum_{p\neq 0} u(p)[a_p^\dagger a_{-p}^\dagger a_0 a_0 + a_0^\dagger a_0^\dagger a_p a_{-p}]. \tag{2}$$

Now, since $a_0^\dagger a_0 = n_0 = O(N)$ and $(a_0 a_0^\dagger - a_0^\dagger a_0) = 1 \ll N$, it follows that $a_0 a_0^\dagger = (n_0 + 1) \simeq a_0^\dagger a_0$. The operators a_0 and $a_0^\dagger$ may, therefore, be treated as c-numbers, each equal to $n_0^{1/2} \simeq N^{1/2}$. At the same time, the amplitude $u(p)$ in the case of low-lying states may be replaced by u_0/V, as before. Expression (2) then becomes

$$\frac{u_0 N}{V}\sum_{p\neq 0}(a_p^\dagger a_{-p}^\dagger + a_p a_{-p}). \tag{3}$$

In view of these results, the Hamiltonian of the system assumes the form

$$\hat{H} = \sum_p \frac{p^2}{2m} a_p^\dagger a_p + \frac{u_0}{2V}\left[N^2 + N\sum_{p\neq 0}(2a_p^\dagger a_p + a_p^\dagger a_{-p}^\dagger + a_p a_{-p})\right] \tag{4}$$

Our next task consists in determining an improved relationship between the quantity u_0 and the scattering length a. While the (approximate) result stated in eqn. (10.1.51) is good enough for evaluating the term involving $N\sum_{p\neq 0}$, it is not so for evaluating the term involving N^2. For this, we note that "if the probability of a particular quantum transition in a given system under the influence of a constant perturbation $\hat{V}$ is, in the *first* approximation, determined by the matrix element V_0^0 then in the *second* approximation we have instead

$$V_0^0 + \sum_{n\neq 0} \frac{V_n^0 V_0^n}{E_0 - E_n},$$

the summation going over the various states of the unperturbed system". In the present case, we are dealing with a collision process in the two-particle system (with reduced mass $\frac{1}{2}m$), and the role of V_0^0 is played by the quantity

$$u_{00}^{00} = \frac{1}{V}\int u(r)\, d^3r = \frac{u_0}{V};$$

see eqn. (10.1.44) for the matrix element $u_{p_1, p_2}^{p_1', p_2'}$. Making use of the other matrix elements, we find that in going from first to second approximation we have to replace u_0/V by

$$\frac{u_0}{V} + \frac{1}{V^2}\sum_{p\neq 0} \frac{|\int d^3r\, e^{i p\cdot r/\hbar} u(r)|^2}{-p^2/m} \simeq \frac{u_0}{V} - \frac{u_0^2 m}{V^2}\sum_{p\neq 0}\frac{1}{p^2}. \tag{5}$$

Equating (5) with the standard expression $4\pi a\hbar^2/mV$, we obtain. instead of (10.1.51),

$$u_0 \simeq \frac{4\pi a\hbar^2}{m}\left(1 + \frac{4\pi a\hbar^2}{V}\sum_{p\neq 0}\frac{1}{p^2}\right). \tag{6}$$

Substituting (6) into (4), we get

$$\begin{aligned}\hat{H} = {} & \frac{2\pi a\hbar^2}{m}\frac{N^2}{V}\left(1 + \frac{4\pi a\hbar^2}{V}\sum_{p\neq 0}\frac{1}{p^2}\right) \\ & + \frac{2\pi a\hbar^2}{m}\frac{N}{V}\sum_{p\neq 0}(2a_p^\dagger a_p + a_p^\dagger a_{-p}^\dagger + a_p a_{-p}) + \sum_{p\neq 0}\frac{p^2}{2m}a_p^\dagger a_p. \end{aligned} \tag{7}$$

To evaluate the energy levels of the system one would like to diagonalize the Hamiltonian (7). which can be done with the help of a *linear transformation* of the operators a_p and $a_p^\dagger$, first employed by Bogoliubov (1947):

$$b_p = \frac{a_p + \alpha_p a_{-p}^\dagger}{\sqrt{(1-\alpha_p^2)}}, \quad b_p^\dagger = \frac{a_p^\dagger + \alpha_p a_{-p}}{\sqrt{(1-\alpha_p^2)}}. \tag{8}$$

where

$$\alpha_p = \frac{mV}{4\pi a\hbar^2 N}\left\{\frac{4\pi a\hbar^2 N}{mV} + \frac{p^2}{2m} - \varepsilon(p)\right\}, \tag{9}$$

with

$$\varepsilon(p) = \left\{ \frac{4\pi a\hbar^2 N}{mV} \frac{p^2}{m} + \left(\frac{p^2}{2m} \right)^2 \right\}^{1/2}; \tag{10}$$

clearly, each $\alpha_p < 1$. Relations inverse to (8) are

$$a_p = \frac{b_p - \alpha_p b_{-p}^\dagger}{\sqrt{(1-\alpha_p^2)}}, \quad a_p^\dagger = \frac{b_p^\dagger - \alpha_p b_{-p}}{\sqrt{(1-\alpha_p^2)}}. \tag{11}$$

It is straightforward to check that the new operators b_p and $b_p^\dagger$ satisfy the same commutation rules as the old operators a_p and $a_p^\dagger$ did, namely

$$[b_p, b_{p'}^\dagger] = \delta_{pp'} \tag{12a}$$

$$[b_p, b_{p'}] = [b_p^\dagger, b_{p'}^\dagger] = 0. \tag{12b}$$

Substituting (11) into (7), we obtain our Hamiltonian in the diagonalized form:

$$\hat{H} = E_0 + \sum_{p \neq 0} \varepsilon(p) b_p^\dagger b_p, \tag{13}$$

where

$$E_0 = \frac{2\pi a\hbar^2 N^2}{mV} + \frac{1}{2} \sum_{p \neq 0} \left\{ \varepsilon(p) - \frac{p^2}{2m} - \frac{4\pi a\hbar^2 N}{mV} + \left(\frac{4\pi a\hbar^2 N}{mV} \right)^2 \frac{m}{p^2} \right\} \tag{14}$$

In view of the commutation rules (12) and expression (13) for the Hamiltonian operator $\hat{H}$, it seems natural to infer that the operators b_p and $b_p^\dagger$ are the annihilation and creation operators of certain "quasi-particles"–which represent *elementary excitations* of the system–with the energy–momentum relation given by (10); it is also clear that these quasi-particles obey Bose–Einstein statistics. The quantity $b_p^\dagger b_p$ is then the particle-number operator for the quasi-particles (or elementary excitations) of momentum p, whereby the second part of the Hamiltonian (13) becomes the energy operator corresponding to the elementary excitations present in the system. The first part of the Hamiltonian, given explicitly by eqn. (14), is therefore the *ground state* energy of the system. Replacing the summation over p by an integration and introducing a dimensionless variable x, defined by

$$x = p \left(\frac{V}{8\pi a\hbar^2 N} \right)^{1/2}$$

we obtain for the ground state energy of the system

$$E_0 = \frac{2\pi a\hbar^2 N^2}{mV} \left[1 + \left(\frac{128 N a^3}{\pi V} \right)^{1/2} \int_0^\infty dx \left[x^2 \left(x\sqrt{(x^2+2)} - x^2 - 1 + \frac{1}{2x^2} \right) \right] \right]. \tag{15}$$

The value of the integral turns out to be $(128)^{1/2}/15$, with the result

$$\frac{E_0}{N} = \frac{2\pi a\hbar^2 n}{m}\left[1 + \frac{128}{15\pi^{1/2}}(na^3)^{1/2}\right], \tag{16}$$

where n denotes the particle density in the system. Equation (16) represents the first two terms of the expansion of the quantity E_0/N in terms of the low-density parameter $(na^3)^{1/2}$; the first term was already obtained in Sec. 10.2.[7]

The foregoing result was first derived by Lee and Yang (1957) using the binary collision method; the details of this calculation, however, appeared somewhat later (see Lee and Yang, 1960a; see also Problem 10.6). Using the pseudopotential method, this result was rederived by Lee, Huang and Yang (1957).

The ground state pressure of the system is now given by

$$\begin{aligned} P_0 &= -\left(\frac{\partial E_0}{\partial V}\right)_N = n^2\frac{\partial(E_0/N)}{\partial n} \\ &= \frac{2\pi a\hbar^2 n^2}{m}\left[1 + \frac{64}{5\pi^{1/2}}(na^3)^{1/2}\right], \end{aligned} \tag{17}$$

whence one obtains for the velocity of sound

$$c_0^2 = \frac{1}{m}\frac{dP_0}{dn} = \frac{4\pi a\hbar^2 n}{m^2}\left[1 + \frac{16}{\pi^{1/2}}(na^3)^{1/2}\right]. \tag{18}$$

Equations (17) and (18) are an improved version of the results obtained in Sec. 10.2.

The ground state of the system is characterized by a total absence of excitations; accordingly, the eigenvalue of the (number) operator $b_p^\dagger b_p$ of the "quasi-particles" must be zero for all $p \neq 0$. As for the real particles, there must be some that possess nonzero energies even at absolute zero, for otherwise the system cannot have a finite amount of energy in the ground state. The momentum distribution of the real particles can be determined by evaluating the ground-state expectation values of the number operators $a_p^\dagger a_p$. Now, in the ground state of the system,

$$a_p|\Psi_0\rangle = \frac{1}{\sqrt{(1-\alpha_p^2)}}(b_p - \alpha_p b_{-p}^\dagger)|\Psi_0\rangle = \frac{-\alpha_p}{\sqrt{(1-\alpha_p^2)}}b_{-p}^\dagger|\Psi_0\rangle \tag{19}$$

because $b_p|\Psi_0\rangle \equiv 0$. Constructing the hermitian conjugate of (19) and remembering that α_p is real, we have

$$\langle\Psi_0|a_p^\dagger = \frac{-\alpha_p}{\sqrt{(1-\alpha_p^2)}}\langle\Psi_0|b_{-p}. \tag{20}$$

The scalar product of expressions (19) and (20) gives

$$\langle\Psi_0|a_p^\dagger a_p|\Psi_0\rangle = \frac{\alpha_p^2}{1-\alpha_p^2}\langle\Psi_0|b_{-p}b_{-p}^\dagger|\Psi_0\rangle = \frac{\alpha_p^2}{1-\alpha_p^2}; \tag{21}$$

here, use has been made of the facts that (i) $b_p b_p^\dagger - b_p^\dagger b_p = 1$ and (ii) in the ground state, for all $p \neq 0$, $b_p^\dagger b_p = 0$ (and hence $b_p b_p^\dagger = 1$). Thus, for $p \neq 0$,

$$\bar{n}_p = \frac{\alpha_p^2}{1-\alpha_p^2} = \frac{x^2+1}{2x\sqrt{(x^2+2)}} - \frac{1}{2}, \tag{22}$$

where $x = p(8\pi a\hbar^2 n)^{-1/2}$. The total number of "excited" particles in the ground state of the system is, therefore, given by

$$\sum_{p\neq 0} \bar{n}_p = \sum_{p\neq 0} \frac{\alpha_p^2}{1-\alpha_p^2} = \sum_{x>0} \frac{1}{2}\left(\frac{x^2+1}{x\sqrt{(x^2+2)}} - 1\right)$$

$$\simeq N\left\{\frac{32}{\pi}(na^3)\right\}^{1/2} \int\limits_0^\infty dx \left[x^2\left(\frac{x^2+1}{x\sqrt{(x^2+2)}} - 1\right)\right] \tag{23}$$

The value of the integral turns out to be $(2)^{1/2}/3$, with the result

$$\sum_{p\neq 0} \bar{n}_p \simeq N \frac{8}{3\pi^{1/2}}(na^3)^{1/2}. \tag{24}$$

Accordingly,

$$\bar{n}_0 = N - \sum_{p\neq 0} \bar{n}_p \simeq N\left[1 - \frac{8}{3\pi^{1/2}}(na^3)^{1/2}\right]. \tag{25}$$

The foregoing result was first obtained by Lee, Huang and Yang (1957), using the pseudopotential method. It may be noted here that the importance of the *real-particle* occupation numbers n_p in the study of the ground state of an interacting Bose system had been emphasized earlier by Penrose and Onsager (1956).

10.4. Energy spectrum of a Bose liquid

In this section we propose to study the most essential features of the energy spectrum of a Bose liquid and to examine the relevance of this study to the problem of liquid He^4. In this context we have seen that the low-lying states of a *low-density gaseous* system composed of weakly interacting bosons are characterized by the presence of the so-called *elementary excitations* (or "quasi-particles"), which are themselves bosons and whose energy spectrum is given by

$$\varepsilon(p) = \{p^2u^2 + (p^2/2m)^2\}^{1/2}, \tag{1}$$

where

$$u = (4\pi an)^{1/2}(\hbar/m); \tag{2}$$

see eqns (10.3.10, 12 and 13).[8] For $p \ll mu$, i.e. $p \ll \hbar(an)^{1/2}$, the spectrum is essentially linear: $\varepsilon \simeq pu$. The initial slope of the (ε, p)-curve is, therefore, given by the parameter u which is identical with the limiting value of the velocity of sound in the system; compare (2) with (10.3.18). It is then natural that these low-momentum excitations be identified as *phonons* – the quanta of the sound field. For $p \gg mu$, the spectrum approaches essentially the classical limit: $\varepsilon \simeq p^2/2m + \Delta^*$, where $\Delta^* = mu^2 = 4\pi an\hbar^2/m$. It is important to note that, all along, this energy–momentum relationship is strictly *monotonic* and does not display any "dip" of the kind propounded by Landau (for liquid He^4) and observed by Yarnell *et al.* and by Henshaw and Woods; see Sec. 7.5. Thus, the spectrum provided by the theory of the preceding sections simulates Landau spectrum only to the extent of phonons; it does not account for rotons. This should not be surprising, for the theory in question was intended only for a low-density Bose gas ($na^3 \ll 1$) and not for liquid $\mathrm{He}^4 (na^3 \simeq 0.2)$.

The problem of elementary excitations in liquid He^4 was tackled successfully by Feynman who, in 1953–54, developed an atomic theory of a *Bose liquid* at low temperatures. In a series of three fundamental papers starting from first principles, Feynman established the following important results.[9]

(i) In spite of the presence of interatomic forces, a Bose liquid undergoes a phase transition analogous to the momentum-space condensation occurring in the ideal Bose gas; in other words, the original suggestion of London (1938a,b) regarding liquid He^4, see Sec. 7.1, is essentially correct.

(ii) At sufficiently low temperatures, the only excited states possible in the liquid are the ones related to compressional waves, viz. *phonons*. Long-range motions, which leave the density of the liquid unaltered (and consequently imply nothing more than a simple "stirring" of the liquid), do not constitute excited states because they differ from the ground state only in the "permutation" of certain atoms. Motions on an atomic scale are indeed possible, but they require a *minimum* energy Δ for their excitation; clearly, these excitations would show up only at comparatively higher temperatures ($T \sim \Delta/k$) and might well turn out to be Landau's *rotons*.

(iii) The wave function of the liquid, in the presence of *an* excitation, should be approximately of the form

$$\Psi = \Phi \sum_i f(\boldsymbol{r}_i), \tag{3}$$

where Φ denotes the ground state wave function of the system while the summation of $f(\boldsymbol{r}_i)$ goes over all the N coordinates $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N$; the wave function Ψ is, clearly, symmetric in its arguments. The exact character of the function $f(\boldsymbol{r})$ can be determined by making use of a variational principle which requires the energy of the state Ψ (and hence the energy associated with the excitation in question) to be a minimum.

The optimal choice for $f(\boldsymbol{r})$ turns out to be, see Problem 10.8,

$$f(\boldsymbol{r}) = \exp i(\boldsymbol{k} \cdot \boldsymbol{r}), \tag{4}$$

with the (minimized) energy value

$$\varepsilon(\boldsymbol{k}) = \frac{\hbar^2 k^2}{2mS(\boldsymbol{k})}, \tag{5}$$

where $S(\boldsymbol{k})$ is the *structure factor* of the liquid, i.e. the Fourier transform of the *pair distribution function* $g(\boldsymbol{r})$:

$$S(\boldsymbol{k}) = \int g(\boldsymbol{r}) \exp i(\boldsymbol{k} \cdot \boldsymbol{r})\, d^3r; \tag{6}$$

it may be recalled here that the function $g(r_2 - r_1)$ denotes the probability density for finding an atom in the neighborhood of the point r_2 when another one is known to be at the point r_1. The optimal wave function is, therefore, given by

$$\Psi = \Phi \sum_i e^{i(k \cdot r_i)}. \tag{7}$$

Now the momentum associated with this excited state is $\hbar k$ because

$$P\Psi = \left(-i\hbar \sum_i \nabla_i\right) \Psi = \hbar k \Psi, \tag{8}$$

$P\Phi$ being identically equal to zero. Naturally, this would be interpreted as the momentum p associated with the excitation. One thus obtains, from first principles, the energy–momentum relationship for the elementary excitations in a Bose liquid.

On physical grounds one can show that, for small k, the structure factor $S(k)$ rises linearly as $\hbar k/2mc$, reaches a maximum near $k = 2\pi/r_0$ (corresponding to a maximum in the pair distribution function at the nearest-neighbor spacing r_0, which for liquid He^4 is about 3.6 Å) and thereafter decreases to approach, with

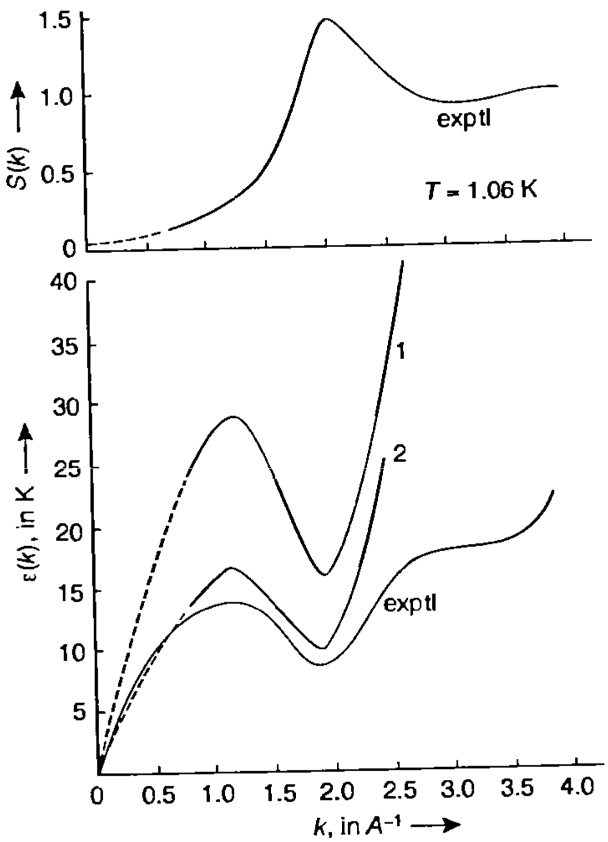


FIG. 10.1. The energy spectrum of the elementary excitations in liquid He^4. The upper portion shows the structure factor of the liquid, as derived by Henshaw (1960) from experimental data on neutron diffraction. Curve 1 in the lower portion shows the energy–momentum relationship based on the Feynman formula (5) while curve 2 is based on an improved formula due to Feynman and Cohen (1956). For comparison, the experimental results of Woods (1966) obtained directly from neutron scattering are also included.

minor oscillations (corresponding to the subsidiary maxima in the pair distribution function at the spacings of the next nearest neighbours), the limiting value 1 for large k; the limiting value 1 arises from the presence of a delta function in the expression for $g(\boldsymbol{r})$ (because, as $\boldsymbol{r}_2 \to \boldsymbol{r}_1$, one is sure to find an atom there).[10] Accordingly, the energy $\varepsilon(k)$ of an elementary excitation in liquid He^4 would start linearly as $\hbar kc$, show a "dip" at $k_0 \simeq 2\,\text{Å}^{-1}$ [11] and rise again to approach the eventual limit of $\hbar^2k^2/2m$. These features are shown in Fig. 10.1. Clearly, Feynman's approach merges both phonons and rotons into a single, unified scheme in which they represent different parts of a common (and continuous) energy spectrum $\varepsilon(k)$, as determined by the structure of the liquid through the function $S(k)$. Since no motion of a rotational character is involved, the name "roton" is clearly a misnomer.

It seems appropriate to mention here that, soon after the work of London which advocated a connection between the phase transition in liquid He^4 and the phenomenon of Bose–Einstein condensation, Bijl (1940) investigated the mathematical structure of the wave functions appropriate to an interacting Bose gas and the excitation energy associated with those wave functions. His picture corresponded very closely to Feynman's and indeed led to the wave function (7). Bijl also derived an expression for $\varepsilon(k)$ which was exactly the same as (5). Unfortunately, he could not make much use of his results–primarily because he leaned heavily on the expansion

$$S(k) = S(0) + C_2k^2 + C_4k^4 + \cdots \tag{9}$$

which, as we now know, represents neither phonons nor rotons.

10.5. States with quantized circulation

We now proceed to examine the possibility of "organized motion" in the ground state of a Bose fluid. In this context, the most important concept is embodied in the *circulation theorem* of Feynman (1955), which establishes a physical basis for the existence of "quantized vortex motion" in the fluid. In the case of liquid helium II, this concept has successfully resolved some of the vital questions which had been baffling superfluid physicists for a long time.

The ground state wave function of a superfluid, composed of N bosons, may be denoted by a symmetric function $\Phi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$; if the superfluid does not partake in any organized motion, then Φ will be a pure real number. If, on the other hand, it possesses a *uniform* mass-motion with velocity $\boldsymbol{v}_s$, then its wave function would be

$$\Psi = \Phi e^{i(\boldsymbol{P}_s \cdot \boldsymbol{R})/\hbar} = \Phi e^{im(\boldsymbol{v}_s \cdot \sum_i \boldsymbol{r}_i)/\hbar} \tag{1}$$

where $\boldsymbol{P}_s$ denotes the total momentum of the fluid and $\boldsymbol{R}$ its center of mass:

$$\boldsymbol{P}_s = Nm\boldsymbol{v}_s; \quad \boldsymbol{R} = N^{-1}\sum_i \boldsymbol{r}_i. \tag{2}$$

Equation (1) gives the desired wave function *exactly* if the drift velocity $\boldsymbol{v}_s$ is uniform throughout the fluid. If $\boldsymbol{v}_s$ is non-uniform, then the present wave function would still be good locally–in the sense that the phase change $\Delta\phi$ resulting from a “set of local displacements” of the atoms (over distances too small for velocity variations to be appreciable) would be practically the same as the one following from expression (1). Thus, for a given set of displacements $\Delta \boldsymbol{r}_i$ of the atoms constituting the fluid, the change in the phase of the wave function would very nearly be

$$\Delta\phi = \frac{m}{\hbar}\sum_i (\boldsymbol{v}_{si}\cdot\Delta\boldsymbol{r}_i), \tag{3}$$

where $\boldsymbol{v}_s$ is now a function of $\boldsymbol{r}$.

The foregoing result may be used for calculating the net phase change resulting from a displacement of atoms along a ring, from their original positions in the ring to the neighboring ones, so that after displacement we obtain a configuration which is physically identical with the one we started with; see Fig. 10.2. In view of the symmetry of the wave function, the net phase change resulting from such a displacement must be an integral multiple of 2π (so that the wave function after the displacement is identical with the one before the displacement):

$$\frac{m}{\hbar}\sum_i{}' (\boldsymbol{v}_{si}\cdot\Delta\boldsymbol{r}_i) = 2\pi n, \qquad n = 0, \pm 1, \pm 2, \ldots; \tag{4}$$

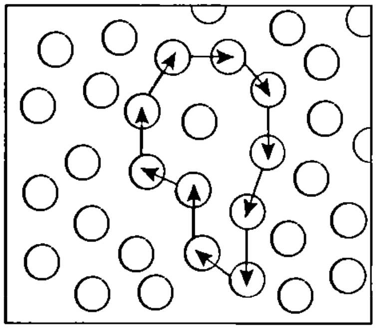


FIG. 10.2. The wave function of the fluid must not change as a result of a permutation of the atoms If all the atoms are displaced around a ring, as shown, the phase change must be a multiple of 2π.

the summation $\sum'$ here goes over all the atoms constituting the ring. We note that for the foregoing result to be valid it is *only* the individual $\Delta \boldsymbol{r}_i$ that have to be small, *not* the whole perimeter of the ring. Now, for a ring of a *macroscopic* size, one may regard the fluid as a continuum; eqn. (4) then becomes

$$\oint \boldsymbol{v}_s\cdot d\boldsymbol{r} = n\frac{h}{m}, \qquad n = 0, \pm 1, \pm 2, \ldots. \tag{5}$$

The quantity on the left-hand side of this equation is, by definition, the *circulation* (of the flow) associated with the circuit of integration and is clearly quantized, the “quantum of circulation” being h/m. Equation (5) constitutes the *circulation theorem* of Feynman; it bears a striking resemblance to the quantum condition of Bohr, namely

$$\oint p\,dq = nh, \tag{6}$$

though the region of application here is macroscopic rather than microscopic.[12]

By Stokes's theorem, eqn. (5) may be written as

$$\int_S (\operatorname{curl} \boldsymbol{v}_s)\cdot d\boldsymbol{S} = n\frac{h}{m}. \quad n = 0, \pm 1, \pm 2, \ldots, \tag{7}$$

where S denotes the area enclosed by the circuit of integration. If this area is "simply-connected" and the velocity $\boldsymbol{v}_s$ is continuous throughout the area, then the domain of integration can be shrunk in a *continuous* manner without limit. The integral on the left-hand side is then expected to decrease *continuously* and finally tend to zero. The right-hand side, however, *cannot* vary continuously. We infer that in this particular case the quantum number n must be zero, i.e. our integral must be identically vanishing. Thus, "in a simply-connected region, in which the velocity field is continuous throughout, the condition

$$\operatorname{curl} \boldsymbol{v}_s = 0 \tag{8}$$

holds everywhere". This is precisely the condition postulated by Landau (1941), which has been the cornerstone of our understanding of the hydrodynamic behavior of superfluid helium.[13]

Clearly, the Landau condition is only a special case of the Feynman theorem. It is quite possible that in a "multiply-connected" domain, which cannot be shrunk *continuously* to zero (without encountering singularities in the velocity field), the Landau condition may not hold everywhere. A typical example of such a domain is provided by the flow of a *vortex*, which is a planar flow with cylindrical symmetry such that

$$v_\rho = 0. \quad v_\phi = \frac{K}{2\pi\rho}, \quad v_z = 0, \tag{9}$$

where ρ is the distance measured perpendicular to the axis of symmetry while K is the circulation of the flow:

$$\oint \boldsymbol{v}\cdot d\boldsymbol{r} = \oint v_\phi(\rho\, d\phi) = K; \tag{10}$$

note that the circuit of integration in (10) must enclose the axis of the vortex. Another version of the foregoing result is

$$\int_S (\operatorname{curl} \boldsymbol{v})\cdot d\boldsymbol{S} = \int_S \left\{\frac{1}{\rho}\frac{d}{d\rho}(\rho v_\phi)\right\}(2\pi\rho\, d\rho) = K. \tag{11}$$

Now, at all $\rho \neq 0$, $\operatorname{curl} \boldsymbol{v} = 0$ but at $\rho = 0$, where v_ϕ is singular. $\operatorname{curl} \boldsymbol{v}$ appears to be indeterminate; it is not difficult to see that, at $\rho = 0$, $\operatorname{curl} \boldsymbol{v}$ diverges (in such a way that the integral in (11) turns out to be finite). In this context, it seems worthwhile to point out that if we carry out the integration in (10) along a circuit which *does not* enclose the axis of the vortex, or in (11) over a region that does not include the point $\rho = 0$, the result would be identically zero. At this stage we note that the energy associated with a unit length of a *classical* vortex is given by

$$\frac{\mathscr{E}}{L} = \int_a^b \frac{1}{2}\{2\pi\rho\, d\rho(mn_0)\}\left(\frac{K}{2\pi\rho}\right)^2 = \frac{mn_0K^2}{4\pi}\ln(b/a). \tag{12}$$

Here, (mn_0) is the mass density of the fluid (which is assumed to be uniform), the upper limit b is related to the size of the container while the lower limit a depends upon the structure of the vortex; in our study, a would be comparable to the interatomic separation.

In the *quantum-mechanical* case we may describe our vortex through a self-consistent wave function $\psi(\boldsymbol{r})$ which, in the case of cylindrical symmetry, see eqn. (9), may be written as

$$\psi(\boldsymbol{r}) = n^{*1/2} e^{is\phi} f_s(\rho), \qquad (13)$$

so that

$$n(\boldsymbol{r}) \equiv |\psi(\boldsymbol{r})|^2 = n^* f_s^2(\rho). \qquad (14)$$

As $\rho \to \infty$, $f_s(\rho) \to 1$, so that n^* becomes the limiting particle density in the fluid in regions far away from the axis of the vortex. The velocity field associated with this wave function will be

$$\begin{aligned} \boldsymbol{v}(\boldsymbol{r}) &= \frac{\hbar}{2im(\psi^*\psi)}(\psi^* \nabla\psi - \psi \nabla\psi^*) \\ &= \frac{\hbar}{m} \nabla(s\phi) = \left(0, s\frac{\hbar}{m\rho}, 0\right). \qquad (15)[14] \end{aligned}$$

Comparing (15) with (9), we conclude that the circulation K in the present case is sh/m; by the circulation theorem, s must be an integer:

$$s = 0, \pm 1, \pm 2, \ldots. \qquad (16)$$

Clearly, the value 0 is of no interest to us. Furthermore, the negative values of s differ from the positive ones only in the "sense of rotation" of the fluid. It is, therefore, sufficient to consider the positive values alone, viz.

$$s = 1, 2, 3, \ldots. \qquad (17)$$

The function $f_s(\rho)$ appearing in eqn. (13) may be determined with the help of a Schrödinger equation in which the potential term is itself ψ-dependent, viz.

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + u_0|\psi|^2\right)\psi = \varepsilon\psi, \qquad (18)$$

where u_0 is given by eqn. (10.1.51):

$$u_0 = 4\pi a\hbar^2/m, \qquad (19)$$

a being of the scattering length of the interparticle interaction operating in the fluid. The characteristic energy ε follows from the observation that at large distances from the axis of the vortex the fluid is essentially uniform in density, with $n(\boldsymbol{r}) \to n^*$; eqn. (18) then gives

$$\varepsilon = u_0 n^* = 4\pi a\hbar^2 n^*/m, \qquad (20)$$

which may be compared with eqn. (10.2.9). Substituting (20) into (18) and remembering that the flow is cylindrically symmetrical, we get

$$-\left[\frac{1}{\rho}\frac{d}{d\rho}\left\{\rho\frac{d}{d\rho}f_s(\rho)\right\} - \frac{s^2}{\rho^2}f_s(\rho)\right] + 8\pi a n^* f_s^3(\rho) = 8\pi a n^* f_s(\rho). \qquad (21)$$

Expressing ρ in terms of a characteristic length l,

$$\rho = l\rho' \quad \left\{l = (8\pi a n^*)^{-1/2}\right\}, \tag{22}$$

we obtain

$$\frac{d^2 f_s}{d\rho'^2} + \frac{1}{\rho'}\frac{df_s}{d\rho'} + \left(1 - \frac{s^2}{\rho'^2}\right) f_s - f_s^3 = 0. \tag{23}$$

Towards the axis of the vortex, where $\rho \to 0$, the very high velocity of the fluid (and the very large centrifugal force accompanying it) will push the particles outward, thus causing an enormous decrease in the density of the fluid. Consequently, the function f_s should tend to zero as $\rho \to 0$. This will make the last term in eqn. (23) negligible and thereby reduce it to the familiar Bessel's equation. Accordingly, for small ρ,

$$f_s(\rho') \sim J_s(\rho') \sim \rho^s, \tag{24}$$

J_s being the *ordinary Bessel function* of order s. For $\rho' \gg 1$, $f_s \simeq 1$; then, the first two terms of eqn. (23) become negligible, with the result

$$f_s(\rho') \simeq 1 - \frac{s^2}{2\rho'^2}. \tag{25}$$

The full solution is obtained by integrating the equation numerically; the results so obtained are shown in Fig. 10.3 where solutions for $s = 1$, 2 and 3 are displayed.

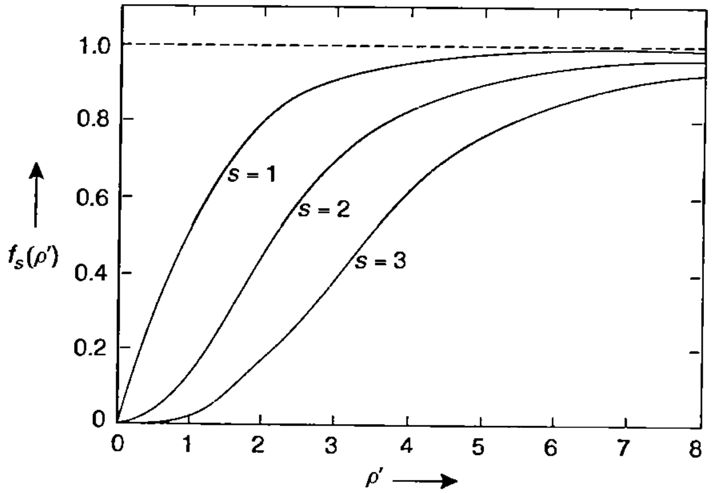


FIG. 10.3. Solutions of eqn. (23) for various values of s (after Kawatra and Pathria, 1966).

We thus find that our model of an imperfect Bose gas does allow the presence of quantized vortices in the system. Not only that, we do not have to invoke here any special assumptions regarding the nature of the "core" of the vortex (as one has to do in the classical theory); our treatment naturally leads to a continual diminution of the particle density n as the axial line is approached, so that there does not exist any specific distribution of vorticity around this line. The distance scale, which governs the spatial variation of n, is provided by the parameter l of eqn. (22); for liquid He^4, $l \simeq 1$ Å.

Pitaevskii 61), who was among the first to demonstrate the possibility of obtaining solutions whose natural interpretation lay in quantized vortex motion (see also Gross 1961; Weller 1963), also evaluated the energy per unit length of the vortex. Employing the wave function (13), with known values of the functions $f_s(\rho)$, Pitaevskii obtained the following results for the energy per unit length of the vortex, with $s = 1$, 2 or 3,

$$\frac{n^*h^2}{4\pi m}\{1 \ln(1.46R/l), \quad 4\ln(0.59R/l), \quad 9\ln(0.38R/l)\}, \tag{26}$$

where R denotes the outer radius of the domain involved. The above results may be compared with the "semi-classical" ones, viz.

$$\frac{n_0h^2}{4\pi m}\{1 \ln(R/a), \quad 4\ln(R/a), \quad 9\ln(R/a)\}, \tag{27}$$

which obtain from formula (12), with K replaced by sh/m and b by R. It is obvious that vortices with $s > 1$ would be relatively unstable because energetically it would be cheaper for a system to have s vortices of unit circulation rather than a single vortex of circulation s.

The existence of quantized vortex lines in liquid helium II has been demonstrated convincingly by the ingenious experiments of Vinen (1958–61) in which the circulation K around a fine wire immersed in the liquid was measured through the influence it exerts on the transverse vibrations of the wire. Vinen found that while vortices with unit circulation were exceptionally stable those with higher circulation too made appearance. Repeating Vinen's experiment with thicker wires, Whitmore and Zimmermann (1965) were able to observe stable vortices with circulation up to three quantum units. For a survey of this and other aspects of the superfluid behavior, see Vinen (1968) and Betts (1969).

For the relevance of quantized vortex lines to the problem of "rotation" of the superfluid, see Sec. 10.7 of the first edition of this book.

10.6. Quantized vortex rings and the breakdown of superfluidity

Feynman (1955) was the first to suggest that the formation of vortices in liquid helium II might provide the mechanism responsible for the breakdown of superfluidity in the liquid. He considered the flow of liquid helium II from an orifice of diameter D and, by tentative arguments, found that the velocity v_0 at which the flow energy available would just be sufficient to create quantized vortices in the liquid is given by

$$v_0 = \frac{\hbar}{mD}\ln(D/l). \tag{1}$$

Thus, for an orifice of diameter 10^{-5} cm, v_0 would be of the order of 1 m/s.[15] It is tempting to identify v_0 with v_c, the *critical velocity* of superflow through the given capillary, despite the fact that the theoretical estimate for v_0 is an order of magnitude higher than the corresponding experimental values of v_c; the latter, for instance, are 13 cm/s, 8 cm/s and 4 cm/s for capillary diameters 1.2×10^{-5} cm, 7.9×10^{-5} cm and 3.9×10^{-4} cm, respectively. Nevertheless, the present estimate is far more desirable than the prohibitively large ones obtained earlier on the basis of a possible creation of phonons or rotons in the liquid; see Sec. 7.5. Moreover, one obtains here a definitive dependence of the critical velocity of superflow on the width of the capillary employed which, at least qualitatively, agrees with the trend seen in the experimental findings. In what follows, we propose to develop Feynman's idea further along the lines suggested by the preceding section.

So far we have been dealing with the so-called *linear vortices* whose velocity field possesses cylindrical symmetry. More generally, however, a vortex line need not be straight–it may be curved and, if it does not terminate on the walls of the container or on the free surface of the liquid, may close on itself. We then speak of a *vortex ring*, which is very much like a smoke ring. Of course, the quantization condition (10.5.5) is as valid for a vortex ring as for a vortex line. However, the dynamical properties of a ring are quite different from those of a line; see, for instance, Fig. 10.4 which shows schematically a vortex ring in cross-section, the radius r of the ring being much larger than the core dimension l. The flow velocity v_r at any point in the field is determined by a superposition of the flow velocities due to the various elements of the ring. It is not difficult to see that the velocity field of the ring, including the ring itself, moves in a direction perpendicular to the plane of the ring, with a velocity[16]

$$v \sim \hbar/2mr; \tag{2}$$

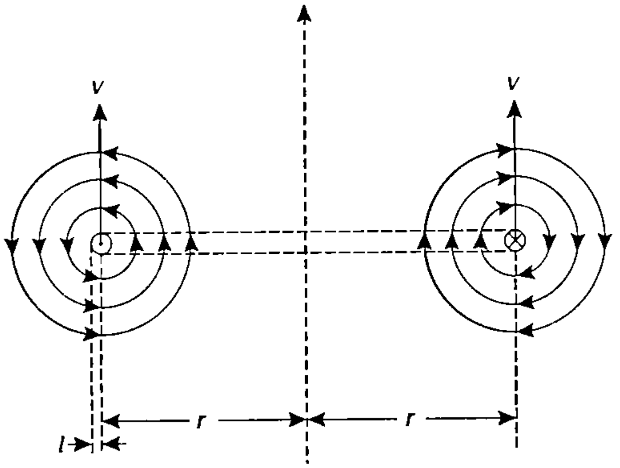


FIG. 10.4. Schematic illustration of a quantized vortex ring in cross-section.

see eqn. (10.5.15), with $s = 1$ and $\rho \sim 2r$. An estimate of the energy associated with the flow may be obtained from expression (10.5.12), with $L = 2\pi r$, $K = h/m$ and $b \sim r$; thus

$$\varepsilon \sim 2\pi^2\hbar^2 n_0 m^{-1} r \ln(r/l). \tag{3}$$

Clearly, the dependence of ε on r arises mainly from the factor r and only slightly from the factor $\ln(r/l)$. Therefore, with good approximation $v \propto \varepsilon^{-1}$, i.e. a ring with larger energy moves slower! The reason behind this somewhat startling result is that firstly the larger the ring the larger the distances between the various circulation-carrying elements of the ring (thus reducing the velocity imparted by one element to another) and secondly a larger ring carries with it a larger amount of the fluid ($M \propto r^3$), so the total energy associated with the ring is also larger (essentially proportional to Mv^2, i.e. $\propto r$). The product $v\varepsilon$, apart from the slowly-varying factor $\ln(r/l)$, is thus a constant, which is equal to $\pi^2\hbar^3 n_0/m^2$.

It is gratifying that vortex rings such as the ones discussed here have been observed and the circulation around them is found to be as close to the quantum h/m as one could expect under the conditions of the experiment. Figure 10.5 shows the experimental results of Rayfield and Reif (1964) for the velocity–energy relationship of free-moving, charge-carrying vortex rings created in liquid helium II by suitably accelerated helium ions. Vortex rings carrying positive as well as negative charge were observed; dynamically, however, they behaved *alike*, as one indeed expects because both the velocity and the energy associated with a vortex ring are determined by the properties of a large amount of fluid carried along with the ring rather than by the small charge coupled to it. Fitting experimental results with the notion of the vortex rings, Rayfield and Reif concluded that their rings carried a circulation of $(1.00 \pm 0.03) \times 10^{-3}$ cm^2/s, which is close to the Onsager–Feynman unit $h/m (= 0.997 \times 10^{-3}$ cm^2/s); moreover, they seemed to have a core radius of about 1.2 Å, which is comparable with the characteristic parameter l of the fluid.

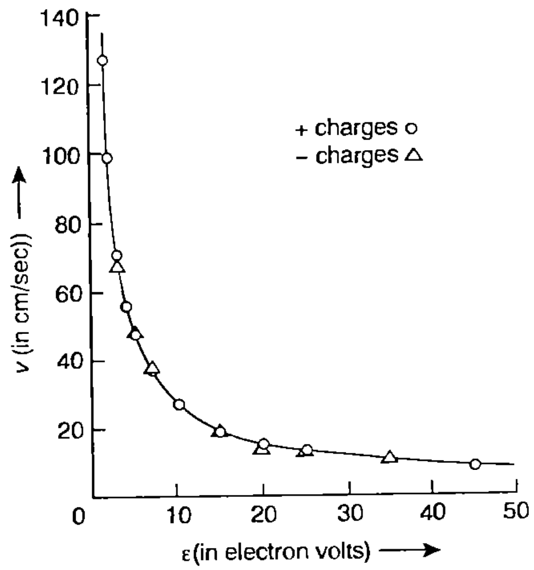


FIG. 10.5. The velocity–energy relationship of the vortex rings formed in liquid helium II (after Rayfield and Reif, 1964). The points indicate the experimental data, while the curve represents the theoretical relationship based on the "quantum of circulation" h/m.

We shall now show that the dynamics of the quantized vortex rings is such that their creation in liquid helium II does provide a mechanism for the *breakdown of superfluidity*. To see this, it is simplest to consider the case of a superfluid flowing through a capillary of radius R. As the velocity of flow increases and approaches the critical value v_c, quantized vortex rings begin to form and energy dissipation sets in, which in turn brings about the rupture of the superflow. By symmetry, the rings will be so formed that their central plane will be perpendicular to the axis of the capillary and they will be moving in the direction of the main flow. Now, by the Landau criterion (7.5.24), the critical velocity of superflow is directly determined by the energy spectrum of the excitations [illegible]ated:

$$v_c = (\varepsilon/p)_{\text{min}}. \tag{4}$$

We therefore require an expression for the momentum p of the vortex ring. In analogy with the classical vortex ring, we may take

$$p = 2\pi^2 \hbar n_0 r^2, \tag{5}$$

which seems satisfactory because (i) it conforms to the general result: $v = (\partial\varepsilon/\partial p)$, though only to a first approximation, and (ii) it leads to the (approximate) dispersion relation: $\varepsilon \propto p^{1/2}$, which has been separately verified by Rayfield and Reif by subjecting their rings to a transverse electric field. Substituting (3) and (5) into (4), we obtain

$$v_c \sim \left\{ \frac{\hbar}{mr} \ln(r/l) \right\}_{\text{min}} \tag{6}$$

Now, since the r-dependence of the quantity ε/p arises mainly from the factor $1/r$, the minimum in (6) will obtain when r has its largest value, namely R, the radius of the capillary. We thus obtain

$$v_c \sim \frac{\hbar}{mR} \ln(R/l), \tag{7}$$

which is very much the same as the original estimate of Feynman–with D replaced by R. Naturally, then, the numerical values of v_c obtained from the new expression (7) continue to be significantly larger than the corresponding experimental values; however, the theory is now much better founded.

Fetter (1963) was the first to account for the fact that, as the radius r of the ring approaches the radius R of the capillary, the influence of the "image vortex" becomes important. The energy of the flow now falls below the asymptotic value given by (3) by a factor of 10 or so which, in turn, reduces the critical velocity by a similar factor. The actual result obtained by Fetter was

$$v_c \simeq \frac{11}{24} \frac{\hbar}{mR} = 0.46 \frac{\hbar}{mR}. \tag{8}$$

Kawatra and Pathria (1966) extended Fetter's calculation by taking into account the boundary effects arising explicitly from the walls of the capillary as well as the ones arising implicitly from the "image vortex"; moreover, in the computation of ε, they employed actual wave functions, obtained by solving eqn. (10.5.23), rather than the analytical approximation employed by Fetter. They obtained

$$v_c \simeq 0.59 \frac{\hbar}{mR}, \tag{9}$$

which is about 30 per cent higher than Fetter's value; for comments regarding the "most favorable" location for the formation of the vortex ring in the capillary, see the original reference.

1.7. Low-lying states of an imperfect Fermi gas

The Hamiltonian of the quantized field for spin-half fermions ($\sigma = +\frac{1}{2}$ or $-\frac{1}{2}$) is given by eqn. (10.1.48), viz.

$$\hat{H} = \sum_{p,\sigma} \frac{p^2}{2m} a^\dagger_{p\sigma} a_{p\sigma} + \frac{1}{2} \sum{}' u^{p_1'\sigma_1', p_2'\sigma_2'}_{p_1\sigma_1, p_2\sigma_2} a^\dagger_{p_1'\sigma_1'} a^\dagger_{p_2'\sigma_2'} a_{p_2\sigma_2} a_{p_1\sigma_1}, \tag{1}$$

where the matrix elements u are related to the scattering length a of the two-body interaction; the summation in the second part of this expression goes only over those states (of the two particles) that conform to the principles of momentum and spin conservation. As in the Bose case, the matrix elements u in the second sum may be replaced by their values at $p = 0$, i.e.

$$u^{p_1'\sigma_1', p_2'\sigma_2'}_{p_1\sigma_1, p_2\sigma_2} \simeq u^{0\sigma_1', 0\sigma_2'}_{0\sigma_1, 0\sigma_2}. \tag{2}$$

Then, in view of the antisymmetric character of the product $a_{p_1\sigma_1} a_{p_2\sigma_2}$, see eqn. (10.1.24c), all those terms (of the second sum) which contain identical indices σ_1 and σ_2 vanish on summation over p_1 and p_2. Similarly, all those terms which contain identical indices σ_1' and σ_2' vanish on summation over p_1' and p_2'.[17] Thus, for a given set of values of the particle momenta, the only choices of the spin components remaining in the sum are

(i) $\sigma_1 = +\frac{1}{2}, \quad \sigma_2 = -\frac{1}{2}; \quad \sigma_1' = +\frac{1}{2}, \quad \sigma_2' = -\frac{1}{2}$
(ii) $\sigma_1 = +\frac{1}{2}, \quad \sigma_2 = -\frac{1}{2}; \quad \sigma_1' = -\frac{1}{2}, \quad \sigma_2' = +\frac{1}{2}$
(iii) $\sigma_1 = -\frac{1}{2}, \quad \sigma_2 = +\frac{1}{2}; \quad \sigma_1' = -\frac{1}{2}, \quad \sigma_2' = +\frac{1}{2}$
(iv) $\sigma_1 = -\frac{1}{2}, \quad \sigma_2 = +\frac{1}{2}; \quad \sigma_1' = +\frac{1}{2}, \quad \sigma_2' = -\frac{1}{2}$.

It is not difficult to see that the contribution arising from choice (i) will be identically equal to the one arising from choice (iii), while the contribution arising from choice (ii) will be identically equal to the one arising from choice (iv). We may, therefore, write

$$\hat{H} = \sum_{p,\sigma} \frac{p^2}{2m} a^\dagger_{p\sigma} a_{p\sigma} + \frac{u_0}{V} \sum{}' a^\dagger_{p_1'+} a^\dagger_{p_2'-} a_{p_2-} a_{p_1+}, \tag{3}$$

where

$$\frac{u_0}{V} = \left(u^{0+,0-}_{0+,0-} - u^{0-,0+}_{0+,0-}\right), \tag{4}$$

while the indices + and − denote the spin states $\sigma = +\frac{1}{2}$ and $\sigma = -\frac{1}{2}$, respectively; the summation in the second part of (3) now goes over all momenta that conform to the conservation law

$$p_1' + p_2' = p_1 + p_2. \tag{5}$$

To evaluate the eigenvalues of the Hamiltonian (3) we shall employ the techniques of the *perturbation theory.*

First of all, we note that the main term in the expression for $\hat{H}$ is already diagonal, and its eigenvalues are

$$E^{(0)} = \sum_{p,\sigma} \frac{p^2}{2m} n_{p\sigma}, \tag{6}$$

where $n_{p\sigma}$ is the occupation number of the single-particle state (p, σ); its mean value, in equilibrium, is given by the Fermi distribution function

$$\overline{n}_{p\sigma} = \frac{1}{z_0^{-1} \exp(p^2/2mkT) + 1}. \tag{7}$$

The sum in (6) may be replaced by an integral, with the result (see Sec. 8.1, with $g = 2$)

$$E^{(0)} = V \frac{3kT}{\lambda^3} f_{5/2}(z_0), \tag{8}$$

where λ is the mean thermal wavelength of the particles,

$$\lambda = h/(2\pi mkT)^{1/2}, \tag{9}$$

while $f_\nu(z_0)$ is the Fermi–Dirac function

$$f_\nu(z_0) = \frac{1}{\Gamma(\nu)} \int_0^\infty \frac{x^{\nu-1}\, dx}{z_0^{-1} e^x + 1} = \sum_{l=1}^{\infty} (-1)^{l-1} \frac{z_0^l}{l^\nu}; \tag{10}$$

the ideal-gas fugacity z_0 is determined by the total number of particles in the system:

$$N = \sum_{p,\sigma} n_{p\sigma} = V \frac{2}{\lambda^3} f_{3/2}(z_0). \tag{11}$$

The *first-order* correction to the energy of the system is given by the diagonal elements of the interaction term, namely the ones for which $p_1' = p_1$ and $p_2' = p_2$; thus

$$E^{(1)} = \frac{u_0}{V} \sum_{p_1, p_2} n_{p_1 +} n_{p_2 -} = \frac{u_0}{V} N^+ N^-, \tag{12}$$

where N^+ (N^-) denotes the total number of particles with spin up (down). Substituting the *equilibrium* values $\overline{N^+} = \overline{N^-} = \frac{1}{2}N$, we obtain

$$E^{(1)} = \frac{u_0}{4V} N^2 = V \frac{u_0}{\lambda^6} \{f_{3/2}(z_0)\}^2 \tag{13}$$

Substituting $u_0 \simeq 4\pi a\hbar^2/m$, see eqn. (10.1.51), we obtain to *first order* in a

$$E_1^{(1)} = \frac{\pi a\hbar^2}{m} \frac{N}{V} N = V \frac{2kT}{\lambda^3} \left(\frac{a}{\lambda}\right) \{f_{3/2}(z_0)\}^2 \tag{14}$$

The *second-order* correction to the energy of the system can be obtained with the help of the formula

$$E_n^{(2)} = \sum_{m \neq n} \frac{|V_{nm}|^2}{E_n - E_m}, \tag{15}$$

where the indices n and m pertain to the unperturbed states of the system. A simple calculation yields

$$E^{(2)} = 2\frac{u_0^2}{V^2}\sum_{p_1,p_2,p_1'}{}' \frac{n_{p_1+}n_{p_2-}(1-n_{p_1'+})(1-n_{p_2'-})}{(p_1^2+p_2^2-p_1'^2-p_2'^2)/2m}, \tag{16}$$

where the summation goes over *all* p_1, p_2 and p_1' (the value of p_2' being fixed by the requirement of momentum conservation); it is understood that we do not include in the sum (16) any terms for which $p_1^2+p_2^2 = p_1'^2+p_2'^2$. It will be noted that the numerator of the summand in (16) is closely related to the fact that the squared matrix element for the transition $(p_1, p_2) \to (p_1', p_2')$ is directly proportional to the probability that "the states p_1 and p_2 are occupied and at the same time the states p_1' and p_2' are unoccupied". Now, expression (16) does not in itself exhaust terms of *second order* in a. A contribution of the same order of magnitude arises from expression (12) if for u_0 we employ an expression more accurate than the one just employed. The desired expression can be obtained in the same manner as in the Bose case; check the steps leading to eqns (10.2.5,6). In the present case, we obtain

$$\frac{4\pi a\hbar^2}{mV} \simeq \frac{u_0}{V} + 2\frac{u_0^2}{V^2}\sum_{p_1,p_2,p_1'}{}' \frac{1}{(p_1^2+p_2^2-p_1'^2-p_2'^2)/2m},$$

whence it follows that

$$u_0 \simeq \frac{4\pi a\hbar^2}{m}\left[1 - \frac{8\pi a\hbar^2}{mV}\sum_{p_1,p_2,p_1'}{}' \frac{1}{(p_1^2+p_2^2-p_1'^2-p_2'^2)/2m}\right] \tag{17}$$

Substituting (17) into (12), we obtain, apart from the first-order term already given in (14), a second as in eqn. (18), (19) + (20) order term, namely

$$E_2^{(1)} = -2\left(\frac{4\pi a\hbar^2}{mV}\right)^2 \sum_{p_1,p_2,p_1'}{}' \frac{n_{p_1+}n_{p_2-}}{(p_1^2+p_2^2-p_1'^2-p_2'^2)/2m}. \tag{18}$$

For the second-order term as given in (16), the approximation $u_0 \simeq 4\pi a\hbar^2/m$ suffices, with the result

$$E_2^{(2)} = 2\left(\frac{4\pi a\hbar^2}{mV}\right)^2 \sum_{p_1,p_2,p_1'}{}' \frac{n_{p_1+}n_{p_2-}(1-n_{p_1'+})(1-n_{p_2'-})}{(p_1^2+p_2^2-p_1'^2-p_2'^2)/2m}. \tag{19}$$

Combining (18) and (19), we obtain[18]

$$E_2 = E_2^{(1)} + E_2^{(2)} = -2\left(\frac{4\pi a\hbar^2}{mV}\right)^2 \sum_{p_1,p_2,p_1'}{}' \frac{n_{p_1+}n_{p_2-}(n_{p_1'+} + n_{p_2'-})}{(p_1^2+p_2^2-p_1'^2-p_2'^2)/2m} \tag{20}$$

To evaluate (20), we prefer to write it as a *symmetrical* summation over the four momenta p_1, p_2, p_1' and p_2' by introducing a Kronecker delta to take care of the momentum conservation; thus

$$E_2 = -2\left(\frac{4\pi a\hbar^2}{mV}\right)^2 \sum_{p_1\cdot p_2\cdot p'_1\cdot p'_2} \frac{n_{p_1+}n_{p_2-}(n_{p'_1+} + n_{p'_2-})\delta_{p_1+p_2,p'_1+p'_2}}{(p_1^2+p_2^2-p_1'^2-p_2'^2)/2m}. \tag{21}$$

It is obvious that the two parts of the sum (21), one arising from the factor $n_{p'_1+}$ and the other from the factor $n_{p'_2-}$, would give identical results on summation. We may therefore write

$$E_2 = -4\left(\frac{4\pi a\hbar^2}{mV}\right)^2 \sum_{p_1\cdot p_2, p'_1\cdot p'_2} \frac{n_{p_1+}n_{p_2-}n_{p'_1+}\delta_{p_1+p_2,p'_1+p'_2}}{(p_1^2+p_2^2-p_1'^2-p_2'^2)/2m}. \tag{22}$$

The sum in (22) can be evaluated by following a procedure due to Huang, Yang and Luttinger (1957), with the result[19]

$$E_2 = V\frac{8kT}{\lambda^3}\left(\frac{a^2}{\lambda^2}\right)F(z_0), \tag{23}$$

where

$$F(z_0) = -\sum_{r,s,t=1}^{\infty} \frac{(-z_0)^{r+s+t}}{\surd(rst)(r+s)(r+t)}. \tag{24}$$

Combining (8), (14) and (23), we obtain to *second order* in a

$$E = V\frac{kT}{\lambda^3}\left[3f_{5/2}(z_0)+\frac{2a}{\lambda}\{f_{3/2}(z_0)\}^2+\frac{8a^2}{\lambda^2}F(z_0)\right], \tag{25}$$

where z_0 is determined by (11).

It is now straightforward to obtain the *ground state energy* of the imperfect Fermi gas ($z_0 \to \infty$); we have only to know the asymptotic behavior of the functions involved. For the $f\nu$ we have from the Sommerfeld lemma (see Appendix E)

$$f_\nu(z_0) \approx (\ln z_0)^\nu/\Gamma(\nu+1), \tag{26}$$

so that

$$f_{5/2}(z_0) \approx \frac{8}{15\pi^{1/2}}(\ln z_0)^{5/2}; \quad f_{3/2}(z_0) \approx \frac{4}{3\pi^{1/2}}(\ln z_0)^{3/2}. \tag{27}$$

Equation (11) then gives

$$n = \frac{N}{V} \approx \frac{8}{3\pi^{1/2}\lambda^3}(\ln z_0)^{3/2}, \tag{28}$$

so that

$$\ln z_0 \approx \lambda^2\left(\frac{3\pi^{1/2}n}{8}\right)^{2/3}. \tag{29}$$

The asymptotic behaviour of $F(z_0)$ is given by

$$F(z_0) \approx \frac{16(11-2\ln 2)}{105\pi^{3/2}}(\ln z_0)^{7/2}; \tag{30}$$

see Problem 10.12. Substituting (27) and (30) into (25), and making use of the relation (29), we finally obtain

$$\frac{E_0}{N} = \frac{3}{10}\frac{\hbar^2}{m}(3\pi^2 n)^{2/3} + \frac{\pi a\hbar^2}{m} n\left\{1 + \frac{6}{35}(11 - 2\ln 2)\left(\frac{3}{\pi}\right)^{1/3} n^{1/3} a\right\}. \tag{31}$$

The ground state pressure of the gas is then given by

$$\begin{aligned} P_0 &= n^2 \frac{\partial(E_0/N)}{\partial n} \\ &= \frac{1}{5}\frac{\hbar^2}{m}(3\pi^2 n)^{2/3} n + \frac{\pi a\hbar^2}{m} n^2\left\{1 + \frac{8}{35}(11 - 2\ln 2)\left(\frac{3}{\pi}\right)^{1/3} n^{1/3} a\right\}. \end{aligned} \tag{32}$$

We may also calculate the velocity of sound, which directly involves the compressibility of the system, with the result

$$\begin{aligned} c_0^2 &= \frac{\partial P_0}{\partial(mn)} \\ &= \frac{1}{3}\frac{\hbar^2}{m^2}(3\pi^2 n)^{2/3} + \frac{2\pi a\hbar^2}{m^2} n\left\{1 + \frac{4}{15}(11 - 2\ln 2)\left(\frac{3}{\pi}\right)^{1/3} n^{1/3} a\right\}. \end{aligned} \tag{33}$$

The leading terms of the foregoing expressions represent the ground-state results for an ideal Fermi gas, while the remaining terms represent corrections arising from the interparticle interactions.

The result embodied in eqn. (31) was first obtained by Huang and Yang (1957) by the method of pseudopotentials; Martin and De Dominicis (1957) were the first to attempt an estimate of the third-order correction.[20] Lee and Yang (1957) obtained (31) on the basis of the binary collision method; for the details of their calculation, see Lee and Yang (1959b, 1960a). The same result was derived later by Galitskii (1958) who employed the method of Green's functions.

10.8. Energy spectrum of a Fermi liquid: Landau's phenomenological theory[21]

In Sec. 10.4 we discussed the main features of the energy spectrum of a Bose liquid; such a spectrum is generally referred to as a *Bose type* spectrum. A liquid consisting of spin-half fermions, such as liquid He^3, is expected to have a different kind of spectrum which, by contrast, may be called a *Fermi type* spectrum. Right away we should emphasize that a liquid consisting of fermions may not necessarily possess a spectrum of the Fermi type; the spectrum actually possessed by such a liquid depends crucially on the nature of the interparticle interactions operating in the liquid. For instance, if the interparticle interactions are such that they tend to associate particles into "pairs", then in a particular limit we may expect to obtain a "molecular" liquid whose constituents possess "integral" spin and hence a "Bose type" spectrum. This indeed happens with the electron gas in superconductors which, at low temperatures, displays a remarkable preference for the formation of the so-called "Cooper pairs" of electrons. The formation of these pairs is essentially due to the electron–phonon–electron interaction operating in the material and lends to the electron gas the same properties of "superfluidity" as are witnessed in the case of liquid He^4 (which, as we know, is a Bose liquid). For the details of the mechanism of pair formation, reference may be made to the original papers of Cooper (1956) and of Bardeen, Cooper and Schrieffer (1957); a simpler version of the theory is given in Cooper (1960). In the present section we propose to discuss the main features of a spectrum which is characteristically of the Fermi type.

According to Landau (1956), whose work provides the basic framework for our discussion, the Fermi type spectrum of a quantum liquid is constructed in analogy with the spectrum of an ideal Fermi gas. As is well known, the ground state of the ideal Fermi gas corresponds to a "complete filling up of the single-particle states with $p \leq p_F$ and a complete absence of particles in the states with $p > p_F$"; the excitation of the system corresponds to a transition of one or more particles from the occupied states to the unoccupied states. The limiting momentum p_F is related to the particle density in the system and for spin-half particles is given by

$$p_F = \hbar(3\pi^2 N/V)^{1/3} \tag{1}$$

In a liquid, we cannot speak of quantum states of *individual* particles. However, as a basis for constructing the desired spectrum, we may assume that as interparticle interactions are gradually "switched on" and a transition made from the gaseous to the liquid state, the ordering of the energy levels (in the momentum space) remains unchanged. Of course, in this ordering, the role of the gas particles is passed on to the "elementary excitations" of the liquid (also referred to as "quasi-particles") whose number coincides with the number of particles in the liquid and which also obey Fermi statistics. Each "quasi-particle" possesses a definite momentum $\boldsymbol{p}$, so we can speak of a *distribution function* $n(\boldsymbol{p})$ such that

$$\int n(\boldsymbol{p})\, d\tau = N/V, \tag{2}$$

where $d\tau = 2\,d^3p/h^3$. We then expect that the specification of the function $n(\boldsymbol{p})$ uniquely determines the total energy E of the liquid. Of course, E will not be given by a simple sum of the energies $\varepsilon(\boldsymbol{p})$ of the quasi-particles; it will rather be a *functional* of the distribution function $n(\boldsymbol{p})$. In other words, the energy E will not reduce to the simple integral $\int \varepsilon(\boldsymbol{p}) n(\boldsymbol{p}) V\, d\tau$, though in the first approximation a *variation* in its value may be written as

$$\delta E = V \int \varepsilon(\boldsymbol{p}) \delta n(\boldsymbol{p})\, d\tau, \tag{3}$$

where $\delta n(\boldsymbol{p})$ is an *assumed variation* in the distribution function of the "quasi-particles". The reason why E does not reduce to an integral of the quantity $\varepsilon(\boldsymbol{p}) n(\boldsymbol{p})$ is related to the fact that the quantity $\varepsilon(\boldsymbol{p})$ is itself a functional of the distribution function. If the initial distribution function is a step function (which corresponds to the ground state of the system), then the variation in $\varepsilon(\boldsymbol{p})$ due to a *small* deviation of the distribution function from the step function (which implies only *low-lying* excited states of the system) would be given by a *linear* functional relationship:

$$\delta\varepsilon(\boldsymbol{p}) = \int f(\boldsymbol{p}, \boldsymbol{p}') \delta n(\boldsymbol{p}')\, d\tau'. \tag{4}$$

Thus, the quantities $\varepsilon(\boldsymbol{p})$ and $f(\boldsymbol{p}, \boldsymbol{p}')$ are the first and second *functional derivatives* of E with respect to $n(\boldsymbol{p})$. Inserting spin dependence, we may now write

$$\delta E = \sum_{p,\sigma} \varepsilon(p,\sigma)\delta n(p,\sigma) + \frac{1}{2V}\sum_{p,\sigma;\,p',\sigma'} f(p,\sigma;p',\sigma')\delta n(p,\sigma)\delta n(p',\sigma'), \tag{5}$$

where δn are *small* variations in the distribution function $n(p)$ from the step function (that characterizes the ground state of the system); it is obvious that these variations will be significant only in the vicinity of the limiting momentum p_F, which continues to be given by eqn. (1). It is thus understood that the quantity $\varepsilon(p,\sigma)$ in (5) corresponds to the distribution function $n(p,\sigma)$ being infinitesimally close to the step function (of the ground state). One may also note that the function $f(p,\sigma;\, p',\sigma')$, being a second functional derivative of E, must be symmetric in its arguments; often, it is of the form $a + b\ \hat{s}_1 \cdot \hat{s}_2$, where the coefficients a and b depend only upon the angle between the momenta p and p'.[22] The function f plays a central role in the theory of the Fermi liquid; for an ideal gas, f vanishes.

To discover the formal dependence of the distribution function $n(p)$ on the quantity $\varepsilon(p)$, we note that, in view of the one-to-one correspondence between the energy levels of the liquid and of the ideal gas, the number of microstates (and hence the entropy) of the liquid is given by the same expression as for the ideal gas; see eqn. (6.1.15), with all $g_i = 1$ and $a = +1$, or Problem 6.1:

$$\frac{S}{k} = -\sum_p \{n \ln n + (1-n)\ln(1-n)\} \approx -V\int\{n\ln n + (1-n)\ln(1-n)\}\,d\tau. \tag{6}$$

Maximizing this expression, under the constraints $\delta E = 0$ and $\delta N = 0$, we obtain for the *equilibrium distribution function*

$$\bar{n} = \frac{1}{\exp\{(\varepsilon - \mu)/kT\} + 1}. \tag{7}$$

It should be noted here that, despite its formal similarity with the standard expression for the Fermi–Dirac distribution function, formula (7) is different insofar as the quantity ε appearing here is itself a function of $\bar{n}$; consequently, this formula gives only an *implicit*, and probably a very complicated, expression for the function $\bar{n}$.

A word may be said about the quantity ε appearing in eqn. (5). Since this ε corresponds to the *limiting case* of n being a step function, it is expected to be a well-defined function of p. Equation (7) then reduces to the usual Fermi–Dirac distribution function, which is indeed an *explicit* function of ε. It is not difficult to see that this reduction remains valid so long as expression (5) is valid, i.e. so long as the variations δn are small, which in turn means that $T \ll T_F$. As mentioned earlier, the variation δn will be significant only in the vicinity of the Fermi momentum p_F; accordingly, we will not have much to do with the function $\varepsilon(p)$ except when $p \simeq p_F$. We may, therefore, write

$$\varepsilon(p \simeq p_F) = \varepsilon_F + \left(\frac{\partial\varepsilon}{\partial p}\right)_{p=p_F}(p - p_F) + \cdots \simeq \varepsilon_F + u_F(p - p_F), \tag{8}$$

where u_F denotes the "velocity" of the quasi-particles at the Fermi surface. In the case of an ideal gas ($\varepsilon = p^2/2m$), $u_F = p_F/m$. By analogy, we define a parameter m^* such that

$$m^* \equiv \frac{p_F}{u_F} = \frac{p_F}{(\partial\varepsilon/\partial p)_{p=p_F}} \tag{9}$$

and call it the *effective mass* of the quasi-particle with momentum p_F (or with $p \simeq p_F$). Another way of looking at the parameter m^* is due to Brueckner and Gammel (1958), who wrote

$$\varepsilon(p \simeq p_F) = \frac{p^2}{2m} + V(p) = \frac{p^2}{2m^*} + \text{const.}; \tag{10}$$

the philosophy behind this expression is that "for quasi-particles with $p \simeq p_F$, the modification, $V(p)$, brought into the quantity $\varepsilon(p)$ by the presence of interparticle interactions in the liquid may be represented by a constant term while the kinetic energy, $p^2/2m$, is modified so as to replace the particle mass m by an effective, quasi-particle mass m^*"; in other words, we adopt a *mean field* point of view. Differentiating (10) with respect to p and setting $p = p_F$, we obtain

$$\frac{1}{m^*} = \frac{1}{m} + \frac{1}{p_F}\left(\frac{dV(p)}{dp}\right)_{p=p_F}. \tag{11}$$

The quantity m^*, in particular, determines the low-temperature specific heat of the Fermi liquid. We can readily see that, for $T \ll T_F$, the ratio of the specific heat of a Fermi liquid to that of an ideal Fermi gas is precisely equal to the ratio m^*/m:

$$\frac{(C_V)\text{ real}}{(C_V)\text{ ideal}} = \frac{m^*}{m}. \tag{12}$$

This follows from the fact that (i) expression (6) for the entropy S, in terms of the distribution function n, is the same for the liquid as for the gas, (ii) the same is true of relation (7) between $\bar{n}$ and ε, and (iii) for the evaluation of the integral in (6) *at low temperatures* only momenta close to p_F are important. Consequently, the result stated in Problem 8.13, namely

$$C_V \simeq S \simeq \frac{\pi^2}{3}k^2 T a(\varepsilon_F), \tag{13}$$

continues to hold–with the sole difference that in the expression for the density of states $a(\varepsilon_F)$, in the vicinity of the Fermi surface, the particle mass m gets replaced by the effective mass m^*; see eqn. (8.1.21).

We now proceed to establish a relationship between the parameters m and m^* in terms of the characteristic function f. In doing so, we neglect the spin-dependence of f, if any; the necessary modification can be introduced without much difficulty. The guiding principle here is that, in the absence of external forces, the momentum density of the liquid must be equal to the density of mass transfer. The former is given by $\int pn\,d\tau$, while the latter is given by $m\int(\partial\varepsilon/\partial p)n\,d\tau$, $(\partial\varepsilon/\partial p)$ being the "velocity" of the quasi-particle with momentum p and energy ε.[23] Thus

$$\int pn\,d\tau = m\int \frac{\partial\varepsilon}{\partial p} n\,d\tau. \tag{14}$$

Varying the distribution function by δn and making use of eqn. (4), we obtain

$$\int p\delta n\,d\tau = m\int \frac{\partial\varepsilon}{\partial p}\delta n\,d\tau + m\int\int\left\{\frac{\partial f(p,p')}{\partial p}\delta n'\,d\tau'\right\}n\,d\tau$$

$$= m\int \frac{\partial \varepsilon}{\partial p}\delta n\, d\tau - m\iint f(p,p')\frac{\partial n'}{\partial p'}\delta n\, d\tau\, d\tau'; \tag{15}$$

in obtaining the last expression, we have interchanged the variables p and p' and have also carried out an integration *by parts*. In view of the arbitrariness of the variation δn, eqn. (15) requires that

$$\frac{p}{m} = \frac{\partial \varepsilon}{\partial p} - \int f(p,p')\frac{\partial n'}{\partial p'}\, d\tau'. \tag{16}$$

We apply this result to quasi-particles with momenta close to p_F; at the same time, we replace the distribution function n' by a "step" function, whereby

$$\frac{\partial n'}{\partial p'} = -\frac{p'}{p'}\delta(p' - p_F).$$

This enables us to carry out integration over the magnitude p' of the momentum, so that

$$\int f(p,p')\frac{\partial n'}{\partial p'}\frac{2p'^2\, dp'\, d\omega'}{h^3} = -\frac{2p_F}{h^3}\int f(\theta)p'_F\, d\omega', \tag{17}$$

$d\omega'$ being the element of a solid angle; note that we have contracted the arguments of the function f because in simple situations it depends only upon the angle between the two momenta. Inserting (17) into (16), with $p = p_F$, making a scalar product with p_F and dividing by p_F^2, we obtain the desired result

$$\frac{1}{m} = \frac{1}{m^*} + \frac{p_F}{2h^3}\cdot 4\int f(\theta)\cos\theta\, d\omega'. \tag{18}$$

If the function f depends upon the spins s_1 and s_2 of the particles involved, then the factor 4 in front of the integral will have to be replaced by a summation over the spin variables.

We now derive a formula for the velocity of sound at absolute zero. From first principles, we have[24]

$$c_0^2 = \frac{\partial P_0}{\partial (mN/V)} = -\frac{V^2}{mN}\left(\frac{\partial P_0}{\partial V}\right)_N$$

In the present context, it is preferable to have an expression in terms of the chemical potential of the liquid. This can be obtained by making use of the formula $N d\mu_0 = V dP_0$, see Problem 1.16, whence it follows that[25]

$$\left(\frac{\partial \mu_0}{\partial N}\right)_V = -\frac{V}{N}\left(\frac{\partial \mu_0}{\partial V}\right)_N = -\frac{V^2}{N^2}\left(\frac{\partial P_0}{\partial V}\right)_N$$

and hence

$$c_0^2 = \frac{N}{m}\left(\frac{\partial \mu_0}{\partial N}\right)_V \tag{19}$$

Now, $\mu_0 = \varepsilon(p_F) = \varepsilon_F$; therefore, the change $\delta\mu_0$ arising from a change δN in the total number of particles in the system is given by

$$\delta\mu_0 = \frac{\partial \varepsilon_F}{\partial p_F}\delta p_F + \int f(p_F,p')\delta n'\, d\tau'. \tag{20}$$

The first part in (20) arises from the fact that a change in the total number of particles in the system inevitably alters the value of the limiting momentum p_F; see eqn. (1), whence (for constant V)

$$\delta p_F/p_F = \tfrac{1}{3}\delta N/N$$

and hence

$$\frac{\partial\varepsilon_F}{\partial p_F}\delta p_F = \frac{p_F^2}{3m^*}\frac{\delta N}{N}. \tag{21}$$

The second part arises from eqn. (4). It will be noted that the variation $\delta n'$ appearing in the integral of eqn. (20) is significant only for $p' \simeq p_F$; we may, therefore, write

$$\int f(p_F, p')\delta n'\,d\tau' \simeq \frac{\delta N}{4\pi V}\int f(\theta)\,d\omega'. \tag{22}$$

Substituting (21) and (22) into (20), we obtain

$$\left(\frac{\partial\mu_0}{\partial N}\right)_V = \frac{p_F^2}{3m^*N} + \frac{1}{4\pi V}\int f(\theta)\,d\omega'. \tag{23}$$

Making use of eqns (18) and (1), we finally obtain

$$c_0^2 = \frac{N}{m}\left(\frac{\partial\mu_0}{\partial N}\right)_V = \frac{p_F^2}{3m^2} + \frac{p_F^3}{6m\hbar^3}\cdot 4\int f(\theta)(1-\cos\theta)\,d\omega'. \tag{24}$$

Once again, if the function f depends upon the spins of the particles, then the factor 4 in front of the integral will have to be replaced by a summation over the spin variables.

For illustration, we shall apply this theory to the imperfect Fermi gas studied in Sec. 10.7. To calculate $f(p,\sigma;p',\sigma')$, we have to differentiate twice the sum of expression (10.7.12), with $u_0 = 4\pi a\hbar^2/m$, and expression (10.7.22) with respect to the distribution function $n(p,\sigma)$ and then substitute $p = p' = p_F$. Performing the desired calculation, then changing summations into integrations and carrying out integrations by simple means, we find that the function f is spin-dependent–the spin-dependent term being in the nature of an *exchange term*, proportional to $\hat{s}_1\cdot\hat{s}_2$. The complete result, according to Abrikosov and Khalatnikov (1957), is

$$f(p,\sigma;p',\sigma') = A(\theta) + B(\theta)\hat{s}_1\cdot\hat{s}_2, \tag{25}$$

where

$$A(\theta) = \frac{2\pi a\hbar^2}{m}\left[1 + 2a\left(\frac{3N}{\pi V}\right)^{1/3}\left\{2 + \frac{\cos\theta}{2\sin(\theta/2)}\ln\frac{1+\sin(\theta/2)}{1-\sin(\theta/2)}\right\}\right]$$

and

$$B(\theta) = -\frac{8\pi a\hbar^2}{m}\left[1 + 2a\left(\frac{3N}{\pi V}\right)^{1/3}\left\{1 - \frac{1}{2}\sin\left(\frac{\theta}{2}\right)\ln\frac{1+\sin(\theta/2)}{1-\sin(\theta/2)}\right\}\right],$$

a being the scattering length of the two-body potential and θ the angle between the momentum vectors p_F and p'_F. Substituting (25) into formulae (18) and (24), in which the factor 4 is now supposed to be replaced by a summation over the spin variables, we find that while the spin-dependent term $B(\theta)\hat{s}_1 \cdot \hat{s}_2$ does not make any contribution towards the final results, the spin-independent term $A(\theta)$ leads to[26]

$$\frac{1}{m^*} = \frac{1}{m} - \frac{8}{15m}(7 \ln 2 - 1)\left(\frac{3N}{\pi V}\right)^{2/3} a^2 \tag{26}$$

and

$$c_0^2 = \frac{p_F^2}{3m^2} + \frac{2\pi a\hbar^2}{m^2}\frac{N}{V}\left[1 + \frac{4}{15}(11 - 2\ln 2)\left(\frac{3N}{\pi V}\right)^{1/3} a\right]; \tag{27}$$

the latter result is identical with expression (10.7.33) derived in the preceding section. Proceeding backward, one can obtain from eqn. (27) corresponding expressions for the ground-state pressure P_0 and the ground-state energy E_0, namely eqns (10.7.32) and (10.7.31), as well as the ground-state chemical potential μ_0, as quoted in Problem 10.15.

Problems

10.1. (a) Show that, for bosons as well as fermions,

$$[\psi(\boldsymbol{r}_j), \hat{H}] = \left(-\frac{\hbar^2}{2m}\nabla_j^2 + \int d^3r\, \psi^\dagger(\boldsymbol{r})u(\boldsymbol{r}, \boldsymbol{r}_j)\psi(\boldsymbol{r})\right)\psi(\boldsymbol{r}_j),$$

where $\hat{H}$ is the Hamiltonian operator defined by eqn. (10.1.4).

(b) Making use of the foregoing result, show that the equation

$$\frac{1}{\sqrt{N!}}\langle 0|\psi(\boldsymbol{r}_1)\ldots\psi(\boldsymbol{r}_N)\hat{H}|\Psi_{NE}\rangle = E\frac{1}{\sqrt{N!}}\langle 0|\psi(\boldsymbol{r}_1)\ldots\psi(\boldsymbol{r}_N)|\Psi_{NE}\rangle$$
$$= E\Psi_{NE}(\boldsymbol{r}_1, \ldots \boldsymbol{r}_N)$$

is equivalent to the Schrödinger equation (10.1.15).

10.2. The grand partition function of a gaseous system composed of mutually interacting bosons is given by

$$\ln \mathcal{Q} \equiv \frac{PV}{kT} = \frac{V}{\lambda^3}\left[g_{5/2}(z) - 2\{g_{3/2}(z)\}^2\frac{a}{\lambda} + O\left(\frac{a^2}{\lambda^2}\right)\right].$$

Study the analytic behavior of this expression near $z = 1$ and show that the system exhibits the phenomenon of Bose–Einstein condensation when its fugacity assumes the critical value

$$z_c = 1 + 4\zeta\left(\frac{3}{2}\right)\frac{a}{\lambda_c} + O\left(\frac{a^2}{\lambda_c^2}\right)$$

Further show that the pressure of the gas at the critical point is given by (Lee and Yang 1958, 1960b)

$$\frac{P_c}{kT_c} = \frac{1}{\lambda_c^3}\left[\zeta\left(\frac{5}{2}\right) + 2\left\{\zeta\left(\frac{3}{2}\right)\right\}^2\frac{a}{\lambda_c} + O\left(\frac{a^2}{\lambda_c^2}\right)\right];$$

cf. eqns (10.2.16–18).

10.3. For the imperfect Bose gas studied in Sec. 10.2, calculate the specific heat C_V near absolute zero and show that, as $T \to 0$, the specific heat vanishes in a manner characteristic of a system with an "energy gap" $\Delta = 4\pi a\hbar^2 n/m$.

10.4. (a) Show that, to *first* order in the scattering length a, the discontinuity in the specific heat C_V of an imperfect Bose gas at the transition temperature T_c is given by

$$(C_V)_{T=T_c-} - (C_V)_{T=T_c+} = Nk\frac{9a}{2\lambda_c}\zeta(3/2),$$

while the discontinuity in the bulk modulus K is given by

$$(K)_{T=T_c-} - (K)_{T=T_c+} = -\frac{4\pi a\hbar^2}{mv_c^2}.$$

(b) Examine the discontinuities in the quantities $(\partial^2 P/\partial T^2)_v$ and $(\partial^2 \mu/\partial T^2)_v$ as well, and show that your results are consistent with the thermodynamic relationship

$$C_V = VT\left(\frac{\partial^2 P}{\partial T^2}\right)_v - NT\left(\frac{\partial^2 \mu}{\partial T^2}\right)_v$$

10.5. (a) Complete the mathematical steps leading to eqns (10.3.15) and (10.3.16).
(b) Complete the mathematical steps leading to eqns (10.3.23) and (10.3.24).

10.6. The ground-state pressure of an interacting Bose gas turns out to be (see Lee and Yang, 1960a)

$$P_0 = \frac{\mu_0^2 m}{8\pi a\hbar^2}\left[1 - \frac{64}{15\pi}\frac{\mu_0^{1/2} m^{1/2} a}{\hbar} + \cdots\right],$$

where μ_0 is the ground-state chemical potential of the gas. It follows that

$$n \equiv \left(\frac{dP_0}{d\mu_0}\right) = \frac{\mu_0 m}{4\pi a\hbar^2}\left[1 - \frac{16}{3\pi}\frac{\mu_0^{1/2} m^{1/2} a}{\hbar} + \cdots\right]$$

and

$$\frac{E_0}{V} \equiv (n\mu_0 - P_0) = \frac{\mu_0^2 m}{8\pi a\hbar^2}\left[1 - \frac{32}{5\pi}\frac{\mu_0^{1/2} m^{1/2} a}{\hbar} + \cdots\right].$$

Eliminating μ_0, derive eqns (10.3.16) and (10.3.17).

10.7. Show that the mean occupation number $\bar{n}_p$ of the *real* particles and the mean occupation number $\bar{N}_p$ of the *quasi*-particles, in an interacting Bose gas, are connected by the relationship

$$\bar{n}_p = \frac{\bar{N}_p + \alpha_p^2(\bar{N}_p + 1)}{1 - \alpha_p^2} \quad (p \neq 0),$$

where α_p is given by eqns (10.3.9) and (10.3.10).

Note that eqn. (10.3.22) corresponds to the special case $\bar{N}_p = 0$.

10.8. The excitation energy of liquid He^4, carrying a *single* excitation above the ground state, is determined by the *minimum* value of the quantity

$$\varepsilon = \int \Psi^*\left\{-\frac{\hbar^2}{2m}\sum_i \nabla_i^2 + V - E_0\right\}\Psi\, d^{3N}r \Big/ \int \Psi^*\Psi\, d^{3N}r,$$

where E_0 denotes the ground-state energy of the liquid while Ψ, according to Feynman, is given by eqn. (10.4.3). Show that the process of minimization leads to expression (10.4.5) for the energy of excitation.

[*Hint*: First express ε in the form

$$\varepsilon = \frac{\hbar^2}{2m}\int |\nabla f(r)|^2 d^3r \Big/ \int f^*(r_1) f(r_2) g(r_2 - r_1)\, d^3r_1\, d^3r_2.$$

Then show that ε is *minimum* when $f(r)$ is of the form (10.4.4).]

10.9. Show that, for a sufficiently large momentum $\hbar k$ (in fact, such that the slope $d\varepsilon/dk$ of the energy spectrum is greater than the initial slope $\hbar c$), a state of *double excitation* in liquid He^4 is energetically more favorable than a state of *single excitation*, i.e. there exist wave vectors k_1 and k_2 such that, while $k_1 + k_2 = k, \varepsilon(k_1) + \varepsilon(k_2) < \varepsilon(k)$.

10.10. Using Fetter's analytical approximation,

$$f_1(\rho') = \frac{\rho'}{\sqrt{(1 + \rho'^2)}},$$

for the soluti[illegible] f eqn. (10.5.23) with $s = 1$, calculate the energy (per unit length) associated with a quantized vo[illegible]line of unit circulation. Compare your result with the one quoted in (10.5.26).

10.11. (a) Study the nature of the velocity field arising from a pair of parallel vortex lines, with $s_1 = +1$ and $s_2 = -1$, separated by a distance d. Derive and discuss the general equation of the stream lines.

(b) Again, using Fetter's analytical approximation for the functions $f(\rho_1')$ and $f(\rho_2')$, calculate the energy (per unit length) of the system and show that its limiting value, as $d \to 0$, is $11\pi\hbar^2 n_0/12m$. Making use of this result, derive expression (10.6.8) for the critical velocity of superflow.

10.12. Establish the asymptotic formula (10.7.30) for the function $F(z_0)$.
[*Hint*: Write the coefficient that appears in the expansion (10.7.24) in the form

$$\frac{1}{\sqrt{(rst)(r+s)(r+t)}} = \left(\frac{2}{\sqrt{\pi}}\right)^3 \int_0^\infty e^{-X^2 r - Y^2 s - Z^2 t - \xi(r+s) - \eta(r+t)}\, dX\, dY\, dZ\, d\xi\, d\eta.$$

Insert this expression into (10.7.24) and carry out summations over r, s and t, with the result

$$F(z_0) = \frac{8}{\pi^{3/2}} \int_0^\infty \frac{1}{z_0^{-1} e^{X^2+\xi+\eta} + 1} \frac{1}{z_0^{-1} e^{Y^2+\xi} + 1} \frac{1}{z_0^{-1} e^{Z^2+\eta} + 1}\, dX\, dY\, dZ\, d\xi\, d\eta.$$

In the limit $z_0 \to \infty$, the integrand is essentially equal to 1 in the region R defined by

$$X^2 + \xi + \eta < \ln z_0, \quad Y^2 + \xi < \ln z_0 \quad \text{and} \quad Z^2 + \eta < \ln z_0;$$

outside this region, it is essentially 0. Hence, the dominant term of the asymptotic expansion is

$$\frac{8}{\pi^{3/2}} \int_R 1 \cdot dX\, dY\, dZ\, d\xi\, d\eta,$$

which, in turn, reduces to the double integral

$$\frac{8}{\pi^{3/2}} \iint (\ln z_0 - \xi - \eta)^{1/2} (\ln z_0 - \xi)^{1/2} (\ln z_0 - \eta)^{1/2}\, d\xi\, d\eta;$$

the limits of integration here are such that not only $\xi < (\ln z_0)$ and $\eta < (\ln z_0)$ but also $(\xi + \eta) < (\ln z_0)$. The rest of the calculation is straightforward.]

10.13. The grand partition function of a gaseous system composed of mutually interacting, spin-half fermions has been evaluated by Lee and Yang (1957), with the result[27]

$$\ln \mathcal{Q} \equiv \frac{PV}{kT} = \frac{V}{\lambda^3}\left[2 f_{5/2}(z) - \frac{2a}{\lambda}\{f_{3/2}(z)\}^2 + \frac{4a^2}{\lambda^2} f_{1/2}(z)\{f_{3/2}(z)\}^2 - \frac{8a^2}{\lambda^2} F(z) + \cdots\right],$$

where z is the fugacity of the *actual* system (not of the corresponding non-interacting system, which was denoted by z_0 in the text); the functions $f_\nu(z)$ and $F(z)$ are defined in a manner similar to eqns (10.7.10) and (10.7.24). From here, one can derive expressions for the quantities $E(z, V, T)$ and $N(z, V, T)$ by using the formulae

$$E(z, V, T) \equiv kT^2 \frac{\partial(\ln \mathcal{Q})}{\partial T} \quad \text{and} \quad N(z, V, T) \equiv \frac{\partial(\ln \mathcal{Q})}{\partial(\ln z)} \left\{= \frac{2V}{\lambda^3} f_{3/2}(z_0)\right\}.$$

(a) Eliminating z between these two results, derive eqn. (10.7.25) for E.

(b) Obtain the zero-point value of the chemical potential μ, correct to *second* order in (a/λ), and verify, with the help of eqns (10.7.31, 32), that

$$(E + PV)_{T=0} = N(\mu)_{T=0}.$$

[*Hint*: At $T = 0$ K, $\mu = (\partial E/\partial N)_V$.]

(c) Show that the low-temperature specific heat and the low-temperature entropy of the gas are given by (see Pathria and Kawatra, 1962)

$$\frac{C_V}{Nk} \simeq \frac{S}{Nk} \simeq \frac{\pi^2}{2}\left(\frac{kT}{\varepsilon_F}\right)\left[1 + \frac{8}{15\pi^2}(7\ln 2 - 1)(k_F a)^2 + \cdots\right],$$

where $k_F = (3\pi^2 n)^{1/3}$. Clearly, the factor within long brackets is to be identified with the ratio m^*/m; see eqns (10.8.12, 26).

[*Hint*: To determine C_V to the *first* power in T, we must know E to the *second* power in T. For this, we require higher-order terms of the asymptotic expansions of the functions $f_\nu(z)$ and $F(z)$; these are given by

$$f_{5/2}(z) = \frac{8}{15\sqrt{\pi}}(\ln z)^{5/2} + \frac{\pi^{3/2}}{3}(\ln z)^{1/2} + O(1),$$

$$f_{3/2}(z) = \frac{4}{3\sqrt{\pi}}(\ln z)^{3/2} + \frac{\pi^{3/2}}{6}(\ln z)^{-1/2} + O(\ln z)^{-5/2}$$

$$f_{1/2}(z) = \frac{2}{\sqrt{\pi}}(\ln z)^{1/2} - \frac{\pi^{3/2}}{12}(\ln z)^{-3/2} + O(\ln z)^{-7/2}.$$

and

$$F(z) = \frac{16(11 - 2\ln 2)}{105\pi^{3/2}}(\ln z)^{7/2} - \frac{2(2\ln 2 - 1)}{3}\pi^{1/2}(\ln z)^{3/2} + O(\ln z)^{5/4}.]$$

The first three results follow from the Sommerfeld lemma (E. 15); for the last one, see Yang (1962).

10.14. The energy spectrum $\varepsilon(p)$ of mutually interacting, spin-half fermions is given by (Galitskii, 1958; Mohling, 1961)

$$\frac{\varepsilon(p)}{p_F^2/2m} \simeq x^2 + \frac{4}{3\pi}(k_F a) + \frac{4}{15\pi^2}(k_F a)^2$$
$$\times \left[11 + 2x^4 \ln\frac{x^2}{|x^2 - 1|} - 10\left(x - \frac{1}{x}\right)\ln\left|\frac{x+1}{x-1}\right| - \frac{(2 - x^2)^{5/2}}{x}\ln\left(\frac{1 + x\sqrt{(2 - x^2)}}{1 - x\sqrt{(2 - x^2)}}\right)\right],$$

where $x = p/p_F \le \sqrt{2}$ and $k = p/\hbar$. Show that, for k close to k_F, this spectrum reduces to

$$\frac{\varepsilon(p)}{p_F^2/2m} \simeq x^2 + \frac{4}{3\pi}(k_F a) + \frac{4}{15\pi^2}(k_F a)^2\left[(11 - 2\ln 2) - 4(7\ln 2 - 1)\left(\frac{k}{k_F} - 1\right)\right].$$

Using eqns (10.8.10, 11), check that this expression leads to the result

$$\frac{m^*}{m} \simeq 1 + \frac{8}{15\pi^2}(7\ln 2 - 1)(k_F a)^2.$$

10.15. In the ground state of a Fermi system, the chemical potential is identical with the Fermi energy: $(\mu)_{T=0} = \varepsilon(p_F)$. Making use of the energy spectrum $\varepsilon(p)$ of the previous problem, we obtain

$$(\mu)_{T=0} \simeq \frac{p_F^2}{2m}\left[1 + \frac{4}{3\pi}(k_F a) + \frac{4}{15\pi^2}(11 - 2\ln 2)(k_F a)^2\right].$$

Integrating this result, rederive eqn. (10.7.31) for the ground-state energy of the system.

10.16. The energy levels of an imperfect Fermi gas in the presence of an external magnetic field B, to *first* order in a, may be written as

$$E_n = \sum_p (n_p^+ + n_p^-)\frac{p^2}{2m} + \frac{4\pi a\hbar^2}{mV}N^+N^- - \mu^* B(N^+ - N^-);$$

see eqns (8.2.8) and (10.7.12). Using this expression for E_n and following the procedure adopted in Sec. 8.2A, study the magnetic behavior of this gas–in particular, the zero-field susceptibility $\chi(T)$. Also examine the possibility of spontaneous magnetization arising from the interaction term with $a > 0$.

Notes

[1] Alternatively, one may employ eqns (20) and (21), and make (f rules (1) or (2) along with eqn. (17).

[2] This representation of the field is generally referred to as the *particle-number representation*.

[3] This can also be seen by noting that the fermion operators $\hat{N}_\alpha$ satisfy the identity

$$\hat{N}_\alpha^2 = a_\alpha^\dagger a_\alpha a_\alpha^\dagger a_\alpha = a_\alpha^\dagger(1 - a_\alpha^\dagger a_\alpha)a_\alpha = a_\alpha^\dagger a_\alpha = \hat{N}_\alpha \quad (\text{since } a_\alpha^\dagger a_\alpha^\dagger a_\alpha a_\alpha \equiv 0).$$

The same would be true of the eigenvalues n_α. Hence, $n_\alpha^2 = n_\alpha$ which means that $n_\alpha = 0$ or 1.

[4] This result is consistent with the *pseudopotential approach* of Huang and Yang (1957) in which $u(r)$ is replaced by the singular potential $(4\pi a\hbar^2/m)\delta(r)$, so the integral u_0 becomes $4\pi a\hbar^2/m$. For an exposition of the pseudopotential approach, see Chapter 10 of the first edition of this book.

[5] In the last step we have replaced the sum $\sum\limits_p n_p^2$ by the single term n_0^2, neglecting the partial sum $\sum\limits_{p\neq 0} n_p^2$ in comparison with the number $(2N^2 - n_0^2)$. Justification for this lies in the fact that, by the theory of fluctuations, the neglected part will be $O(N)$ and not $O(N^2)$.

[6] This and the subsequent results were derived by Lee and Yang (1958, 1960c) using the binary collision method and by Huang (1959, 1960) using the pseudopotential method.

[7] The evaluation of higher-order terms of this expansion necessitates consideration of three-body collisions as well; hence, in general, they cannot be expressed in terms of the scattering length alone. The exceptional case of a hard-sphere gas has been studied by Wu (1959), who obtained (using the pseudopotential method)

$$\frac{E_0}{N} = \frac{2\pi a\hbar^2 n}{m}\left[1 + \frac{128}{15\pi^{1/2}}(na^3)^{1/2} + 8\left(\frac{4\pi}{3} - \sqrt{3}\right)(na^3)\ln(12\pi na^3) + O(na^3)\right],$$

which shows that the expansion does not proceed in simple powers of $(na^3)^{1/2}$.

[8] Spectrum (1) was first obtained by Bogoliubov (1947) by the method outlined in the preceding sections. Using the pseudopotential method, it was rederived by Lee, Huang and Yang (1957).

[9] The reader interested in pursuing Feynman's line of argument should refer to Feynman's original papers or to a recent review of Feynman's work on superfluidity by Mehra and Pathria (1994).

[10] For a microscopic study of the structure factor $S(k)$, see Huang and Klein (1964); also Jackson and Feenberg (1962).

[11] It is natural that at some value of $k < k_0$, the (ε, k)-curve passes through a *maximum*; this happens when $dS/dk = 2S/k$.

[12] That the vortices in a superfluid may be quantized, the quantum of circulation being h/m, was first suggested by Onsager (1949) in a footnote to a paper dealing with the classical vortex theory and the theory of turbulence!

[13] Drawing upon the well-known analogy between the phenomena of superfluidity and superconductivity, and the resulting correspondence between the mechanical momentum mv_s of a superfluid particle and the electromagnetic momentum $2eA/c$ of a Cooper pair of electrons, we observe that the relevant counterpart of the Landau condition, in superconductors, would be

$$\operatorname{curl} A \equiv B = 0, \tag{8a}$$

which is precisely the *Meissner effect*; furthermore, the appropriate counterpart of the Feynman theorem would be

$$\int_S B \cdot dS = n\frac{hc}{2e}, \tag{7a}$$

which leads to the "quantization of the magnetic flux", the quantum of flux being $hc/2e$.

[14] It is of interest to see that the angular momentum per particle in the fluid is given by

$$\frac{1}{\psi}\left(\frac{\hbar}{i}\frac{\partial}{\partial\phi}\psi\right) = s\hbar(= mv_\phi \rho);$$

this is again reminiscent of the quantum condition of Bohr.

[15] We have taken here: $l \simeq 1$ Å, so that $\ln(D/l) \simeq 7$.

[16] This result would be exact if we had a pair of oppositely directed *linear* vortices, with the same cross-section as shown in Fig. 10.4. In the case of a ring, the velocity would be somewhat larger.

[17] Physically means that, in the limiting case of *slow collisions*, only particles with opposite spins interact with one another.

[18] We have omitted here terms containing a “product of *four n*’ r the following reason: in view of the fact that the numerator of such a term would be symmetric the denominator antisymmetric with respect to the exchange operation $(p_1, p_2) \leftrightarrow (p_1', p_2')$, its sum over the variables p_1, p_2, p_1' (and p_2') would vanish identically.

[19] For a direct evaluation of the sum (22), in the limit $T \to 0$, see Abrikosov and Khalatnikov (1957). See also Problem 10.12.

[20] The third-order correction has also been discussed by Mohling (1961).

[21] For a *microscopic* theory of a Fermi liquid, see Nozières (1964); see also Tuttle and Mohling (1966).

[22] Of course, if the functions involved are spin-dependent, then the factor 2 in the element $d\tau$ (as well as in $d\tau'$) must be replaced by a summation over the spin variable(s).

[23] Since the total number of quasi-particles in the liquid is the same as the total number of real particles, to obtain the net transport of mass by the quasi-particles one has to multiply their number by the mass m of the *real* particle.

[24] At $T = 0$, $S = 0$; so there is no need to distinguish between the isothermal and adiabatic compressibilities of the liquid.

[25] Since μ_0 is an intensive quantity and, therefore, it depends upon N and V only through the ratio N/V, we can write: $\mu_0 = \mu_0(N/V)$. Consequently,

$$\left(\frac{\partial \mu_0}{\partial N}\right)_V = \mu_0' \left(\frac{\partial (N/V)}{\partial N}\right)_V = \mu_0' \frac{1}{V}$$

and

$$\left(\frac{\partial \mu_0}{\partial V}\right)_N = \mu_0' \left(\frac{\partial (N/V)}{\partial V}\right)_N = -\mu_0' \frac{N}{V^2}$$

Hence

$$\left(\frac{\partial \mu_0}{\partial N}\right)_V = -\frac{V}{N} \left(\frac{\partial \mu_0}{\partial V}\right)_N$$

[26] In a dense system, such as liquid He^3, the ratio m^*/m would be significantly larger than unity. The experimental work of Roberts and Sydoriak (1955), on the specific heat of liquid He^3, and the theoretical work of Brueckner and Gammel (1958), on the thermodynamics of a dense Fermi gas, suggest that the ratio $(m^*/m)_{\mathrm{He}^3} \simeq 1.85$.

[27] For the details of this calculation, see Lee and Yang (1959b) where the case of bosons, as well as of fermions, with spin J has been treated using the *binary collision* method. The second-order result for the case of *spinless* bosons was first obtained by Huang, Yang and Luttinger (1957) using the *method of pseudopotentials*.

CHAPTER 11

PHASE TRANSITIONS: CRITICALITY, UNIVERSALITY AND SCALING

VARIOUS physical phenomena to which the formalism of statistical mechanics has been applied may, in general, be divided into two categories. In the first category, the microscopic constituents of the given system are, or can be regarded as, practically non-interacting; as a result, the thermodynamic functions of the system follow straightforwardly from a knowledge of the energy levels of the individual constituents. Notable examples of phenomena belonging to this category are the specific heats of gases (Secs 1.4 and 6.5), the specific heats of solids (Sec. 7.3), chemical reactions and equilibrium constants (Problem 3.14), the condensation of an ideal Bose gas (Sec. 7.1), the spectral distribution of the black-body radiation (Sec. 7.2), the elementary electron theory of metals (Sec. 8.3), the phenomenon of paramagnetism (Secs 3.9 and 8.2), etc. In the case of solids, the interatomic interaction does, in fact, play an important physical role; however, since the actual positions of the atoms, over a substantial range of temperatures, do not depart significantly from their mean values, we can rewrite our problem in terms of the so-called *normal* coordinates and treat the given solid as an "assembly of practically non-interacting harmonic oscillators". We note that the most significant feature of the phenomena falling in the first category is that, with the *sole* exception of Bose–Einstein condensation, the thermodynamic functions of the systems involved are smooth and continuous!

Phenomena belonging to the second category, however, present a very different situation. In most cases, one encounters analytic discontinuities or singularities in the thermodynamic functions of the given system which, in turn, correspond to the occurrence of various kinds of *phase transitions*. Notable examples of phenomena belonging to this category are the condensation of gases, the melting of solids, phenomena associated with the co-existence of phases (especially in the neighborhood of a critical point), the behavior of mixtures and solutions (including the onset of phase separation), phenomena of ferromagnetism and antiferromagnetism, the order–disorder transitions in alloys, the superfluid transition from liquid He I to liquid He II, the transition from a normal to a superconducting material, etc. The characteristic feature of the interparticle interactions in these systems is that they *cannot* be "removed" by means of a transformation of the coordinates of the problem; accordingly, the energy levels of the total system cannot, in any simple manner, be related to the energy levels of the individual constituents.

One finds instead that, under favorable circumstances, a large number of microscopic constituents of the system may exhibit a tendency of interacting with one another in a rather *strong, cooperative* fashion. This cooperative behavior assumes macroscopic significance at a particular temperature T_c, known as the *critical temperature* of the system, and gives rise to the kind of phenomena listed above.

Mathematical problems associated with the study of cooperative phenomena are quite formidable.[1] To facilitate calculations, one is forced to introduce models in which the interparticle interactions are considerably simplified, yet retaining characteristics that are essential to the cooperative aspect of the problem. One then hopes that a theoretical study of these simplified models, which still involves serious difficulties of analysis, will reproduce the most basic features of the phenomena exhibited by actual physical systems. For instance, in the case of a magnetic transition, one may consider a lattice structure in which all interactions other than the ones among nearest-neighbor spins are neglected. It turns out that a model as simplified as that captures practically all the essential features of the phenomenon–especially in the close neighborhood of the critical point. The inclusion of interactions among spins farther out than the nearest neighbors does not change these features in any significant manner, nor are they affected by the replacement of one lattice structure by another so long as the dimensionality of the lattice is the same. Not only this, these features may also be shared, with little modification, by many other physical systems undergoing very different kinds of phase transitions, e.g. gas–liquid instead of paramagnetic–ferromagnetic. This "unity in diversity" turns out to be a hallmark of the phenomena associated with phase transitions–a subject we propose to explore in considerable detail in this and the following two chapters, but first a few preliminaries.

11.1. General remarks on the problem of condensation

We consider an N-particle system, obeying classical or quantum statistics, with the *proviso* that the total potential energy of the system can be written as a sum of two-particle terms $u(r_{ij})$, with $i < j$. The function $u(r)$ is supposed to satisfy the conditions

$$\left.\begin{array}{lll} u(r) = +\infty & \text{for} & r \le \sigma, \\ 0 > u(r) > -\varepsilon & \text{for} & \sigma < r < r^*, \\ u(r) = 0 & \text{for} & r \ge r^* \end{array}\right\}; \qquad (1)$$

see Fig. 11.1. Thus, each particle may be looked upon as a hard sphere of diameter σ, surrounded by an attractive potential of range r^* and of (maximum) depth ε. From a practical point of view, conditions (1) do not entail any "serious restriction" on the two-body potential, for the interparticle potentials ordinarily met with in nature are not materially different from the one satisfying these conditions. We therefore expect that the conclusions drawn from the use of this potential will not be very far from the realities of the actual physical phenomena.

Suppose that we are able to evaluate the *exact* partition function, $Q_N(V, T)$, of the given system. This function will possess certain properties which have been recognized and accepted for quite some time, though a rigorous proof of these was

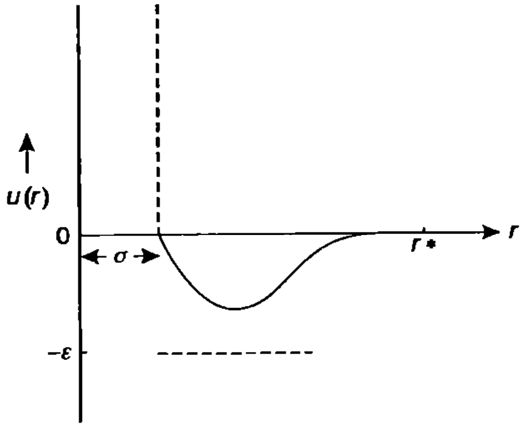


FIG. 11.1. A sketch of the interparticle potential $u(r)$, as given by eqn. (1).

first attempted by van Hove as late as in 1949.[2] These properties can be expressed as follows:

(i) In the thermodynamic limit, i.e. when N and $V \to \infty$ while the ratio N/V stays constant, the quantity $N^{-1} \ln Q$ tends to be a function only of the specific volume $v(= V/N)$ and the temperature T; this limiting form may be denoted by the symbol $f(v, T)$. It is natural to identify $f(v, T)$ with the *intensive* variable $-A/NkT$, where A is the Helmholtz free energy of the system. The thermodynamic pressure P is then given by

$$P(v, T) = -\left(\frac{\partial A}{\partial V}\right)_{N,T} = kT\left(\frac{\partial f}{\partial v}\right)_T, \tag{2}$$

which turns out to be a strictly *non-negative* quantity.

(ii) The function $f(v, T)$ is everywhere *concave*, so the slope $(\partial P \partial v)_T$ of the (P, v)-curve is *never* positive. While at high temperatures the slope is negative for all v, at lower temperatures there *can* exist a region (or regions) in which the slope is zero and, consequently, the system is infinitely compressible! The existence of such regions, in the (P, v)-diagram, corresponds to the co-existence of two or more phases of different density in the given system; in other words, it constitutes direct evidence of the onset of a phase transition in the system. In this connection it is important to note that, so long as one uses the *exact* partition function of the system, isotherms of the van der Waals type, which possess unphysical regions of positive slope, *never* appear. On the other hand, if the partition function is evaluated under approximations, as we did in the derivation of the van der Waals equation of state in Sec. 9.3, isotherms with unphysical regions may indeed appear. In that case the isotherms in question have got to be "corrected", by introducing a region of "flatness" $(\partial P/\partial v = 0)$, with the help of the *Maxwell construction of equal areas*; see Fig. 11.2.[3] The real reason for the appearance of unphysical regions in the isotherms is that the approximate evaluations of the partition function introduce, almost invariably (though implicitly), the restraint of a *uniform* density throughout the system. This restraint eliminates the very possibility of the system passing through states in which there exist, side by side, two phases of different densities; in other

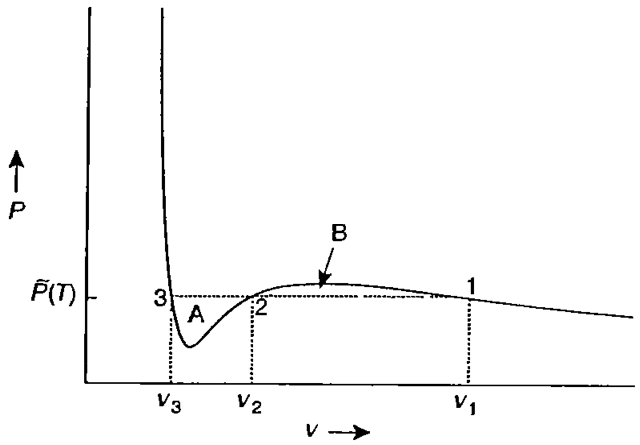


FIG. 11.2. An unphysical isotherm corrected with the help of the Maxwell construction; the horizontal line is such that the areas A and B are equal. The "corrected" isotherm corresponds to a phase transition, taking place at pressure $\tilde{P}(T)$, with densities v_1^{-1} and v_3^{-1} of the respective phases.

words, the existence of a region of "flatness" in the (P, v)-diagram is automatically ruled out. On the other hand, an exact evaluation of the partition function must allow for *all* possible configurations of the system, including the ones characterized by a simultaneous existence of two or more phases of different densities. Under suitable conditions (for instance, when the temperature is sufficiently low), such a configuration might turn out to be the *equilibrium configuration* of the system, with the result that the system shows up in a multiphase, rather than a single-phase, state. We should, in this context, mention that if in the evaluation of the partition function one introduces no other approximation except the assumption of a uniform density in the system, then the resulting isotherms, corrected with the help of the Maxwell construction, would be the *exact* isotherms of the problem.

(iii) The presence of an absolutely flat portion in an isotherm, with mathematical singularities at its ends, is, strictly speaking, a consequence of the limiting process $N \to \infty$. If N were *finite*, and if the exact partition function were used, then the quantity P', defined by the relation

$$P' = kT \left(\frac{\partial \ln Q}{\partial V} \right)_{N,T}, \tag{3}$$

would be free from mathematical singularities. The ordinarily sharp corners in an isotherm would be rounded off; at the same time, the ordinarily flat portion of the isotherm would not be strictly flat–it would have for large N a small, negative slope. In fact, the quantity P' in this case would not be a function of v and T alone; it would depend on the number N as well, though in a thermodynamically negligible manner.

If we employ the grand partition function $\mathcal{Q}$, as obtained from the exact partition functions Q_N, viz.

$$\mathcal{Q}(z, V, T) = \sum_{N \geq 0} Q_N(V, T) z^N, \tag{4}$$

a similar picture results. To see this, we note that for real molecules, with a given V, the variable N will be bounded by an upper limit, say N_m, which is the number of molecules that fill the volume V "tight-packed"; obviously, $N_m \sim V/\sigma^3$. For $N > N_m$, the potential energy of the system will be infinite; accordingly,

$$Q_N(N > N_m) \equiv 0. \tag{5}$$

Hence, for all practical purposes, our power series in (4) is a polynomial in z (which is ≥ 0) and is of degree N_m. Since the coefficients Q_N are all positive and $Q_0 \equiv 1$, the sum $\mathcal{Q} \geq 1$. The thermodynamic potential $\ln \mathcal{Q}$ is, therefore, a *well-behaved* function of the parameters z, V and T. Consequently, so long as V (and hence N_m) remains finite, we do not expect any singularities or discontinuities in any of the functions derived from this potential. A non-analytic behavior could appear only in the limit $(V, N_m) \to \infty$.

We now define P' by the relation

$$P' = \frac{kT}{V} \ln \mathcal{Q} \quad (V \text{ finite}); \tag{6}$$

since $\mathcal{Q} \geq 1$, $P' \geq 0$. The mean number of particles and the mean square deviation in this number are given by the formulae

$$\overline{N} = \left(\frac{\partial \ln \mathcal{Q}}{\partial \ln z}\right)_{V,T} \tag{7}$$

and

$$\overline{N^2} - \overline{N}^2 \equiv \overline{(N - \overline{N})^2} = \left(\frac{\partial \overline{N}}{\partial \ln z}\right)_{V,T}, \tag{8}$$

respectively; see Secs 4.2 and 4.5. Accordingly,

$$\left(\frac{\partial \ln \mathcal{Q}}{\partial \overline{N}}\right)_{V,T} = \left(\frac{\partial \ln \mathcal{Q}}{\partial \ln z}\right)_{V,T} \Big/ \left(\frac{\partial \overline{N}}{\partial \ln z}\right)_{V,T} = \frac{\overline{N}}{\overline{N^2} - \overline{N}^2}. \tag{9}$$

On the other hand, writing $\overline{v}$ for $V/\overline{N}$ and using (6), we have

$$\left(\frac{\partial \ln \mathcal{Q}}{\partial \overline{N}}\right)_{V,T} = \frac{V}{kT}\left(\frac{\partial P'}{\partial \overline{N}}\right)_{V,T} = -\frac{\overline{v}^2}{kT}\left(\frac{\partial P'}{\partial \overline{v}}\right)_{V,T}. \tag{10}$$

Comparing (9) and (10), we obtain

$$\left(\frac{\partial P'}{\partial \overline{v}}\right)_{V,T} = -\frac{kT}{V^2}\frac{\overline{N}^3}{\overline{N^2} - \overline{N}^2}, \tag{11}$$

which is clearly non-positive.[4] For finite V, expression (11) will never vanish; accordingly, P' will never be strictly constant. Nevertheless, the slope $(\partial P'/\partial \overline{v})$ can, in a certain region, be extremely small–in fact, as small as $O(1/\overline{N})$; such a region would hardly be distinguishable from a phase transition because, on a macroscopic scale, the value of P' in such a region would be as good as a constant.[5]

If we now define the pressure of the system by the limiting relationship

$$P(\overline{v}, T) = \lim_{V\to\infty} P'(\overline{v}, T; V) = kT \lim_{V\to\infty}\left(\frac{1}{V}\ln \mathcal{Q}(z, V, T)\right), \tag{12}$$

then we *can* expect, in a set of isotherms, an absolutely flat portion ($\partial P/\partial \bar{v} \equiv 0$), with sharp corners implying mathematical singularities. The mean particle density $\bar{n}$ would now be given by

$$\bar{n} = \lim_{V \to \infty} \left[\frac{1}{V} \frac{\partial \ln \mathscr{Q}(z, V, T)}{\partial \ln z} \right]; \tag{13}$$

it seems important to mention here that the operation $V \to \infty$ and the operation $\partial/\partial \ln z$ cannot be interchanged freely.

In passing, we note that the picture emerging from the grand partition function $\mathscr{Q}$, which is supposed to have been obtained from the exact partition functions Q_N, remains unchanged even if one had employed a set of approximate Q_N. This is so because the argument developed in the preceding paragraphs makes no use whatsoever of the actual form of the functions Q_N. Thus, if an approximate Q_N leads to the van der Waals type of loop in the canonical ensemble, as in Fig. 11.2, the corresponding set of Q_N, when employed in a grand canonical ensemble, would lead to isotherms free from such loops (Hill, 1953).

Subsequent to van Hove, Yang and Lee (1952) suggested an alternative approach that enables one to carry out a rigorous mathematical discussion of the phenomenon of condensation and of other similar transitions. In their approach one is primarily concerned with the analytic behavior of the quantities P and $\bar{n}$, of eqns (12) and (13), as functions of z at different values of T. The problem is examined in terms of the "zeros of the grand partition function $\mathscr{Q}$ in the *complex* z-plane", with attention focused on the way these zeros are distributed in the plane and the manner in which they evolve as the volume of the system is increased. Since for real, positive z, $\mathscr{Q} \geq 1$, none of the zeros will lie on the real, positive axis in the z-plane. However, as $V \to \infty$ (and hence the degree of the polynomial (4) and, with it, the number of zeros itself grows to infinity), the distribution of zeros is expected to become *continuous* and, depending on T, may in fact *converge* upon the real, positive axis at one or more points z_c. If so, our functions $P(z)$ and $\bar{n}(z)$, even with z varied along the real axis only, may, by virtue of their relationship to the function $\ln \mathscr{Q}$, turn out to be singular at the points $z = z_c$. The presence of such a singularity would imply the onset of a phase transition in the system. For further details of this approach, see Secs 12.3 and 12.4 of the first edition of this book. See also Griffiths (1972), pp. 50–58.

11.2. Condensation of a van der Waals gas

We start with the simplest, and historically the first, theoretical model that undergoes a gas–liquid phase transition. This model is generally referred to as the van der Waals gas and obeys the equation of state, see eqn. (9.3.9),

$$P = \frac{RT}{v - b} - \frac{a}{v^2}, \tag{1}$$

v being the *molar* volume of the gas; the parameters a and b then also pertain to one mole of the gas. We recall that, while a is a measure of the attractive forces among the molecules of the system, b is a measure of the repulsive forces that come into play when two molecules come too close to one another; accordingly,

b is also a measure of the "effective space" occupied by the molecules (by virtue of a finite volume that may be associated with each one of them). In Sec. 9.3, the equation of state (1) was derived under the express assumption that $v \gg b$; here, we shall pretend, with van der Waals, that this equation holds even when v is comparable to b.

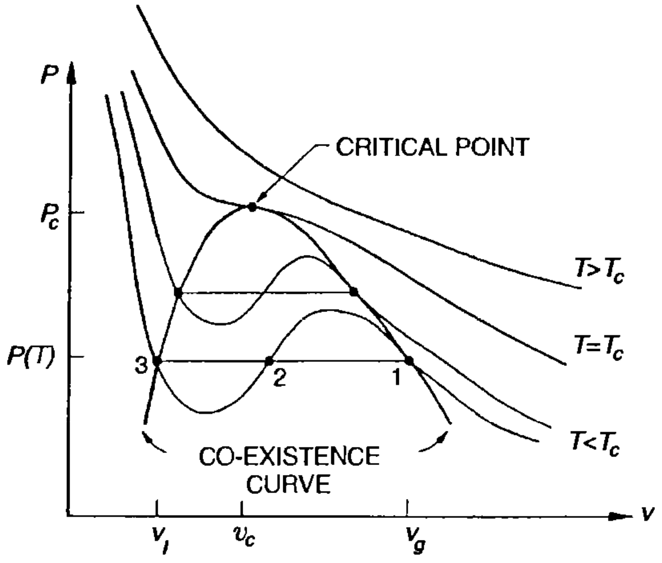


FIG. 11.3. The isotherms of a van der Waals system; those for $T < T_c$ are "corrected" with the help of the Maxwell construction, thus leading to the co-existence curve at the top of which sits the critical point.

The isotherms following from eqn. (1) are shown in Fig. 11.3. We note that, for temperatures above a *critical temperature* T_c, P decreases monotonically with v. For $T < T_c$, however, the relationship is non-monotonic, so that over a certain range of v we encounter a region where $(\partial P/\partial v) > 0$; such a region is unphysical and must be "corrected" with the help of the Maxwell construction,[6] leading to an isotherm with a flat portion signalling transition from the gaseous state with molar volume $v_g(T)$ to the liquid state with molar volume $v_l(T)$ at a constant pressure $P(T)$. For $v_l < v < v_g$, the system resides in a state of mixed phases–partly liquid, partly gaseous–and, since the passage from one end of the flat portion to the other takes place with $\Delta v \neq 0$ but $\Delta P = 0$, the system is all along in a state of *infinite* compressibility; clearly, we are encountering here a brand of behavior that is patently *singular*. As T increases towards T_c, the transition takes place at a comparatively higher value of P, with v_g less than and v_l more than before–so that, as $T \to T_c$, both v_g and v_l approach a common value v_c which may be referred to as the *critical volume*; the corresponding value of P may then be designated by P_c, and we find ourselves located at the *critical point* of the system. The locus of all such points as $v_l(T)$ and $v_g(T)$ is generally referred to as the *co-existence curve*, for the simple reason that in the region enclosed by this curve the gaseous and the liquid phases mutually co-exist; the top of this curve, where $v_l = v_g$, coincides with the critical point itself. Finally, the isotherm pertaining to $T = T_c$ which, of course, passes through the critical point is referred to as the *critical isotherm* of the system; it is straightforward to see that the critical point is a point of inflection of this isotherm, so that both $(\partial P/\partial v)_T$ and $(\partial^2 P/\partial^2 v)_T$ vanish at this point. Using (1), we obtain for the coordinates of the critical point

$$P_c = \frac{a}{27b^2}, \quad v_c = 3b, \quad T_c = \frac{8a}{27bR}, \tag{2}$$

so that the number

$$\mathcal{K} \equiv RT_c/P_c v_c = 8/3 = 2.666\ldots \tag{3}$$

We thus find that, while P_c, v_c and T_c vary from system to system (through the interaction parameters a and b), the quantity $\mathcal{K}$ has a common, *universal* value for all of them–so long as they all obey the same, i.e. van der Waals, equation of state. The experimental results for $\mathcal{K}$ indeed show that it is nearly the same over a large group of substances; for instance, its value for carbon tetrachloride, ethyl ether and ethyl formate is 3.677, 3.814 and 3.895, respectively–close, though not exactly the same, and also a long way from the van der Waals value. The concept of universality is, nonetheless, there (even though van der Waals equation of state may not truly apply).

It is now tempting to see if the equation of state itself can be written in a universal form. We find that this indeed can be done by introducing *reduced variables*

$$P_r = \frac{P}{P_c}, \quad v_r = \frac{v}{v_c}, \quad T_r = \frac{T}{T_c}. \tag{4}$$

Using (1) and (2), we readily obtain the reduced equation of state

$$\left(P_r + \frac{3}{v_r^2}\right)(3v_r - 1) = 8T_r, \tag{5}$$

which is clearly universal for all systems obeying van der Waals' original equation of state (1); all we have done here is to rescale the observable quantities P, v and T in terms of their critical values and thereby "push the interaction parameters a and b into the background". Now, if two different systems happen to be in states characterized by the same values of v_r and T_r, then their P_r would also be the same; the systems are then said to be in "corresponding states" and, for that reason, the statement just made is referred to as the "law of corresponding states". Clearly, the passage from eqn. (1) to eqn. (5) takes us from an expression of diversity to a statement of unity!

We shall now examine the behavior of the van der Waals system in the *close* neighborhood of the critical point. For this, we write

$$P_r = 1 + \pi, \quad v_r = 1 + \psi, \quad T_r = 1 + t. \tag{6}$$

Equation (5) then takes the form

$$\pi(2 + 7\psi + 8\psi^2 + 3\psi^3) + 3\psi^3 = 8t(1 + 2\psi + \psi^2). \tag{7}$$

First of all, along the critical isotherm ($t = 0$) and in the close vicinity of the critical point ($|\pi|, |\psi| \ll 1$), we obtain the simple, asymptotic result

$$\pi \approx -\tfrac{3}{2}\psi^3, \tag{8}$$

which is indic[illegible] of the "degree of flatness" of the critical isotherm at the critical point. Next, we examine the dependence of ψ on t as we approach the critical point from below. For this, we write (7) in the form

$$3\psi^3 + 8(\pi - t)\psi^2 + (7\pi - 16t)\psi + 2(\pi - 4t) \simeq 0. \tag{9}$$

Now, a close look at the (symmetric) shape of the co-existence curve near its top (where $|t| \ll 1$) shows that the three roots ψ_1, ψ_2 and ψ_3 of eqn. (9), which arise from the limiting behavior of the roots v_1, v_2 and v_3 of the original equation of state (1) as $T \to T_c-$, are such that $|\psi_2| \ll |\psi_{1,3}|$ and $|\psi_1| \simeq |\psi_3|$. This means that, in the region of interest,

$$\pi \approx 4t, \tag{10}$$

so that one of the roots, say ψ_2, of eqn. (9) essentially vanishes while the other two are given by

$$\psi^2 + 8t\psi + 4t \simeq 0. \tag{9a}$$

We expect the middle term here to be negligible (as will be confirmed by the end result), yielding

$$\psi_{1,3} \approx \pm 2|t|^{1/2}; \tag{11}$$

note that the upper sign here pertains to the gaseous phase and the lower sign to the liquid phase.

Finally, we consider the *isothermal compressibility* of the system which, in terms of reduced variables, is determined essentially by the quantity $-(\partial\psi/\partial\pi)_t$. Retaining only the dominant terms, we obtain from (7)

$$-\left(\frac{\partial\psi}{\partial\pi}\right)_t \approx \frac{2}{7\pi + 9\psi^2 - 16t}. \tag{12}$$

For $t > 0$, we approach the critical point along the critical isochore ($\psi = 0$); eqn. (12), with the help of eqn. (10), then gives

$$-\left(\frac{\partial\psi}{\partial\pi}\right)_{t\to 0+} \approx \frac{1}{6t}. \tag{13}$$

For $t < 0$, we approach the critical point along the co-existence curve (on which $\psi^2 \approx -4t$); we now obtain

$$-\left(\frac{\partial\psi}{\partial\pi}\right)_{t\to 0-} \approx \frac{1}{12|t|}. \tag{14}$$

For the record, we quote here results for the specific heat, C_v, of the van der Waals gas (Uhlenbeck, 1966; Thompson, 1988)

$$C_V \approx \begin{cases} (C_V)_{\text{ideal}} + \frac{9}{2}Nk\left(1 + \frac{28}{25}t\right) & (t \leq 0) \qquad \text{(15a)} \\ (C_V)_{\text{ideal}} & (t > 0), \qquad \text{(15b)} \end{cases}$$

which imply a finite jump at the critical point.

Equations (8), (11), (13), (14) and (15) illustrate the nature of the *critical behavior* displayed by a van der Waals system undergoing the gas–liquid transition. While it differs in several important respects from the critical behavior of real physical systems, it shows up again and again in studies pertaining to other critical phenomena that have apparently nothing to do with the gas–liquid phase transition. In fact, this particular brand of behavior turns out to be a bench mark against which the results of more sophisticated theories are automatically compared.

11.3. A dynamical model of phase transitions

A number of physico-chemical systems which undergo phase transitions can be represented, to varying degrees of accuracy, by an "array of lattice sites, with only nearest-neighbor interaction that depends upon the manner of occupation of the neighboring sites". This simple-minded model turns out to be good enough to provide a unified, theoretical basis for understanding a variety of phenomena such as ferromagnetism and antiferromagnetism, gas–liquid and liquid–solid transitions, order–disorder transitions in alloys, phase separation in binary solutions, etc. There is no doubt that this model considerably oversimplifies the actual physical systems it is supposed to represent; nevertheless, it does retain the essential physical features of the problem–features that account for the propagation of *long-range order* in the system. Accordingly, it does lead to the onset of a phase transition in the given system, which arises in the nature of a *cooperative* phenomenon.

We find it convenient to formulate our problem in the language of ferromagnetism; later on, we shall establish correspondence between this language and the languages appropriate to other physical phenomena. We thus regard each of the N lattice sites to be occupied by an atom possessing a magnetic moment μ, of magnitude $g\mu_B\sqrt{[J(J+1)]}$, which is capable of $(2J+1)$ discrete orientations in space. These orientations define "different possible manners of occupation" of a given lattice site; accordingly, the whole lattice is capable of $(2J+1)^N$ different configurations. Associated with each configuration is an energy E that arises from mutual interactions among the neighboring atoms of the lattice and from the interaction of the whole lattice with an external field $\boldsymbol{B}$. A statistical analysis in the canonical ensemble should then enable us to determine the expectation value, $\overline{M}(B, T)$, of the net magnetization M. The presence of a *spontaneous magnetization* $\overline{M}(0, T)$ at temperatures below a certain (critical) temperature T_c and its absence above that temperature will then be interpreted as a ferromagnetic phase transition in the system at $T = T_c$.

Detailed studies, both theoretical and experimental, have shown that, for all ferromagnetic materials, data on the temperature dependence of the spontaneous magnetization, $\overline{M}(0, T)$, fit best with the value $J = \frac{1}{2}$; see Fig. 11.4. One is, therefore, tempted to infer that the phenomenon of ferromagnetism is associated only with the spins of the electrons and not with their orbital motions. This is further confirmed by the gyromagnetic experiments (Barnett, 1944; Scott, 1951, 1952), in which one either reverses the magnetization of a freely suspended specimen and observes the resulting rotation or imparts a rotation to the specimen and observes the resulting magnetization; the former is known as the *Einstein–de Haas method*, the latter the *Barnett method*. From these experiments one can derive the relevant g-value of the specimen which, in each case, turns out to be very close to 2; this, as we know, pertains to the electron spin. Therefore, in discussing the problem

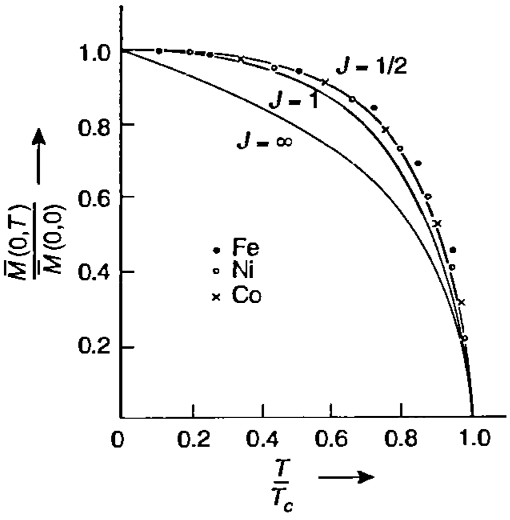


Fig. 11.4. Spontaneous magnetization of iron, nickel and cobalt as a function of temperature. Theoretical curves are based on the Weiss theory of ferromagnetism.

of ferromagnetism, we may specifically take: $\mu = 2\mu_B\sqrt{[s(s+1)]}$, where s is the quantum number associated with the electron spin. With $s = \frac{1}{2}$, only two orientations are possible for each lattice site, namely $s_z = +\frac{1}{2}$ (with $\mu_z = +\mu_B$) and $s_z = -\frac{1}{2}$ (with $\mu_z = -\mu_B$). The whole lattice is then capable of 2^N configurations; one such configuration is shown in Fig. 11.5.

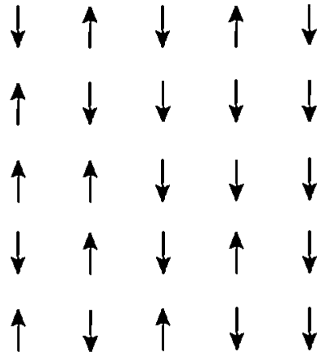


Fig. 11.5. One of the 2^N possible configurations of a system composed of N spins; here, $N = 25$.

We now consider the nature of the interaction energy between two neighboring spins s_i and s_j. According to quantum mechanics, this energy is of the form $K_{ij} \pm J_{ij}$, where the upper sign applies to "antiparallel" spins ($S = 0$) and the lower sign to "parallel" spins ($S = 1$). Here, K_{ij} is the direct or Coulomb energy between the two spins, while J_{ij} is the exchange energy between them:

$$K_{ij} = \int \psi_i^*(1)\psi_j^*(2)u_{ij}\psi_j(2)\psi_i(1)\,d\tau_1\,d\tau_2, \tag{1}$$

while

$$J_{ij} = \int \psi_j^*(1)\psi_i^*(2)u_{ij}\psi_j(2)\psi_i(1)\,d\tau_1\,d\tau_2, \tag{2}$$

u_{ij} being the relevant interaction potential. The energy difference between a state of "parallel" spins and one of "antiparallel" spins is given by

$$\varepsilon_{\uparrow\uparrow} - \varepsilon_{\uparrow\downarrow} = -2J_{ij}. \tag{3}$$

If $J_{ij} > 0$, the state $\uparrow\uparrow$ is energetically favored against the state $\uparrow\downarrow$; we then look for the possibility of *ferromagnetism*. If, on the other hand, $J_{ij} < 0$, the situation is reversed and we see the possibility of *antiferromagnetism.*

It seems useful to express the interaction energy of the two states, $\uparrow\uparrow$ and $\downarrow\downarrow$, by a single expression; to do this, we consider the eigenvalues of the scalar product

$$\begin{aligned} s_i \cdot s_j &= \tfrac{1}{2}\{(s_i + s_j)^2 - s_i^2 - s_j^2\} \\ &= \tfrac{1}{2}S(S+1) - s(s+1), \end{aligned} \tag{4}$$

which equals $+\frac{1}{4}$ if $S = 1$ and $-\frac{3}{4}$ if $S = 0$. We may, therefore, write for the interaction energy of the spins i and j

$$\varepsilon_{ij} = \text{const.} - 2J_{ij}(s_i \cdot s_j). \tag{5}$$

which is consistent with the energy difference (3). The precise value of the constant here is immaterial because the potential energy is arbitrary to the extent of an additive constant anyway. Typically, the exchange interaction J_{ij} falls off rapidly as the separation of the two spins is increased. To a first approximation, therefore, we may regard J_{ij} as negligible for all but nearest-neighbor pairs (for which its value may be denoted by the simpler symbol J). The interaction energy of the whole lattice is then given by

$$E = \text{const.} - 2J \sum_{\text{n.n.}} (s_i \cdot s_j), \tag{6}$$

where the summation goes over all nearest-neighbor pairs in the lattice. The model based on expression (6) for the interaction energy of the lattice is known as the *Heisenberg model* (1928).

A simpler model results if we use, instead of (6), a *truncated* expression in which the product $(s_i \cdot s_j)$, which is equal to the sum $(s_{ix}s_{jx} + s_{iy}s_{jy} + s_{iz}s_{jz})$, is replaced by a single term $s_{iz}s_{jz}$; one reason for adopting this simpler model is that it does not necessarily require a quantum-mechanical treatment (because all the variables in the truncated expression for E commute). Expression (6) may now be written as

$$E = \text{const.} - J \sum_{\text{n.n.}} \sigma_i \sigma_j, \tag{7}$$

where the new symbol σ_i (or σ_j) $= +1$ for an "up" spin and -1 for a "down" spin; note that, with the introduction of the new symbol, we still have: $\varepsilon_{\uparrow\uparrow} - \varepsilon_{\uparrow\downarrow} = -2J$. The model based on expression (7) is known as the *Ising model*; it originated with Lenz (1920) and was subsequently investigated by his student Ising (1925).[7]

A different model results if we suppress the z-components of the spins and retain the x- and y-components instead. This model was originally introduced by Matsubara and Matsuda (1956) as a model of a quantum lattice gas, with possible relevance to the superfluid transition in liquid He^4. The critical behavior of this so-called X–Y model has been investigated in detail by Betts and co-workers, who have also emphasized the relevance of this model to the study of insulating ferromagnets (see Betts *et al.*, 1968–74).

It seems appropriate to regard the Ising and the $X-Y$ models as special cases of an *anisotropic* Heisenberg model with interaction parameters J_x, J_y and J_z; while the Ising model represents the situation $J_x, J_y \ll J_z$, the $X-Y$ model represents just the opposite. Introducing a parameter n, which denotes the number of spin components entering into the Hamiltonian of the system, we may regard the Ising, the $X-Y$ and the Heisenberg models as pertaining to the n-values 1, 2 and 3, respectively. As will be seen later, the parameter n, along with the dimensionality d of the lattice, constitutes the basic set of elements that determine the qualitative nature of the critical behavior of a given system. For the time being, though, we confine our attention to the Ising model which is not only the simplest one to analyze but also unifies the study of phase transitions in systems as diverse as ferromagnets, gas–liquids, liquid mixtures, binary alloys, etc.

To study the statistical mechanics of the Ising model, we disregard the kinetic energy of the atoms occupying the various lattice sites, for the phenomenon of phase transitions is essentially a consequence of the interaction energy among the atoms; in the interaction energy again, we include only the nearest-neighbor contributions, in the hope that the farther-neighbor contributions would not affect the results qualitatively. To fix the z-direction, and to be able to study properties such as magnetic susceptibility, we subject the lattice to an external magnetic field $\boldsymbol{B}$, directed "upward"; the spin σ_i then possesses an additional potential energy $-\mu B\sigma_i$.[8] The Hamiltonian of the system in configuration $\{\sigma_1, \sigma_2, \ldots, \sigma_N\}$ is then given by

$$H\{\sigma_i\} = -J\sum_{\text{n.n.}} \sigma_i\sigma_j - \mu B\sum_i \sigma_i, \tag{8}$$

and the partition function by

$$\begin{aligned} Q_N(B,T) &= \sum_{\sigma_1}\sum_{\sigma_2}\cdots\sum_{\sigma_N} \exp[-\beta H\{\sigma_i\}] \\ &= \sum_{\sigma_1}\sum_{\sigma_2}\cdots\sum_{\sigma_N} \exp\left[\beta J\sum_{n.n.}\sigma_i\sigma_j + \beta\mu B\sum_i \sigma_i\right] \end{aligned} \tag{9}$$

The Helmholtz free energy, the internal energy, the specific heat and the net magnetization of the system then follow from the formulae

$$A(B,T) = -kT\ln Q_N(B,T), \tag{10}$$

$$U(B,T) = -T^2\frac{\partial}{\partial T}\left(\frac{A}{T}\right) = kT^2\frac{\partial}{\partial T}\ln Q_N, \tag{11}$$

$$C(B,T) = \frac{\partial U}{\partial T} = -T\frac{\partial^2 A}{\partial T^2}, \tag{12}$$

and

$$\overline{M}(B,T) = \mu\overline{\left(\sum_i \sigma_i\right)} = \overline{\left(-\frac{\partial H}{\partial B}\right)} = \frac{1}{\beta}\left(\frac{\partial \ln Q_N}{\partial B}\right)_T = -\left(\frac{\partial A}{\partial B}\right)_T. \tag{13}$$

Obviously, the quantity $\overline{M}(0, T)$ gives the *spontaneous magnetization* of the system; if it is nonzero at temperatures below a certain critical temperature T_c, the system would be ferromagnetic for $T < T_c$ and paramagnetic for $T > T_c$. At the transition temperature itself, the system is expected to show some sort of a singular behavior.

It is obvious that the energy levels of the system as a whole will be *degenerate*, in the sense that the various configurations $\{\sigma_i\}$ will not all possess *distinct* energy values. In fact, the energy of a given configuration does not depend upon the *detailed* values of all the variables σ_i; it depends only upon a few numbers such as the total number N_+ of "up" spins, the total number N_{++} of "up–up" nearest-neighbor pairs, and so on. To see this, we define certain other numbers as well, e.g. N_- as the total number of "down" spins, N_{--} as the total number of "down–down" nearest-neighbor pairs and N_{+-} as the total number of nearest-neighbor pairs with opposite spins. The numbers N_+ and N_- must satisfy the relation

$$N_+ + N_- = N. \tag{14}$$

And if q denotes the *coordination number* of the lattice, i.e. the number of nearest neighbors for each lattice site,[9] then we also have the relations

$$qN_+ = 2N_{++} + N_{+-}\,, \tag{15}$$

$$qN_- = 2N_{--} + N_{+-}. \tag{16}$$

With the help of these relations, we can express all our numbers in terms of any two of them, say N_+ and N_{++}. Thus

$$N_- = N - N_+, \quad N_{+-} = qN_+ - 2N_{++}, \quad N_{--} = \tfrac{1}{2}qN - qN_+ + N_{++}; \tag{17}$$

it will be noted that the total number of nearest-neighbor pairs of all types is given, quite expectedly, by the expression

$$N_{++} + N_{--} + N_{+-} = \tfrac{1}{2}qN. \tag{18}$$

Naturally, the Hamiltonian of the system can also be expressed in terms of N_+ and N_{++}; we have from (8), with the help of the relations established above,

$$\begin{aligned} H_N(N_+, N_{++}) &= -J(N_{++} + N_{--} - N_{+-}) - \mu B(N_+ - N_-) \\ &= -J\left(\tfrac{1}{2}qN - 2qN_+ + 4N_{++}\right) - \mu B(2N_+ - N). \end{aligned} \tag{19}$$

Now, let $g_N(N_+, N_{++})$ be "the number of distinct ways in which the N spins of the lattice can be so arranged as to yield certain preassigned values of the numbers N_+ and N_{-+}". The partition function of the system can then be written as

$$Q_N(B, T) = \sum_{N_+, N_{++}}{}' g_N(N_+, N_{++}) \exp\{-\beta H_N(N_+, N_{++})\}, \tag{20}$$

that is,

$$e^{-\beta A} = e^{\beta N\left(\frac{1}{2}qJ - \mu B\right)} \sum_{N_+=0}^{N} e^{-2\beta(qJ-\mu B)N_+} \sum_{N_{++}}{}' g_N(N_+, N_{++}) e^{4\beta J N_{++}}, \tag{21}$$

where the primed summation in (21) goes over all values of N_{++} that are consistent with a *fixed* value of N_+ and is followed by a summation over all possible values of N_+, i.e. from $N_+ = 0$ to $N_+ = N$. The central problem thus consists in determining the *combinatorial function* $g_N(N_+, N_{++})$ for the various lattices of interest.

11.4. The lattice gas and the binary alloy

Apart from ferromagnets, the Ising model can be readily adapted to simulate the behavior of certain other systems as well. More common among these are (i) the lattice gas and (ii) the binary alloy.

(i) *The lattice gas.* Though it had already been recognized that the results derived for the Ising model would apply equally well to a system of "occupied" and "unoccupied" lattice sites, i.e. to a system of "atoms" and "holes" in a lattice, it was Yang and Lee (1952) who first used the term "lattice gas" to describe such a system. By definition, a lattice gas is a collection of atoms, N_a in number, which can occupy only discrete positions in space–positions that constitute a lattice structure with coordination number q. Each lattice site can be occupied by *at most* one atom, and the interaction energy between two occupied sites is nonzero, say $-\varepsilon_0$, only if the sites involved constitute a *nearest-neighbor* pair. The configurational energy of the gas is then given by

$$E = -\varepsilon_0 N_{aa}, \tag{1}$$

where N_{aa} is the total number of nearest-neighbor pairs (of occupied sites) in a given configuration of the system. Let $g_N(N_a, N_{aa})$ denote "the number of distinct ways in which the N_a atoms of the gas, assumed indistinguishable, can be distributed among the N sites of the lattice so as to yield a certain preassigned value of the number N_{aa}". The partition function of the system, neglecting the kinetic energy of the atoms, is then given by

$$Q_{N_a}(N, T) = {\sum_{N_{aa}}}' g_N(N_a, N_{aa}) e^{\beta\varepsilon_0 N_{aa}}, \tag{2}$$

where the primed summation goes over all values of N_{aa} that are consistent with the given values of N_a and N; clearly, the number N here plays the role of the "total volume" available to the gas.

Going over to the grand canonical ensemble, we write for the grand partition function of the system

$$\mathcal{Q}(z, N, T) = \sum_{N_a=0}^{N} z^{N_a} Q_{N_a}(N, T). \tag{3}$$

The pressure P and the mean number $\overline{N}_a$ of the atoms in the gas are then given by

$$e^{\beta PN} = \sum_{N_a=0}^{N} z^{N_a} {\sum_{N_{aa}}}' g_N(N_a, N_{aa}) e^{\beta\varepsilon_0 N_{aa}} \tag{4}$$

and

$$\frac{\overline{N}_a}{N} = \frac{1}{v} = \frac{z}{kT}\left(\frac{\partial P}{\partial z}\right)_T; \tag{5}$$

here, v denotes the average volume per particle of th [illegible] s (measured in terms of the "volume of a primitive cell of the lattice").

To establish a formal correspondence between the lattice gas and the ferromagnet, we compare the present formulae with the ones established in the preceding section–in particular, formula (4) with formula (11.3.21). The first thing we note here is that the canonical ensemble of the ferromagnet corresponds to the grand canonical ensemble of the lattice gas. The rest of the correspondence is summarized in the following chart:

The lattice gas		*The ferromagnet*
$N_a, N - N_a$	$\leftrightarrow$	$N_+, N - N_+ (= N_-)$
ε_0	$\leftrightarrow$	$4J$
z	$\leftrightarrow$	$\exp\{-2\beta(qJ - \mu B)\}$
P	$\leftrightarrow$	$-\left(\frac{A}{N} + \frac{1}{2}qJ - \mu B\right)$
$\frac{\overline{N}_a}{N}\left(= \frac{1}{v}\right)$	$\leftrightarrow$	$\frac{\overline{N}_+}{N}\left(= \frac{1}{2}\left\{\frac{\overline{M}}{N\mu} + 1\right\}\right)$,

where

$$\overline{M} = \mu(\overline{N}_+ - \overline{N}_-) = \mu(2\overline{N}_+ - N). \tag{6}$$

We also note that the ferromagnetic analogue of formula (5) would be

$$\frac{\overline{N}_+}{N} = \frac{1}{kT}\left\{\frac{\partial\left(A/N + \frac{1}{2}qJ - \mu B\right)}{2\beta\partial(qJ - \mu B)}\right\}_T = \frac{1}{2}\left[-\frac{1}{N\mu}\left(\frac{\partial A}{\partial B}\right)_T + 1\right] \tag{7}$$

which, by eqn. (11.3.13), assumes the form

$$\frac{\overline{N}_+}{N} = \frac{1}{2}\left(\frac{\overline{M}}{N\mu} + 1\right). \tag{8}$$

It is quite natural to ask: does lattice gas correspond to any real physical system in nature? The immediate answer is that if we let the lattice constant tend to zero (thus going from a discrete structure to a continuous one) and also add, to the lattice-gas formulae, terms corresponding to an ideal gas (namely, the kinetic energy terms), then the model might simulate the behavior of a gas of real atoms interacting through a delta function potential. A study of the possibility of a phase transition in such a system may, therefore, be of some value in understanding phase transitions in real gases. The case $\varepsilon_0 > 0$, which implies an *attractive* interaction among the nearest neighbors, has been frequently cited as a possible model for a gas–liquid transition. On the other hand, if the interaction is *repulsive* ($\varepsilon_0 < 0$), so that configurations with alternating sites being "occupied" and "unoccupied" are the more favored ones, then we obtain a model which arouses interest in connection with the theory of solidification; in such a study, however, the lattice constant has to stay finite. Thus, several authors have pursued the study of the antiferromagnetic version of this model, hoping that this might throw some light on the liquid–solid transition. For a bibliography of these pursuits, see the review article by Brush (1967).

(ii) *The binary alloy*. Much of the early activity in the theoretical analysis of the Ising model was related to the study of order–disorder phenomena in alloys. In an alloy–to be specific, a *binary* alloy–we have a lattice structure consisting of two types of atoms, say 1 and 2, numbering N_1 and N_2, respectively. In a configuration characterized by the numbers N_{11}, N_{22} and N_{12} of the three types of nearest-neighbor pairs, the configurational energy of the alloy may be written as

$$E = \varepsilon_{11}N_{11} + \varepsilon_{22}N_{22} + \varepsilon_{12}N_{12}, \tag{9}$$

where ε_{11}, ε_{22} and ε_{12} have an obvious meaning. As in the case of a ferromagnet, the various numbers appearing in the expression for E may be expressed in terms of the numbers N, N_1 and N_{11} (of which only N_{11} is variable here). Equation (9) then takes the form

$$\begin{aligned} E &= \varepsilon_{11}N_{11} + \varepsilon_{22}\left(\tfrac{1}{2}qN - qN_1 + N_{11}\right) + \varepsilon_{12}(qN_1 - 2N_{11}) \\ &= \tfrac{1}{2}q\varepsilon_{22}N + q(\varepsilon_{12} - \varepsilon_{22})N_1 + (\varepsilon_{11} + \varepsilon_{22} - 2\varepsilon_{12})N_{11}. \end{aligned} \tag{10}$$

The correspondence between this system and the lattice gas, in the canonical ensemble, is now straightforward:

The lattice gas		*The binary alloy*
$N_a, N - N_a$	↔	$N_1, N - N_1(= N_2)$
$-\varepsilon_0$	↔	$(\varepsilon_{11} + \varepsilon_{22} - 2\varepsilon_{12})$
A	↔	$A - \frac{1}{2}q\varepsilon_{22}N - q(\varepsilon_{12} - \varepsilon_{22})N_1$

The correspondence with a ferromagnet can be established likewise; in particular, this requires that $\varepsilon_{11} = \varepsilon_{22} = -J$ and $\varepsilon_{12} = +J$.

At absolute zero, our alloy will be in the state of minimum configurational energy, which would also be the state of maximum configurational order. We expect that the two types of atoms would then occupy *mutually exclusive* sites, so that one might speak of atoms 1 being only at sites a and atoms 2 being only at sites b. As temperature rises, an exchange of sites would result and, in the face of thermal agitation, the order in the system would start giving way. Ultimately, the two types of atoms would get so "mixed up" that the very notion of the sites a being the "right" ones for atoms 1 and the sites b being the "right" ones for atoms 2 would break down; the system would then behave, from the *configurational* point of view, as an assembly of $N_1 + N_2$ atoms of essentially the same species.

11.5. Ising model in the zeroth approximation

In 1928 Gorsky attempted a statistical study of order–disorder transitions in binary alloys on the basis of the assumption that the work expended in transferring an atom from an "ordered" position to a "disordered" one (or, in other words, from a "right" site to a "wrong" one) is directly proportional to the *degree of order* prevailing in the system. This idea was further developed by Bragg and Williams (1934, 1935) who, for the first time, introduced the concept of *long-range order* in the sense we understand it now and, with relatively simple mathematics, obtained results that could explain the main qualitative features of the relevant experimental data. The basic assumption in the Bragg–Williams approximation is that the energy of an individual atom in the given system is determined by the (*average*) degree of order prevailing in the entire system rather than by the (*fluctuating*) configurations of the neighboring atoms. In this sense, the approximation is equivalent to the *mean molecular field* (or the *internal field*) theory of Weiss, which was put forward in 1907 to explain the magnetic behavior of ferromagnetic domains.

It seems natural to call this approximation the *zeroth* approximation, for its features are totally insensitive to the detailed structure, or even to the dimensionality, of the lattice. We expect that the results following from this approximation will become more reliable as the number of neighbors interacting with a given atom increases (i.e. as $q \to \infty$), thus diminishing the importance of local, fluctuating influences.[10]

We define a long-range order parameter L in a given configuration by the very suggestive relationship

$$L = \frac{1}{N}\sum_i \sigma_i = \frac{N_+ - N_-}{N} = 2\frac{N_+}{N} - 1 \quad (-1 \leq L \leq +1), \tag{1}$$

whence

$$N_+ = \frac{N}{2}(1 + L) \quad \text{and} \quad N_- = \frac{N}{2}(1 - L). \tag{2}$$

The magnetization M is then given by

$$M = (N_+ - N_-)\mu = N\mu L \quad (-N\mu \leq M \leq +N\mu); \tag{3}$$

the parameter L is, therefore, a direct measure of the net magnetization in the system. For a completely random configuration, $\overline{N}_+ = \overline{N}_- = \frac{1}{2}N$; the expectation values of both L and M are then identically zero.

Now, in the spirit of the present approximation, we replace the first part of the Hamiltonian (11.3.8) by the expression $-J\left(\frac{1}{2}q\overline{\sigma}\right)\sum_i \sigma_i$, i.e. for a given σ_i, we replace each of the $q\sigma_j$ by $\overline{\sigma}$ while the factor $\frac{1}{2}$ is included to avoid duplication in the counting of the nearest-neighbor pairs. Making use of eqn. (1), and noting that $\overline{\sigma} \equiv \overline{L}$, we obtain for the total configurational energy of the system

$$E = -\tfrac{1}{2}(qJ\overline{L})NL - (\mu B)NL. \tag{4}$$

The expectation value of E is then given by

$$U = -\tfrac{1}{2}qJN\overline{L}^2 - \mu BN\overline{L}. \tag{5}$$

In the same approximation, the difference $\Delta\varepsilon$ between the overall configurational energy of an "up" spin and the overall configurational energy of a "down" spin–specifically, the energy expended in changing any "up" spin into a "down" one–is given by, see again eqn. (11.3.8),

$$\begin{aligned} \Delta\varepsilon &= -J(q\overline{\sigma})\Delta\sigma - \mu B\Delta\sigma \\ &= 2\mu\left(\frac{qJ}{\mu}\overline{\sigma} + B\right), \end{aligned} \tag{6}$$

for $\Delta\sigma = -2$ here. The quantity $qJ\overline{\sigma}/\mu$ thus plays the role of the *internal (molecular) field* of Weiss; it is determined by (i) the mean value of the long-range order prevailing in the system and by (ii) the strength of the coupling, qJ, between a given spin i and all its q nearest neighbors. The relative values of the equilibrium numbers $\overline{N}_+$ and $\overline{N}_-$ then follow from the *Boltzmann principle*, viz.

$$\overline{N}_-/\overline{N}_+ = \exp(-\Delta\varepsilon/kT) = \exp\{-2\mu(B' + B)/kT\}, \tag{7}$$

where B' denotes the internal molecular field:

$$B' = qJ\overline{\sigma}/\mu = qJ(\overline{M}/N\mu^2). \tag{8}$$

Substituting (2) into (7), and keeping in mind eqn. (8), we obtain for $\overline{L}$

$$\frac{1-\overline{L}}{1+\overline{L}} = \exp\{-2(qJ\overline{L} + \mu B)/kT\} \tag{9}$$

or, equivalently,

$$\frac{qJ\overline{L} + \mu B}{kT} = \frac{1}{2}\ln\frac{1+\overline{L}}{1-\overline{L}} = \tanh^{-1}\overline{L}. \tag{10}$$

To investigate the possibility of spontaneous magnetization, we let $B \to 0$, which leads to the relationship

$$\overline{L}_0 = \tanh\left(\frac{qJ\overline{L}_0}{kT}\right). \tag{11}$$

Equation (11) may be solved graphically; see Fig. 11.6. For any temperature T, the appropriate value of $\overline{L}_0(T)$ is determined by the point of intersection of (i) the straight line $y = L_0$ and (ii) the curve $y = \tanh(qJL_0/kT)$. Clearly, the solution $\overline{L}_0 = 0$ is always there; however, we are interested in *nonzero* solutions, if any. For those, we note that, since the slope of the curve (ii) varies from the initial value qJ/kT to the final value zero while the slope of the line (i) is throughout unity, an intersection other than the one at the origin is possible if, and only if,

$$qJ/kT > 1, \tag{12}$$

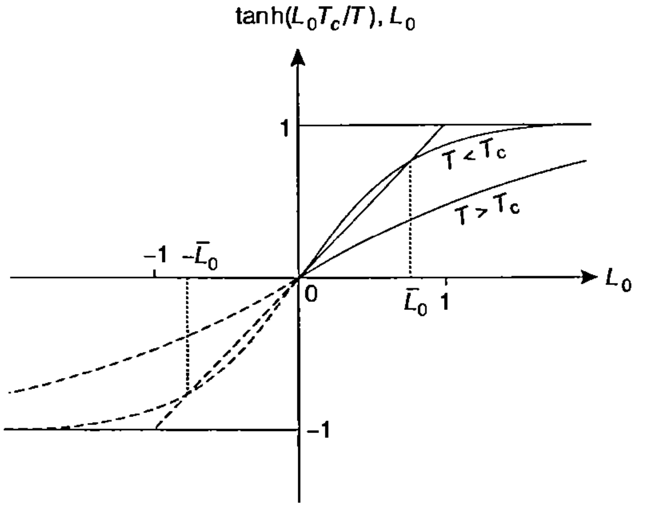


Fig. 11.6. The graphical solution of eqn. (11), with $T_c = qJ/k$.

that is, if

$$T < qJ/k = T_c, \text{ say.} \tag{13}$$

We thus obtain a *critical temperature* T_c, below which the system *can* acquire a nonzero spontaneous magnetization and above which it *cannot.* It is natural to identify T_c with the *Curie temperature* of the system–the temperature that marks a transition from the ferromagnetic to the paramagnetic behavior of the system or vice versa.

It is clear from Fig. 11.6, as well as from eqn. (11), that if $\bar{L}_0$ is a solution of the problem, than $-\bar{L}_0$ is also a solution. The reason for this duplicity of solutions is that, in the absence of an external field, there is no way of assigning a "positive", as opposed to a "negative", direction to the alignment of spins. In fact, if B were zero right from the beginning, then the positive solution of eqn. (11) would be as likely to occur as the negative one–with the result that the *true* expectation value of $L_0(T)$ would be zero. If, on the other hand, B were nonzero to begin with (to be definite, say $B > 0$), then eqn. (10) for $\bar{L}(B, T)$ would admit only positive solutions and, in the limit $B \to 0+$, we would obtain a positive $\bar{L}_0(T)$. The "up–down symmetry" will thus be broken and we will see a *net* alignment of spins in the "up" direction.[11]

The precise variation of $\bar{L}_0(T)$ with T can be obtained by solving eqn. (11) numerically; the general trend, however, can be seen from Fig. 11.6 as well. We note that at $T = qJ/k(= T_c)$ the straight line $y = L_0$ is tangential to the curve $y = \tanh(qJL_0/kT)$ at the origin; the relevant solution then is: $\bar{L}_0(T_c) = 0$. As T decreases, the initial slope of the curve becomes larger and the relevant point of intersection moves rapidly away from the origin; accordingly, $\bar{L}_0(T)$ rises rapidly as T decreases below T_c. To obtain an approximate dependence of $\bar{L}_0(T)$ on T near $T = T_c$, we write (11) in the form $\bar{L}_0 = \tanh(\bar{L}_0 T_c/T)$ and use the approximation $\tanh x \simeq x - x^3/3$, to obtain

$$\bar{L}_0(T) \approx \{3(1 - T/T_c)\}^{1/2} \quad (T \lesssim T_c, B \to 0). \tag{14}$$

On the other hand, as $T \to 0$, $\bar{L}_0 \to 1$, in accordance with the asymptotic relationship

$$\bar{L}_0(T) \approx 1 - 2\exp(-2T_c/T) \quad \{(T/T_c) \ll 1\}. \tag{15}$$

Figure 11.7 shows a plot of $\bar{L}_0(T)$ versus T, along with the relevant experimental results for iron, nickel, cobalt and magnetite; we find the agreement not too bad.

The field-free configurational energy and the field-free specific heat of the system are given by, see eqn. (5),

$$U_0(T) = -\tfrac{1}{2}qJN\bar{L}_0^2 \tag{16}$$

and

$$C_0(T) = -qJN\bar{L}_0\frac{d\bar{L}_0}{dT} = \frac{Nk\bar{L}_0^2}{\dfrac{(T/T_c)^2}{1-\bar{L}_0^2} - \dfrac{T}{T_c}}, \tag{17}$$

where the last step has been carried out with the help of eqn. (11). Thus, for all $T > T_c$, both $U_0(T)$ and $C_0(T)$ are identically zero. However, the value of the

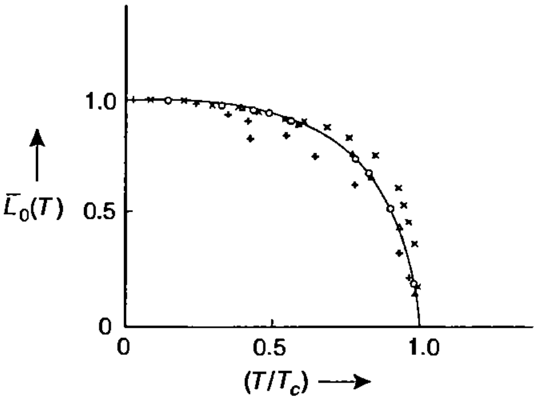


FIG. 11.7. The spontaneous magnetization of a Weiss ferromagnet as a function of temperature. The experimental points (after Becker) are for iron (×), nickel (o), cobalt (Δ) and magnetite (+).

specific heat at the transition temperature T_c, as approached from below, turns out to be, see eqns (14) and (17),

$$C_0(T_c-) = \lim_{x\to 0}\left\{\frac{Nk\cdot 3x}{\dfrac{(1-x)^2}{1-3x}-(1-x)}\right\} = \frac{3}{2}Nk. \tag{18}$$

The specific heat, therefore, displays a discontinuity at the transition point. On the other hand, as $T \to 0$, the specific heat vanishes, in accordance with the formula, see eqns (15) and (17),

$$C_0(T) \approx 4Nk\left(\frac{T_c}{T}\right)^2 \exp(-2T_c/T). \tag{19}$$

The full trend of the function $C_0(T)$ is shown in Fig. 11.8.

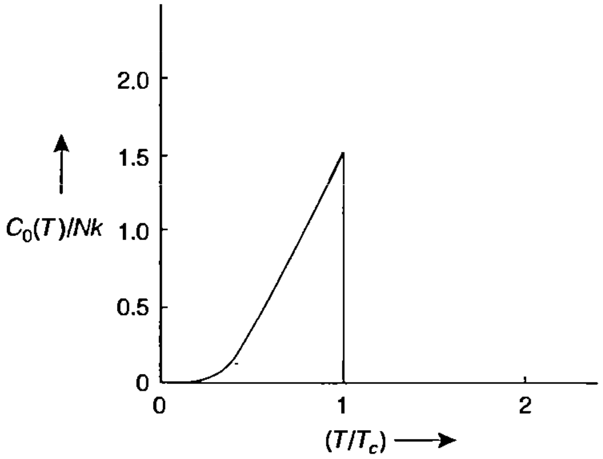


FIG. 11.8. The field-free specific heat of a Weiss ferromagnet as a function of temperature

It is important to note that the vanishing of the configurational energy and the specific heat of the system at temperatures above T_c is directly related to the fact that, in the present approximation, the configurational order prevailing in the system at lower temperatures is completely wiped out as $T \to T_c$. Consequently, the configurational entropy and the configurational energy of the system attain their maximum values at $T = T_c$; beyond that, the system remains thermodynamically "inert". As a check, we evaluate the configurational entropy of the system at $T = T_c$; with the help of eqns (11) and (17), we get

$$S_0(T_c) = \int_0^{T_c} \frac{C_0(T)\,dT}{T} = -qJN \int_1^0 \frac{\bar{L}_0}{T}\,d\bar{L}_0$$

$$= Nk \int_0^1 (\tanh^{-1} \bar{L}_0)\,d\bar{L}_0 = Nk \ln 2, \tag{20}$$

precisely the result we expect for a system capable of 2^N *equally likely* microstates.[12] The fact that all these microstates are *equally likely* to occur is again related to the fact that for $T \geq T_c$ there is no (configurational) order in the system.

We now proceed to study the magnetic susceptibility of the system. Using eqn. (10), we get

$$\chi(B, T) = \left(\frac{\partial \bar{M}}{\partial B}\right)_T = N\mu \left(\frac{\partial \bar{L}}{\partial B}\right)_T = \frac{N\mu^2}{k} \frac{1 - \bar{L}^2(B, T)}{T - T_c\{1 - \bar{L}^2(B, T)\}}. \tag{21}$$

For $\bar{L} \ll 1$ (which is true at high temperatures for a wide range of B but is also true near T_c if B is small), we obtain the *Curie–Weiss law*

$$\chi_0(T) \approx (N\mu^2/k)(T - T_c)^{-1} \quad (T \gtrsim T_c, B \to 0), \tag{22a}$$

which may be compared with the *Curie law* derived earlier for a paramagnetic system; see eqn. (3.9.12). For T less than, but close to, T_c we utilize eqn. (14) as well and get

$$\chi_0(T) \approx (N\mu^2/2k)(T_c - T)^{-1} \quad (T \lesssim T_c, B \to 0). \tag{22b}$$

Experimentally, one finds that the Curie–Weiss law is satisfied with considerable accuracy, except that the empirical value of T_c thus obtained is always somewhat larger than the true transition temperature of the material; for instance, in the case of nickel, the empirical value of T_c obtained in this manner turns out to be about 650 K while the actual transition takes place at about 631 K. In passing, we add that, as $T \to 0$, the low-field susceptibility vanishes, in accordance with the formula

$$\chi_0(T) \approx \frac{4N\mu^2}{kT} \exp(-2T_c/T). \tag{23}$$

Finally, we examine the relationship between $\bar{L}$ and B at $T = T_c$. Using, once again, eqn. (10) and employing the approximation $\tanh^{-1} x \simeq x + x^3/3$, we get

$$\bar{L} \approx (3\mu B/kT_c)^{1/3} \quad (T = T_c, B \to 0). \tag{24}$$

At this point we wish to emphasize the remarkable similarity that exists between the critical behavior of a gas–liquid system obeying van der Waals equation of state and that of a magnetic system treated in the Bragg–Williams approximation. Even though the two systems are physically very different, the level of approximation is such that the exponents governing power-law behavior of the various physical quantities in the critical region turn out to be the same; compare, for instance, eqn. (14) with (11.2.11), eqns (22a, b) with (11.2.13,14), eqn. (24) with (11.2.8)–along with the behavior of the specific heat as well. This sort of similarity will be seen again and again whenever we employ an approach similar in spirit to the mean field approach of this section.

Before we close our discussion of the so-called zeroth approximation, we would like to demonstrate that it corresponds exactly to the *random mixing approximation* (which was employed originally in the theory of binary solutions). According to eqn. (11.3.19), the mean configurational energy in the absence of the external field is given by

$$U_0 = -J\left(\tfrac{1}{2}qN - 2q\overline{N}_+ + 4\overline{N}_{++}\right). \tag{25}$$

On the other hand, eqns (2) and (16) of the present approach give

$$\overline{N}_+ = \tfrac{1}{2}N(1+\overline{L}_0) \quad \text{and} \quad U_0 = -\tfrac{1}{2}qJN\overline{L}_0^2. \tag{26}$$

Combining (25) and (26), we obtain

$$\overline{N}_{++} = \tfrac{1}{8}qN(1+\overline{L}_0)^2, \tag{27}$$

so that

$$\frac{\overline{N}_{++}}{\frac{1}{2}qN} = \left(\frac{\overline{N}_+}{N}\right)^2 \tag{28}$$

Thus, the probability of having an "up–up" nearest-neighbor pair of spins in the lattice is precisely equal to the square of the probability of having an "up" spin; in other words, there does not exist, in spite of the presence of a nearest-neighbor interaction (characterized by the coupling constant J), any specific correlation between the neighboring spins of the lattice. Put differently, there does not exist any *short-range order* in the system, apart from what follows statistically from the long-range order (characterized by the parameter $\overline{L}$). It follows that, in the present approximation, our system consists of a specific number of "up" spins, viz. $N(1+\overline{L})/2$, and a corresponding number of "down" spins, viz. $N(1-\overline{L})/2$, distributed over the N lattice sites *in a completely random manner*–similar to the mixing of $N(1+\overline{L})/2$ atoms of one kind with $N(1-\overline{L})/2$ atoms of another kind in a completely random manner to obtain a binary solution of N atoms; see also Problem 11.4. For this sort of mixing, we obviously have

$$\overline{N}_{++} = \frac{1}{2}qN\left(\frac{1+\overline{L}}{2}\right)^2; \quad \overline{N}_{--} = \frac{1}{2}qN\left(\frac{1-\overline{L}}{2}\right)^2; \tag{29a}$$

$$\overline{N}_{+-} = 2\cdot\frac{1}{2}qN\left(\frac{1+\overline{L}}{2}\right)\left(\frac{1-\overline{L}}{2}\right), \tag{29b}$$

with the result that

$$\frac{\overline{N}_{++}\overline{N}_{--}}{(\overline{N}_{+-})^2} = \frac{1}{4}. \tag{30}$$

11.6. Ising model in the first approximation

The approaches considered in the preceding section have a natural generalization towards an improved approximation. The mean field approach leads naturally to the *Bethe approximation* (Bethe, 1935; Rushbrooke, 1938), which treats the interaction of a given spin with its nearest neighbors somewhat more accurately. The random mixing approach, on the other hand, leads to the *quasi-chemical* approximation (Guggenheim, 1935; Fowler and Guggenheim, 1940), which takes into account the *specific* short-range order of the lattice–over and above the one that follows statistically from the long-range order. As shown by Guggenheim (1938) and by Chang (1939), the two methods yield identical results for the Ising model. It seems worthwhile to mention here that the extension of these approximations to higher orders, or their application to the Heisenberg model, does not produce identical results.

In the Bethe approximation, a given spin σ_0 is regarded as the central member of a group, which consists of this spin and its q nearest neighbors, and in writing down the Hamiltonian of this group the interaction between the central spin and the q neighbors is taken into account exactly while the interaction of these neighbors with other spins in the lattice is taken into account again through a mean molecular field, B' say. Thus

$$H_{q+1} = -\mu B\sigma_0 - \mu(B + B')\sum_{j=1}^{q}\sigma_j - J\sum_{j=1}^{q}\sigma_0\sigma_j, \tag{1}$$

B being the external magnetic field acting on the lattice. The internal field B' is determined by the *condition of self-consistency*, which requires that the mean value, $\overline{\sigma}_0$, of the central spin be the same as the mean value, $\overline{\sigma}_j$, of any of the q neighbors. The partition function Z of this group of spins as a whole is given by

$$\begin{aligned} Z &= \sum_{\sigma_0,\sigma_j=\pm 1} \exp\left[\frac{1}{kT}\left\{\mu B\sigma_0 + \mu(B+B')\sum_{j=1}^{q}\sigma_j + J\sum_{j=1}^{q}\sigma_0\sigma_j\right\}\right] \\ &= \sum_{\sigma_0,\sigma_j=\pm 1} \exp\left[\alpha\sigma_0 + (\alpha+\alpha')\sum_{j=1}^{q}\sigma_j + \gamma\sum_{j=1}^{q}\sigma_0\sigma_j\right], \end{aligned} \tag{2}$$

where

$$\alpha = \frac{\mu B}{kT}, \quad \alpha' = \frac{\mu B'}{kT} \quad \text{and} \quad \gamma = \frac{J}{kT}. \tag{3}$$

Now, the right-hand side of (2) can be written as a sum of two terms, one pertaining to $\sigma_0 = +1$ and the other to $\sigma_0 = -1$, i.e.

$$Z = Z_+ + Z_-,$$

where

$$\begin{aligned} Z_\pm &= \sum_{\sigma_j=\pm 1} \exp\left[\pm\alpha + (\alpha + \alpha' \pm \gamma)\sum_{j=1}^{q}\sigma_j\right] \\ &= e^{\pm\alpha}\left[2\cosh(\alpha + \alpha' \pm \gamma)\right]^q. \end{aligned} \tag{4}$$

The mean val [illegible] of the central spin is then given by

$$\overline{\sigma}_0 = \frac{Z_+ - Z_-}{Z}, \tag{5}$$

while the mean value of any one of its neighbors is given by, see (2),

$$\begin{aligned}\overline{\sigma}_j = \overline{\frac{1}{q}\left(\sum_{j=1}^{q} \sigma_j\right)} &= \frac{1}{q}\left(\frac{1}{Z}\frac{\partial Z}{\partial \alpha'}\right) \\ &= \frac{1}{Z}\{Z_+ \tanh(\alpha + \alpha' + \gamma) + Z_- \tanh(\alpha + \alpha' - \gamma)\}.\end{aligned} \tag{6}$$

Equating (5) and (6), we get

$$Z_+\{1 - \tanh(\alpha + \alpha' + \gamma)\} = Z_-\{1 + \tanh(\alpha + \alpha' - \gamma)\}. \tag{7}$$

Substituting for Z_+ and Z_- from (4), we finally obtain

$$e^{2\alpha'} = \left\{\frac{\cosh(\alpha + \alpha' + \gamma)}{\cosh(\alpha + \alpha' - \gamma)}\right\}^{q-1} \tag{8}$$

Equation (8) determines α' which, in turn, determines the magnetic behavior of the lattice.

To study the possibility of spontaneous magnetization, we set $\alpha (= \mu B/kT) = 0$. Equation (8) then reduces to

$$\alpha' = \frac{q-1}{2}\ln\left\{\frac{\cosh(\alpha' + \gamma)}{\cosh(\alpha' - \gamma)}\right\} \tag{9}$$

In the absence of interactions ($\gamma = 0$), α' is clearly zero. In the presence of interactions ($\gamma \neq 0$), α' may still be zero unless γ exceeds a certain critical value, γ_c say. To determine this value, we expand the right-hand side of (9) as a Taylor series around $\alpha' = 0$, with the result

$$\alpha' = (q-1)\tanh\gamma\left\{\alpha' - \operatorname{sech}^2\gamma\frac{\alpha'^3}{3} + \cdots\right\}. \tag{10}$$

We note that, for all γ, $\alpha' = 0$ is one possible solution of the problem; this, however, does not interest us. A nonzero solution requires that

$$(q-1)\tanh\gamma > 1,$$

that is,

$$\gamma > \gamma_c = \tanh^{-1}\left(\frac{1}{q-1}\right) = \frac{1}{2}\ln\left(\frac{q}{q-2}\right). \tag{11}$$

In terms of temperature, this means that

$$T < T_c = \frac{2J}{k}\Big/\ln\left(\frac{q}{q-2}\right), \tag{12}$$

which determines the *Curie temperature* of the lattice. From (10), it also follows that for temperatures less than, but close to, the Curie temperature

$$\alpha' \simeq \{3\cosh^2\gamma_c[(q-1)\tanh\gamma - 1]\}^{1/2} \simeq \{3(q-1)(\gamma-\gamma_c)\}^{1/2}$$
$$\simeq \left\{3(q-1)\frac{J}{kT_c}\left(1-\frac{T}{T_c}\right)\right\}^{1/2} \tag{13}$$

The parameter $\overline{L}$, which is a measure of the long-range order in the system, is, by definition, equal to $\overline{\sigma}$. From eqns (5) and (7), we get

$$\overline{L} = \frac{(Z_+/Z_-)-1}{(Z_+/Z_-)+1} = \frac{\sinh(2\alpha+2\alpha')}{\cosh(2\alpha+2\alpha')+\exp(-2\gamma)}. \tag{14}$$

In the limit $B \to 0$ (which means $\alpha \to 0$) and at temperatures less than, but close to, the Curie temperature ($\gamma \gtrsim \gamma_c$; $\alpha' \simeq 0$), we obtain

$$\overline{L}_0 = \frac{\sinh(2\alpha')}{\cosh(2\alpha')+\exp(-2\gamma)} \simeq \frac{2\alpha'}{1+(q-2)/q} = \frac{q}{q-1}\alpha'. \tag{15}$$

Substituting from (12) and (13), we get

$$\overline{L}_0 \simeq \left[3\frac{q}{q-1}\left\{\frac{q}{2}\ln\left(\frac{q}{q-2}\right)\right\}\left(1-\frac{T}{T_c}\right)\right]^{1/2} \tag{16}$$

We note that, for $q \gg 1$, eqns (12) and (16) reduce to their zeroth-order counterparts (11.5.13) and (11.5.14), respectively; in either case, as $T \to T_c$ from below, $\overline{L}_0$ vanishes as $(T_c - T)^{1/2}$. We also note that the spontaneous magnetization curve in the present approximation has the same general shape as in the zeroth approximation; see Fig. 11.7. Of course, in the present case the curve depends explicitly on the coordination number q, being steepest for small q and becoming less steep as q increases–tending ultimately to the limiting form given by the zeroth approximation.

We shall now study *correlations* that might exist among neighboring spins in the lattice. For this, we evaluate the numbers $\overline{N}_{++}, \overline{N}_{--}$ and $\overline{N}_{+-}$ in terms of the parameters α, α' and γ, and compare the resulting expressions with the ones obtained under the mean field approximation. Carrying out summations in (2) over all the spins (of the group) except σ_0 and σ_1, we obtain

$$Z = \sum_{\sigma_0,\sigma_1=\pm 1} [\exp\{\alpha\sigma_0 + (\alpha+\alpha')\sigma_1 + \gamma\sigma_0\sigma_1\}\{2\cosh(\alpha+\alpha'+\gamma\sigma_0)\}^{q-1}]. \tag{17}$$

Writing this as a sum of three parts pertaining, respectively, to the cases (i) $\sigma_0 = \sigma_1 = +1$, (ii) $\sigma_0 = \sigma_1 = -1$ and (iii) $\sigma_0 = -\sigma_1 = \pm 1$, we have

$$Z = Z_{++} + Z_{--} + Z_{+-}, \tag{18}$$

where, naturally enough,

$$\overline{N}_{++} : \overline{N}_{--} : \overline{N}_{+-} :: Z_{++} : Z_{--} : Z_{+-}. \tag{19}$$

We thus obtain, using (8) as well,

$$\overline{N}_{++} \propto e^{(2\alpha+\alpha'+\gamma)}\{2\cosh(\alpha+\alpha'+\gamma)\}^{q-1},$$

$$\begin{aligned}\overline{N}_{--} &\propto e^{(-2\alpha-\alpha'+\gamma)}\{2\cosh(\alpha+\alpha'-\gamma)\}^{q-1}\\ &= e^{(-2\alpha-3\alpha'+\gamma)}\{2\cosh(\alpha+\alpha'+\gamma)\}^{q-1},\end{aligned}$$

and

$$\begin{aligned}\overline{N}_{+-} &\propto e^{(-\alpha'-\gamma)}\{2\cosh(\alpha+\alpha'+\gamma)\}^{q-1} + e^{(\alpha'-\gamma)}\{2\cosh(\alpha+\alpha'-\gamma)\}^{q-1}\\ &= 2e^{(-\alpha'-\gamma)}\{2\cosh(\alpha+\alpha'+\gamma)\}^{q-1}.\end{aligned}$$

Normalizing with the help of the relationship

$$\overline{N}_{++} + \overline{N}_{--} + \overline{N}_{+-} = \tfrac{1}{2}qN, \tag{20}$$

we obtain the desired results

$$(\overline{N}_{++}, \overline{N}_{--}, \overline{N}_{+-}) = \frac{1}{2}qN\frac{(e^{2\alpha+2\alpha'+\gamma}, e^{-2\alpha-2\alpha'+\gamma}, 2e^{-\gamma})}{2\{e^{\gamma}\cosh(2\alpha+2\alpha') + e^{-\gamma}\}}, \tag{21}$$

whence

$$\frac{\overline{N}_{++}\overline{N}_{--}}{(\overline{N}_{+-})^2} = \frac{1}{4}e^{4\gamma} = \frac{1}{4}e^{4J/kT}. \tag{22}$$

The last result differs significantly from the one, namely (11.5.30), that followed from the random mixing approximation. The difference lies in the extra factor exp $(4J/kT)$ which, for $J > 0$, favors the formation of parallel-spin pairs ↑↑ and ↓↓, as opposed to antiparallel-spin pairs ↑↓ and ↓↑. In fact, one may regard the elementary process

$$\uparrow\uparrow + \downarrow\downarrow \Leftrightarrow 2\uparrow\downarrow, \tag{23}$$

which leaves the total numbers of "up" spins and "down" spins unaltered, as a kind of a "chemical reaction" which, proceeding from left to right, is endothermic (requiring an amount of energy $4J$ to get through) and, proceeding from right to left, is exothermic (releasing an amount of energy $4J$). Equation (22) then constitutes the *law of mass action* for this reaction, the expression on the right-hand side being the *equilibrium constant* of the reaction. Historically, eqn. (22) was adopted by Guggenheim as the starting point of his "quasi-chemical" treatment of the Ising model; only later on did he show that his treatment was equivalent to the Bethe approximation studied here.

Equation (22) tells us that, for $J > 0$, there exists among *like* neighbors (↑ and ↑ or ↓ and ↓) a *positive* correlation and among *unlike* neighbors (↑ and ↓) a *negative* correlation, and that these correlations are a direct consequence of the nearest-neighbor interaction. Accordingly, there must exist a *specific* short-range order in the system, over and above the one that follows statistically from the long-range order. To see this explicitly, we note that even when long-range order disappears ($\alpha + \alpha' = 0$) some short-range order still persists; from eqn. (21), we obtain

$$(\overline{N}_{++}, \overline{N}_{--}, \overline{N}_{+-})_{\overline{L}=0} = \frac{1}{2}qN\frac{(e^{\gamma}, e^{\gamma}, 2e^{-\gamma})}{4\cosh\gamma} \tag{24}$$

which, only in the limit $\gamma \to 0$, goes over to the rand mixing result, see eqn. (11.5.29) with $\bar{L} = 0$,

$$(\overline{N}_{++}, \overline{N}_{--}, \overline{N}_{+-})_{\bar{L}=0} = \frac{1}{2} qN \frac{(1, 1, 2)}{4}. \tag{25}$$

In the zeroth approximation, eqn. (25) is supposed to hold at *all* temperatures above T_c; we now find that a better approximation at these temperatures is provided by (24).

Next, we evaluate the configurational energy U_0 and the specific heat C_0 of the lattice in the absence of the external field ($\alpha = 0$). In view of eqn. (11.5.25),

$$U_0 = -J\left(\tfrac{1}{2}qN - 2q\overline{N}_+ + 4\overline{N}_{++}\right)_{\alpha=0}. \tag{26}$$

The expression for $\overline{N}_{++}$ is given by eqn. (21) while that for $\overline{N}_+$ can be obtained from (14):

$$(\overline{N}_+)_{\alpha=0} = \frac{1}{2}N(1 + \bar{L}_0) = \frac{1}{2}N \frac{\exp(2\alpha') + \exp(-2\gamma)}{\cosh(2\alpha') + \exp(-2\gamma)}. \tag{27}$$

Equation (26) then gives

$$U_0 = -\frac{1}{2}qJN \frac{\cosh(2\alpha') - \exp(-2\gamma)}{\cosh(2\alpha') + \exp(-2\gamma)}, \tag{28}$$

where α' is determined by eqn. (9). For $T > T_c$, $\alpha' = 0$, so

$$U_0 = -\frac{1}{2}qJN \frac{1 - \exp(-2\gamma)}{1 + \exp(-2\gamma)} = -\frac{1}{2}qJN \tanh \gamma. \tag{29}$$

Obviously, this result arises solely from the *short-range order* that persists in the system even above T_c. As for the specific heat, we get

$$C_0/Nk = \tfrac{1}{2}q\gamma^2 \operatorname{sech}^2 \gamma \quad (T > T_c). \tag{30}$$

As $T \to \infty$, C_0 vanishes like T^{-2}. We note that a nonzero specific heat above the transition temperature is a welcome feature of the present approximation, for it brings our model somewhat closer to real physical systems. In this connection, we recall that in the previous approximation the specific heat was zero for *all* $T > T_c$. Figure 11.9 shows the specific heat of an Ising lattice, with coordination number 4, as given by the Bethe approximation; for comparison, the result of the previous approximation is also included.

We are now in a position to study the specific heat discontinuity at $T = T_c$. The limiting value of C_0, as T approaches T_c from above, can be obtained from eqn. (30) by letting $\gamma \to \gamma_c$. One obtains, with the help of eqn. (11),

$$\frac{1}{Nk} C_0(T_c+) = \frac{1}{2}q\gamma_c^2 \operatorname{sech}^2 \gamma_c = \frac{1}{8}\frac{q^2(q-2)}{(q-1)^2}\left\{\ln\left(\frac{q}{q-2}\right)\right\}^2 \tag{31}$$

To obtain the corresponding result as T approaches T_c from below, we must use the general expression (28) for U_0, with $\alpha' \to 0$ as $\gamma \to \gamma_c$. Expanding (28) in

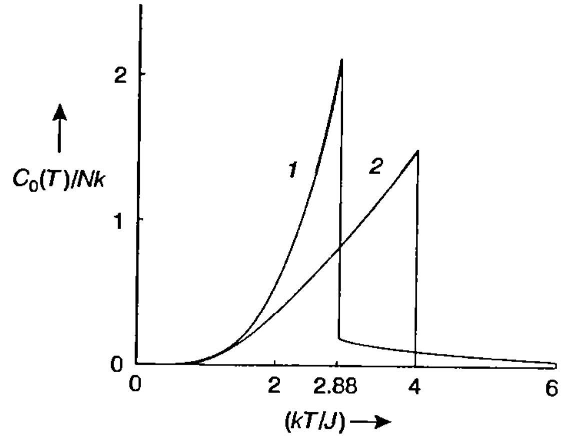


FIG. 11.9. The field-free specific heat of an Ising lattice with coordination number 4. Curve 1 obtains in the Bethe approximation, curve 2 in the Bragg–Williams approximation

powers of the quantities $(\gamma - \gamma_c)$ and α', and making use of eqn. (13), we obtain for $(1 - T/T_c) \ll 1$

$$\begin{aligned} U_0 &= -\frac{1}{2}qJN\left[\frac{1}{(q-1)} + \frac{q(q-2)}{(q-1)^2}(\gamma - \gamma_c) + \frac{q(q-2)}{(q-1)^2}\alpha'^2 + \cdots\right] \\ &= -\frac{1}{2}qJN\left[\frac{1}{(q-1)} + \frac{q(q-2)(3q-2)}{(q-1)^2}\frac{J}{kT_c}\left(1 - \frac{T}{T_c}\right) + \cdots\right] \end{aligned} \tag{32}$$

Differentiating with respect to T and substituting for T_c, we obtain

$$\frac{1}{Nk}C_0(T_c-) = \frac{1}{8}\frac{q^2(q-2)(3q-2)}{(q-1)^2}\left\{\ln\left(\frac{q}{q-2}\right)\right\}^2 \tag{33}$$

which is $(3q-2)$ times larger than the corresponding result for $T = T_c+$; cf. eqn. (31). The specific-heat discontinuity at the transition point is, therefore, given by

$$\frac{1}{Nk}\Delta C_0 = \frac{3}{8}\frac{q^2(q-2)}{(q-1)}\left\{\ln\left(\frac{q}{q-2}\right)\right\}^2. \tag{34}$$

One may check that, for $q \gg 1$, the foregoing results go over to the ones following from the zeroth approximation.

Finally, we examine the relationship between $\bar{L}$ and B at $T = T_c$. Using eqns (8) and (14), with both α and $\alpha' \ll 1$ while $\gamma = \gamma_c$, we get

$$\bar{L} \approx \{3q^2\mu B/(q-1)(q-2)kT_c\}^{1/3} \quad (T = T_c, B \to 0). \tag{35}$$

For the behavior of χ_0, see Problem 11.16.

In passing, we note that, according to eqn. (12), the transition temperature for a lattice with $q = 2$ is zero, which essentially means that a one-dimensional Ising chain *does not* undergo a phase transition. This result is in complete agreement with the one following from an exact treatment of the one-dimensional lattice; see Sec. 12.1. In fact, for a lattice with $q = 2$, any results following from the Bethe approximation are completely identical with the corresponding exact results (see Problem 12.3); on the other hand, the Bragg–Williams approximation is least reliable when $q = 2$.

That T_c for $q = 2$ is zero (rather than $2J/k$) is in line with the fact that for all q the first approximation yields a transition temperature closer to the correct value of T_c than does the zeroth approximation. The same is true of the amplitudes that determine the quantitative behavior of the various physical quantities near $T = T_c$, though the exponents in the various power laws governing this behavior remain the same; cf., for instance, eqn. (16) with (11.5.14), eqn. (35) with (11.5.24) as well as the behavior of the specific heat. In fact, one finds that successive approximations of the mean field approach, while continuing to improve the theoretical value of T_c *and* the quantitative behavior of the various physical quantities (as given by the amplitudes), do not modify their qualitative behavior (as determined by the exponents). For an account of the higher-order approximations, see Domb (1960).

One important virtue of the Bethe approximation is that it brings out the role of the dimensionality of the lattice in bringing about a phase transition in the system. The fact that $T_c = 0$ for $q = 2$ and thereon it increases steadily with q leads one to infer that, while a linear Ising chain does not undergo phase transition at any finite T, higher dimensionality does promote the phenomenon. One may, in fact, argue that the absence of a phase transition in a one-dimensional chain is essentially due to the fact that, the interactions being severely short-ranged, "communication" between any two parts of the chain can be completely disrupted by a single defect in between. The situation remains virtually unaltered even if the range of interactions is allowed to increase–so long as it remains finite. Only when interactions become truly long-ranged, with $J_{ij} \sim |i - j|^{-(1+\sigma)} (\sigma > 0)$, does a phase transition at a finite T become possible–but only if $\sigma < 1$; for $\sigma > 1$, we are back to the case of no phase transition, while the borderline case $\sigma = 1$ remains in doubt. For more details, see Griffiths (1972), pp. 89–94.

Peierls (1936) was the first to demonstrate that at sufficiently low temperatures the Ising model in two or three dimensions must exhibit a phase transition. He considered the lattice as made up of two kinds of domains, one consisting of "up" spins and the other of "down" spins, separated by a set of boundaries between the neighboring domains, and argued on energy considerations that in a two- or three-dimensional lattice the long-range order that exists at 0 K would persist at finite temperatures. Again, for details, see Griffiths (1972), pp. 59–66.

11.7. The critical exponents

A basic problem in the theory of phase transitions is to study the behavior of a given system in the neighborhood of its critical point. We know that this behavior is marked by the fact that the various physical quantities pertaining to the system possess singularities at the critical point. It is customary to express these singularities in terms of power laws characterized by a set of *critical exponents* which determine the qualitative nature of the critical behavior of the given system. To begin with, we identify an *order parameter* m, and the corresponding *ordering field* h, such that, in the limit $h \to 0$, m tends to a limiting value m_0, with the property that $m_0 = 0$ for $T \geq T_c$ and $\neq 0$ for $T < T_c$. F [illegible] magnetic system, the natural candidate for m is the parameter $\overline{L} (= \overline{\sigma})$ of Secs 11.5 and 11.6, while h

is identified with the quantity $\mu B/kT_c$; for a gas–liquid system, one may adopt the density differential $(\rho_l - \rho_c)$ or $|\rho_g - \rho_c|$ for m and the pressure differential $(P - P_c)$ for h. The various critical exponents are then defined as follows.

The manner in which $m_0 \to 0$, as $T \to T_c$ from below, defines the exponent β:

$$m_0 \sim (T_c - T)^{\beta} \quad (h \to 0, T \lesssim T_c). \tag{1}$$

The manner in which the low-field susceptibility χ_0 diverges, as $T \to T_c$ from above (or from below), defines the exponent γ (or γ'):

$$\chi_0 \sim \left(\frac{\partial m}{\partial h}\right)_{T, h \to 0} \sim \begin{cases} (T - T_c)^{-\gamma} & \text{for } h \to 0, T \gtrsim T_c & \text{(2a)} \\ (T_c - T)^{-\gamma'} & \text{for } h \to 0, T \lesssim T_c; & \text{(2b)} \end{cases}$$

in the gas–liquid transition, the role of χ_0 is played by the isothermal compressibility, $\kappa_T = \rho^{-1}(\partial\rho/\partial P)_T$, of the system. Next, we define an exponent δ by the relation

$$m\,|_{T=T_c} \sim h^{1/\delta} \quad (T = T_c, h \to 0); \tag{3}$$

in the case of a gas–liquid system, δ is a measure of the "degree of flatness" of the critical isotherm at the critical point, for then

$$|P - P_c|\,|_{T=T_c} \sim |\rho - \rho_c|^{\delta} \quad (T = T_c, P \to P_c). \tag{4}$$

Finally, we define exponents α and α' on the basis of the specific heat, C_V, of the gas–liquid system:

$$C_V \sim \begin{cases} (T - T_c)^{-\alpha} & \text{for } T \gtrsim T_c & \text{(5a)} \\ (T_c - T)^{-\alpha'} & \text{for } T \lesssim T_c. & \text{(5b)} \end{cases}$$

In connection with the foregoing relations, especially eqns (5), we wish to emphasize that in certain cases the exponent in question is rather small in value; it is then more appropriate to write

$$f(t) \sim \frac{|t|^{-\lambda} - 1}{\lambda} \quad (|t| \ll 1). \tag{6}$$

Now, if $\lambda > 0$, the function $f(t)$ would have a power-law divergence at $t = 0$; in case $\lambda \to 0$, the function $f(t)$ would have a logarithmic divergence instead:

$$f(t) \sim \ln(1/|t|) \quad (|t| \ll 1). \tag{7}$$

In either case, the derivative $f'(t) \sim |t|^{-(1+\lambda)}$.

A survey of the results derived in Secs 11.2–11.6 shows that for a gas–liquid system obeying the van der Waals equation of state or for a magnetic system treated in the mean field approximation (it does not matter what order of approximation one is talking about), the various critical exponents are the same:

$$\beta = \tfrac{1}{2}, \quad \gamma = \gamma' = 1, \quad \delta = 3, \quad \alpha = \alpha' = 0. \tag{8}$$

In Table 11.1 we have compiled experimental data on critical exponents pertaining to a variety of systems including the ones mentioned above; for completeness, we

TABLE 11.1. EXPERIMENTAL DATA ON CRITICAL EXPONENTS

Critical exponents	Magnetic systems(a)	Gas–liquid systems(b)	Binary fluid mixtures(c)	Binary alloys(d)	Ferroelectric systems(e)	Superfluid He^{4}(f)	Mean field results
α,α'	0.0–0.2	0.1–0.2	0.05–0.15	– – –	– – –	−0.026	0
β	0.30–0.36	0.32–0.35	0.30–0.34	0.305 ± 0.005	0.33–0.34	– – –	1/2
γ	1.2–1.4	1.2–1.3	1.2–1.4	1.24 ± 0.015	1.0 ± 0.2	inaccessible	1
γ'	1.0–1.2	1.1–1.2	– – –	1.23 ± 0.025	1.23 ± 0.02	inaccessible	1
δ	4.2–4.8	4.6–5.0	4.0–5.0	– – –	– – –	inaccessible	3
ν	0.62–0.68	– – –	– – –	0.65 ± 0.02	0.5–0.8	0.675	1/2
η	0.03–0.15	– – –	– – –	0.03–0.06	– – –	– – –	0

(a) Stierstadt *et al.* (1990).

(b) Voronel (1976); Rowlinson and Swinton (1982).

(c) Rowlinson and Swinton (1982).

(d) J. Als-Nielsen (1976); data pertain to beta-brass only.

(e) Kadanoff *et al.* (1967); Lines and Glass (1977).

(f) Ahlers (1980).

have included [illegible] data on another two exponents, ν and η, which will be defined in Sec. 11.12. [illegible] find that, while the observed values of an exponent, say β, differ very little as one goes from system to system within a given category (or even from category to category), these values are considerably different from the ones following from the mean field approximation. Clearly, we need a theory of phase transitions which is basically different from the mean field theory.

To begin with, some questions arise:

(i) Are these exponents completely independent of one another *or* are they mutually related? In the latter case, how many of them are truly independent?
(ii) On what characteristics of the given system do they depend? This includes the question why, for systems differing so much from one another, they differ so little.
(iii) How can they be evaluated from first principles?

The answer to question (i) is simple: yes, the various exponents do obey certain relations and hence are *not* completely independent. These relations appear in the form of inequalities, dictated by the principles of thermodynamics, which will be explored in Sec. 11.8; in the modern theory of phase transitions, see Secs 11.10–11.12 and Chap. 13, the same relations turn up as equalities, and the number of these (restrictive) relations is such that, for most situations, only *two* of the exponents are truly independent.

As regards question (ii), it turns out that our exponents depend upon a *very small* number of characteristics, or parameters, of the problem, which explains why they differ so little from one system to another in a given category of systems (and also from one category to another, even though systems in those categories are so different from one another). The characteristics that seem to matter are (a) the dimensionality, d, of the space in which the system is embedded, (b) the number of components, n, of the order parameter of the problem, and (c) the range of microscopic interactions in the system.

Insofar as interactions are concerned, all that matters is whether they are *short-ranged* (which includes the special case of nearest-neighbor interactions) or *long-ranged*. In the former case, the values of the critical exponents resulting from nearest-neighbor interactions remain unaltered–regardless of whether further-neighbor interactions are included or not; in the latter case, assuming $J_{ij} \sim |i-j|^{-(d+\sigma)}$ with $\sigma > 0$, the critical exponents depend on σ. Unless a statement is made to the contrary, the microscopic interactions operating in the given system will be assumed to be short-ranged; the critical exponents then depend only on d and n–both of which, for instructional purposes, may be treated as *continuous* variables.

Insofar as d is concerned, we recall the Bethe approximation which highlighted the special role played by the dimensionality of the lattice through, and only through, the coordination number q. We also recall that, while the theoretical value of T_c and the various amplitudes of the problem were influenced by q, the critical exponents were not. In more accurate theories we find that the critical exponents depend more directly on d and only indirectly on q; however, for a given d, they do *not* depend on the structural details of the lattice (which includes the number q).

Insofar as n is concerned, the major difference lies between the Ising model ($n = 1$) with *discrete* symmetry ($\sigma_i = +1$ or -1) and other models ($n \geq 2$) with *continuous* symmetry ($-1 \leq \sigma_{i\alpha} \leq +1$ for $\alpha = 1, \ldots, n$, with $|\sigma_i| = 1$). In the former case, T_c is zero for $d \leq 1$ and nonzero for $d > 1$; in the latter, T_c is zero for $d \leq 2$ and nonzero for $d > 2$.[13] In either case, the critical exponents depend on both d and n, except that for $d > 4$ they become independent of d and n, and assume values identical with the ones given by the mean field theory; the physical reason behind this overwhelming generality is examined in Sec. 11.13. In passing, we note that, for given d and n, the critical exponents do not depend on whether the spins constituting the system are treated classically or quantum-mechanically.

As regards question (iii), the obvious procedure for evaluating the critical exponents is to carry out exact (or almost exact) analysis of the various models–a task to which the whole of Chap. 12 is devoted. An alternative approach is provided by the renormalization group theory, which is discussed in Chap. 13. A modest attempt to evaluate the critical exponents is made in Sec. 11.9, which yields results far below our expectations but teaches us quite a few lessons about the shortcomings of the so-called classical approaches.

11.8. Thermodynamic inequalities

The first rigorous relation linking critical exponents was derived by Rushbrooke (1963) who, on thermodynamic grounds, showed that for any physical system undergoing a phase transition

$$(\alpha' + 2\beta + \gamma') \geq 2. \tag{1}$$

The proof of inequality (1) is straightforward if one adopts a magnetic system as an example. We start with the thermodynamic formula for the difference between the specific heat at constant field C_H and the specific heat at constant magnetization C_M (see Problem 3.40)

$$C_H - C_M = -T\left(\frac{\partial H}{\partial T}\right)_M \left(\frac{\partial M}{\partial T}\right)_H = T\chi^{-1}\left\{\left(\frac{\partial M}{\partial T}\right)_H\right\}^2 \tag{2}$$

Since $C_M \geq 0$, it follows that

$$C_H \geq T\chi^{-1}\left\{\left(\frac{\partial M}{\partial T}\right)_H\right\}^2 \tag{3}$$

Now, letting $H \to 0$ and $T \to T_c$ from below, we get

$$D_1(T_c - T)^{-\alpha'} \geq D_2 T_c (T_c - T)^{\gamma' + 2(\beta - 1)}, \tag{4}$$

where D_1 and D_2 are positive constants; here, use has been made of power laws (11.7.1, 2b and 5b).[14] Inequality (4) may as well be written as

$$(T_c - T)^{2 - (\alpha' + 2\beta + \gamma')} \geq D_2 T_c / D_1. \tag{5}$$

Since $(T_c - T)$ can be made as small as we like, (5) will not hold if $(\alpha' + 2\beta + \gamma') < 2$. The *Rushbrooke inequality* (1) is thus established.

To establish further inequalities, one utilizes the convexity properties of the Helmholtz free energy $A(T, M)$. Since $dA = -S\,dT + H\,dM$,

$$\left(\frac{\partial A}{\partial T}\right)_M = -S, \left(\frac{\partial^2 A}{\partial T^2}\right)_M = -\left(\frac{\partial S}{\partial T}\right)_M = -\frac{C_M}{T} \leq 0 \qquad \text{(6a, b)}$$

and

$$\left(\frac{\partial A}{\partial M}\right)_T = H, \left(\frac{\partial^2 A}{\partial M^2}\right)_T = \left(\frac{\partial H}{\partial M}\right)_T = \frac{1}{\chi} \geq 0. \qquad \text{(7a, b)}$$

It follows that $A(T, M)$ is concave in T and convex in M. We now proceed to establish the *Griffiths inequality* (1965a, b)

$$\alpha' + \beta(\delta + 1) \geq 2. \qquad (8)$$

Consider a magnetic system in zero field and at a temperature $T_1 < T_c$. Then, by (7a), $A(T, M)$ is a function of T only, so we can write

$$A(T_1, M) = A(T_1, 0) \quad (-M_1 \leq M \leq M_1), \qquad (9)$$

where M_1 is the spontaneous magnetization at temperature T_1; see Fig. 11.10. Applying (6a) to (9), we get

$$S(T_1, M) = S(T_1, 0) \quad (-M_1 \leq M \leq M_1). \qquad (10)$$

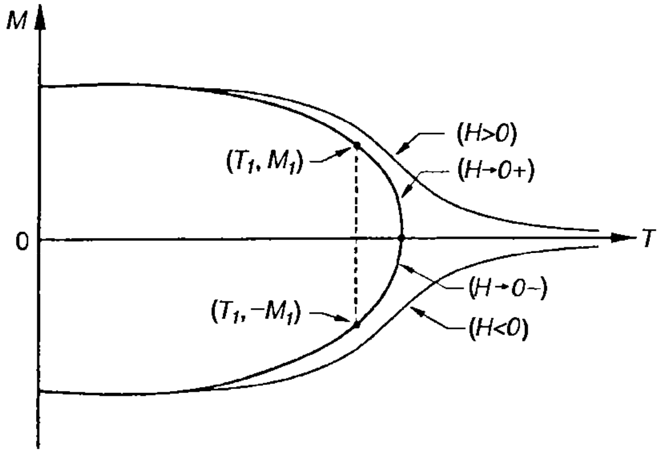


FIG. 11.10. Magnetization, $M(T, H)$, of a magnetic system for $H > 0, H < 0$ and $H \to 0$. Here, M_1 denotes the spontaneous magnetization of the system at a temperature $T_1 < T_c$.

We now define two new functions

$$A^*(T, M) = \{A(T, M) - A_c\} + (T - T_c)S_c \qquad (11)$$

and

$$S^*(T, M) = S(T, M) - S_c, \qquad (12)$$

where $A_c = A(T_c, 0)$ and $S_c = S(T_c, 0)$. It follows th

$$\left(\frac{\partial A^*}{\partial T}\right)_M = -S^*, \quad \left(\frac{\partial^2 A^*}{\partial T^2}\right)_M = -\left(\frac{\partial S^*}{\partial T}\right)_M = -\frac{C_M}{T} \leq 0. \tag{13a, b}$$

Thus, A^* is also concave in T. Geometrically, this means that, for any choice of T_1, the curve $A^*(T)$, with M fixed at M_1, lies below the tangent line at $T = T_1$, i.e.

$$A^*(T, M_1) \leq A^*(T_1, M_1) + \left(\frac{\partial A^*}{\partial T}\right)_{M_1, T=T_1} (T - T_1); \tag{14}$$

see Fig. 11.11. Letting $T = T_c$ in (14), we get

$$A^*(T_c, M_1) \leq A^*(T_1, M_1) - S^*(T_1, M_1)(T_c - T_1) \tag{15}$$

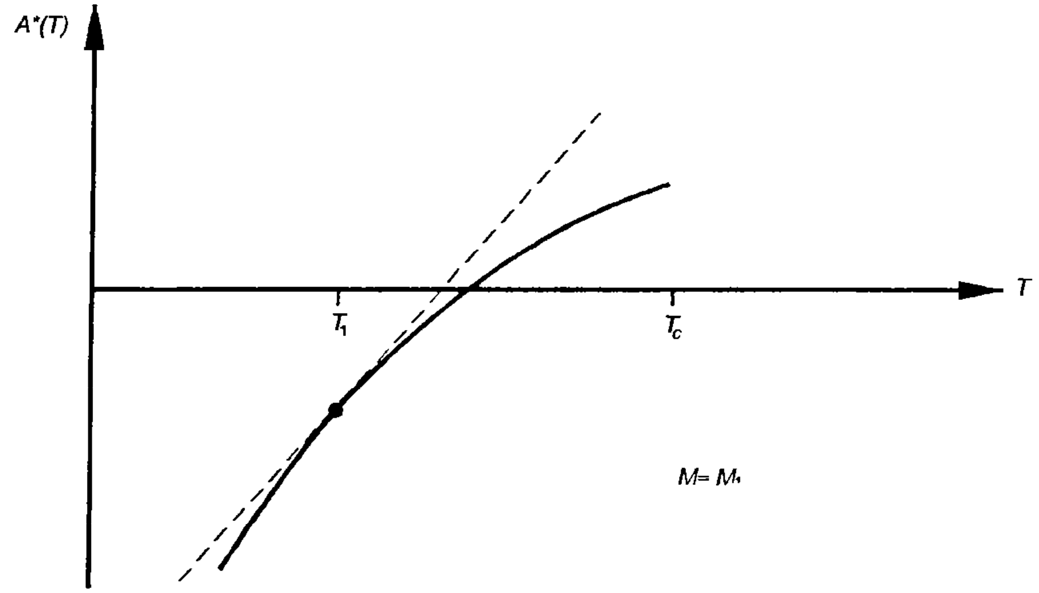


FIG. 11.11. The function $A^*(T, M)$ of a magnetic system, with magnetization M fixed at M_1. The slope of this curve is $S(T_c, 0) - S(T, M_1)$, which is positive for all $T \leq T_c$.

which, in view of eqns (9)–(12), may be written as

$$A^*(T_c, M_1) \leq A^*(T_1, 0) - S^*(T_1, 0)(T_c - T_1). \tag{16}$$

Utilizing, once again, the concavity of the function $A^*(T)$ but this time at $T = T_c$ [with M fixed at zero and the slope $(\partial A^*/\partial T)$ vanishing], we get, see (14),

$$A^*(T, 0) \leq A^*(T_c, 0). \tag{17}$$

Now, letting $T = T_1$ in (17) and noting that $A^*(T_c, 0) = 0$ by definition, we get

$$A^*(T_1, 0) \leq 0. \tag{18}$$

Combining (16) and (18), we finally get

$$A^*(T_c, M_1) \leq -(T_c - T_1)S^*(T_1, 0), \tag{19}$$

valid for all $T_1 < T_c$.

The next step is straightforward. We let $T_1 \to T_c-$, so that $M_1 \to 0$ and along with it

$$A^*(T_c, M_1) = \left[\int_0^{M_1} H\,dM\right]_{T=T_c} \approx DM_1^{\delta+1} \approx D'(T_c - T_1)^{\beta(\delta+1)}, \tag{20}$$

while

$$S^*(T_1, 0) = \int_{T_c}^{T_1} \frac{C(T,0)}{T}\,dT \approx -\frac{D''}{T_c}(T_c - T_1)^{1-\alpha'}. \tag{21}$$

where D, D' and D'' are positive constants; here, use has been made of power laws (11.7.1,3 and 5b). Substituting (20) and (21) into (19), we get

$$(T_c - T_1)^{2-\alpha'-\beta(\delta+1)} \geq D'T_c/D''. \tag{22}$$

Again, since $(T_c - T_1)$ can be made as small as we like, (22) will not hold if $\alpha' + \beta(\delta + 1) < 2$. The Griffiths inequality (8) is thus established. It will be noted that unlike the Rushbrooke inequality, which related critical exponents pertaining only to $T < T_c$, the present inequality relates two such exponents, α' and β, with one, viz. δ, that pertains to the critical isotherm ($T = T_c$).

While inequalities (1) and (8) are thermodynamically exact, Griffiths has derived several others that require certain plausible assumptions on the system in question. We quote two of them here, without proof:

$$\gamma' \geq \beta(\delta - 1) \tag{23}$$

$$\gamma \geq (2 - \alpha)(\delta - 1)/(\delta + 1). \tag{24}$$

For a complete list of such inequalities, see Griffiths (1972), p. 102, where references to original papers are also given.

Before proceeding further, the reader may like to verify that the experimental data on critical exponents, as given in Table 11.1, do indeed conform to the inequalities proved or quoted in this section. It is important in this connection to note that the mean field exponents ($\alpha = \alpha' = 0$, $\beta = 1/2$, $\gamma = \gamma' = 1$ and $\delta = 3$) satisfy all these relations as *equalities*.

11.9. Landau's phenomenological theory

As early as 1937 Landau attempted a unified description of all *second-order* phase transitions–second-order in the sense that the second derivatives of the free energy, namely the specific heat and the magnetic susceptibility (or isothermal compressibility, in the case of fluids), show a divergence while the first derivatives, namely the entropy and the magnetization (or density, in the case of fluids), are continuous at the critical point. He emphasized the importance of an *order parameter* m_0 (which would be zero on the high-temperature side of the transition and nonzero on the low-temperature side) and suggested that the basic features of the critical behavior of a given system may be determined by expanding its free energy in powers of m_0 [for we know that, in the close vicinity of the critical point, $m_0 \ll 1$]. He also argued that in the absence of the *ordering field* ($h = 0$)

the up–down symmetry of the system would require that the proposed expansion contain only *even* powers of m_0. Thus, the zero-field free energy $\psi_0(= A_0/NkT)$ of the system may be written as

$$\psi_0(t, m_0) = q(t) + r(t)m_0^2 + s(t)m_0^4 + \cdots \quad \left(t = \frac{T - T_c}{T_c}, |t| \ll 1\right); \tag{1}$$

at the same time, the coefficients $q(t), r(t), s(t)\ldots$ may be written as

$$q(t) = \sum_{k\geq 0} q_k t^k, \quad r(t) = \sum_{k\geq 0} r_k t^k, \quad s(t) = \sum_{k\geq 0} s_k t^k, . \tag{2}$$

The equilibrium value of the order parameter is then determined by *minimizing* ψ_0 with respect to m_0; retaining terms only up to the order displayed in (1), which for thermodynamic stability requires that $s(t) > 0$, we obtain

$$r(t)m_0 + 2s(t)m_0^3 = 0. \tag{3}$$

The equilibrium value of m_0 is thus either 0 or $\pm\surd[-r(t)/2s(t)]$. The first solution is of lesser interest, though this is the only one we will have for $t > 0$; it is the other solutions that lead to the possibility of spontaneous magnetization in the system. To obtain physically sensible results, see eqns (9)–(11), we must have in eqn. (2): $r_0 = 0$, $r_1 > 0$ and $s_0 > 0$, with the result

$$|m_0| \approx [(r_1/2s_0)|t|]^{1/2} \quad (t \lesssim 0), \tag{4}$$

giving $\beta = 1/2$.

The asymptotic expression for the free energy, viz.

$$\psi_0(t, m_0) \approx q_0 + r_1 t m_0^2 + s_0 m_0^4 \quad (r_1, s_0 > 0), \tag{5}$$

is plotted in Fig. 11.12. We see that, for $t \geq 0$, there is only *one* minimum, which is located at $m_0 = 0$: for $t = 0$, the minimum is rather flat. For $t < 0$, on the other hand, we have *two* minima, located at $m_0 = \pm m_s$, as given by expression (4), with a maximum at $m_0 = 0$. Now, since ψ_0 has to be convex in m_0, so that the susceptibility of the system be non-negative, see eqn. (11.8.7b), we must replace the nonconvex portion of the curve, that lies between the points $m_0 = -m_s$ and $m_0 = +m_s$, by a straight line (along which the susceptibility would be infinite). This replacement is reminiscent of the *Maxwell construction* employed in Secs 11.1 and 11.2.

We now subject the system to an *ordering field h*, assumed positive. If the field is weak, the only change in the expression for the free energy would be the addition of a term $-hm$. Disregarding the appearance of any higher powers of (hm) as well as any modifications of the coefficients already present, we may now write

$$\psi_h(t, m) = -hm + q(t) + r(t)m^2 + s(t)m^4. \tag{6}$$

The equilibrium value of m is then given by[15]

$$-h + 2r(t)m + 4s(t)m^3 = 0. \tag{7}$$

The low-field susceptibility of the system, in units of $N\mu^2/kT$, thus turns out to be

$$\chi = \left(\frac{\partial h}{\partial m}\right)_t^{-1} = \frac{1}{2r(t) + 12s(t)m^2}, \tag{8}$$

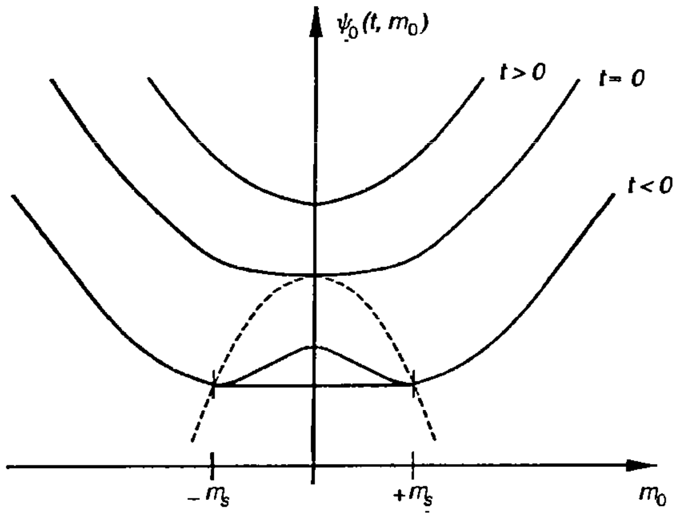


FIG. 11.12. The free energy $\psi_0(t, m_0)$ of the Landau theory, shown as a function of m_0, for three different values of t. The dashed curve depicts spontaneous magnetization $m_s(t)$, while the horizontal line for $t < 0$ provides the Maxwell construction.

valid in the limit $h \to 0$. Now, for $t > 0$, $m \to 0$ and we get

$$\chi \approx 1/2r_1 t \quad (t \gtrsim 0), \tag{9}$$

giving $\gamma = 1$. On the other hand, for $t < 0$, $m \to \surd[(r_1/2s_0)|t|]$, see (4); we then get

$$\chi \approx 1/4r_1|t| \quad (t \lesssim 0), \tag{10}$$

giving $\gamma' = 1$. Finally, if we set $t = 0$ in (7), we obtain the following relation between h and m:

$$h \approx 4s_0 m^3 \quad (h \to 0), \tag{11}$$

giving $\delta = 3$.

We shall now look at the specific heats C_h and C_m. If $t > 0$, then $h \to 0$ implies $m \to 0$, so in this limit there is no difference between C_h and C_m. Equation (1) then gives, in units of Nk,

$$C_h = C_m = -\left(\frac{\partial^2 \psi_0}{\partial t^2}\right)_{m \to 0} = -(2q_2 + 6q_3 t + \cdots) \quad (t \gtrsim 0). \tag{12}$$

For $t < 0$, we have

$$\begin{aligned} C_m &= -\left[(2q_2 + 6q_3 t + \cdots) + (2r_2 + \cdots) m_s^2 + \cdots\right] \\ &= -\left[2q_2 + \{6q_3 - (r_1 r_2/s_0\} t + \ldots\right] \quad (t \lesssim 0). \end{aligned} \tag{13}$$

Next, using eqn. (11.8.2) along with (4) and (10), we have

$$C_h - C_m = \left(\frac{\partial h}{\partial m}\right)_t \left\{\left(\frac{\partial m}{\partial t}\right)_h\right\}^2 \approx \frac{r_1^2}{2s_0} \quad (t \lesssim 0). \tag{14}$$

It follows that, while C_m possesses a cusp-like singularity at $t = 0$, C_h undergoes a jump discontinuity of magnitude

$$(C_h)_{t \to 0-} - (C_h)_{t \to 0+} \approx r_1^2/2s_0. \tag{15}$$

It follows that $\alpha = \alpha' = 0$.

The most striking feature of the Landau theory is that it gives exactly the same critical exponents as the mean field theory of Secs 11.5 and 11.6 (or the van der Waals theory of Sec. 11.2). Actually it goes much further, for it starts with an expression for the free energy of the system containing parameters $q_k, r_k, s_k, \ldots$, which represent the structure of the given system and the interactions operating in it, and goes on to show that, while the amplitudes of the various physical quantities near the critical point do depend on these parameters, the critical exponents do not! This *universality* (of critical exponents) suggests that we are dealing here with a class of systems which, despite their structural differences, display a critical behavior that is qualitatively the same for all members of the class. This leads to the concept of a *universality class* which, if Landau were right, would be a rather large one. The fact of the matter is that the concept of universality is very much overstated in Landau's theory; in reality, there are many different universality classes–each defined by the parameters d and n of Sec. 11.7 and by the range of the microscopic interactions–such that the critical exponents within a class are the same while they vary from one class to another. The way Landau's theory is set up, the parameter n is essentially equal to 1 (because the order parameter m_0 is treated as a scalar), the parameter d plays no role at all (though later on we shall see that the mean field exponents are, in fact, valid for all n if $d > 4$), while the microscopic interactions are implicitly long-ranged.[16]

An objection commonly raised against the Landau theory is that, knowing fully well that the thermodynamic functions of the given system are going to be *singular* at $t = 0$, a *Taylor-type expansion* of the free energy around $m = 0$ is patently a wrong start. While the objection is valid, it is worth noting how a regular function, (1) or (6), leads to an equation of state, (3) or (7), which yields different results for $t \to 0-$ from the ones for $t \to 0+$, the same being true of whether $h \to 0+$ or $0-$. The trick lies in the fact that we are not using eqn. (1) or (6) as such for *all* t; for $t < 0$, we use instead a modified form, as "corrected" by the Maxwell construction (see Fig. 11.12). The spirit of the singularity is thereby captured, though the nature of the singularity, being closely tied with the nature of the original expansion, could not be any different from the mean-field type. The question now arises: how can the Landau theory be improved so that it may provide a more satisfactory picture of the critical phenomena? Pending exact analyses, one wonders if some generalization of the Landau approach, admitting more than one universality class, would provide a better picture than the one we have at present. It turns out that the scaling approach, initiated by Widom (1965), by Domb and Hunter (1965) and by Patashinskii and Pokrovskii (1966), was the next natural step in the right direction.

11.10. Scaling hypothesis for thermodynamic functions

The scaling approach, which took the subject of phase transitions far beyond the mean field theory, emerged independently from three different sources–from Widom (1965) who searched for a generalization of the van der Waals equation of state that could accommodate nonclassical exponents, from Domb and Hunter (1965) who analyzed the behavior of the series expansions of higher derivatives of the free energy with respect to the magnetic field at the critical point of a magnetic system, and from Patashinskii and Pokrovskii (1966) who studied the behavior of multi-correlation functions for the spins constituting a system. All three were led to the same form of the thermodynamic equation of state. Subsequently, Kadanoff (1966a) suggested a *scaling hypothesis* from which not only could this equation of state be derived but one could also obtain a number of relations among the critical exponents, which turned out to be equalities consistent with the findings of Sec. 11.8. This approach also made it clear why one needed only *two* independent numbers to describe the nature of the singularity in question; all other relevant numbers followed as consequences.

To set the stage for this development, we go back to the equation of state following from the Landau theory, viz. eqn. (11.9.7), and write it in the asymptotic form

$$h(t, m) \approx 2r_1 tm + 4s_0 m^3. \tag{1}$$

In view of the relationship (11.9.4), we rewrite (1) in the form

$$h(t, m) \approx \frac{r_1^{3/2}}{s_0^{1/2}} |t|^{3/2} \left[2\,\mathrm{sgn}\,(t) \left(\frac{s_0^{1/2}}{r_1^{1/2}} \frac{m}{|t|^{1/2}} \right) + 4 \left(\frac{s_0^{1/2}}{r_1^{1/2}} \frac{m}{|t|^{1/2}} \right)^3 \right] \tag{2}$$

It follows that

$$m(t, h) \approx \frac{r_1^{1/2}}{s_0^{1/2}} |t|^{1/2} \times \text{a function of} \left(\frac{s_0^{1/2}}{r_1^{3/2}} \frac{h}{|t|^{3/2}} \right) \tag{3}$$

and, within the context of the Landau theory, the function appearing here is *universal* for all systems conforming to this theory. In the same spirit, the relevant part of the free energy $\psi_h(t, m)$–the part that determines the nature of the singularity–may be written in the form

$$\psi_h^{(s)}(t, m) \approx -hm + r_1 tm^2 + s_0 m^4 \tag{4}$$

$$= \frac{r_1^2}{s_0} t^2 \left[-\left(\frac{s_0^{1/2}}{r_1^{3/2}} \frac{h}{|t|^{3/2}} \right) \left(\frac{s_0^{1/2}}{r_1^{1/2}} \frac{m}{|t|^{1/2}} \right) + \mathrm{sgn}\,(t) \left(\frac{s_0^{1/2}}{r_1^{1/2}} \frac{m}{|t|^{1/2}} \right)^2 + \left(\frac{s_0^{1/2}}{r_1^{1/2}} \frac{m}{|t|^{1/2}} \right)^4 \right]. \tag{5}$$

Substituting (3) into (5), one gets

$$\psi^{(s)}(t, h) \approx \frac{r_1^2}{s_0} t^2 \times \text{a function of} \left(\frac{s_0^{1/2}}{r_1^{3/2}} \frac{h}{|t|^{3/2}} \right) \tag{6}$$

where, again, the function appearing here is *universal*. As a check, we see that differentiating (6) with respect to h we readily obtain (3).

The most notable feature of the equation of state, as expressed in (3), is that, instead of being the usual relationship among *three* variables m, h and t, it is now a relationship among only *two* variables, viz. $m/|t|^{1/2}$ and $h/|t|^{3/2}$. Thus, by scaling m with $|t|^{1/2}$ and h with $|t|^{3/2}$, we have effectively reduced the total number of variables by one. Similarly, we have replaced eqn. (4) by (6), which expresses the singular part of the free energy ψ scaled with t^2 as a function of the single variable h scaled with $|t|^{3/2}$. This reduction in the total number of effective variables may be regarded as the first important achievement of the scaling approach.

The next step consists in generalizing (6), to write

$$\psi^{(s)}(t,h) \approx F|t|^{2-\alpha} f(Gh/|t|^{\Delta}) \tag{7}$$

where α and Δ are *universal* numbers common to all systems in the given (universality) class, $f(x)$ is a *universal* function which is expected to have two different branches, f_+ for $t > 0$ and f_- for $t < 0$, while F and G (like r_1 and s_0) are *non-universal* parameters characteristic of the particular system under consideration. We expect α and Δ to determine all the critical exponents of the problem, while the amplitudes appearing in the various power laws will be determined by F, G and the limiting values of the function $f(x)$ and its derivatives (as x tends to zero). Equation (7) constitutes the so-called *scaling hypothesis*, whose status will become much more respectable when it acquires legitimacy from the renormalization group theory; see Secs 13.1 and 13.3.

First of all it should be noted that the exponent of $|t|$, outside the function $f(x)$ in eqn. (7), has been chosen to be $(2-\alpha)$, rather than 2 of the corresponding mean field expression (6), so as to ensure that the specific heat singularity is correctly reproduced. Secondly, the fact that one must not encounter any singularities as one crosses the critical isotherm ($t = 0$) at *nonzero* values of h or m requires that the exponents on the high-temperature side of the critical point be the same as on the low-temperature side, i.e.

$$\alpha' = \alpha \quad \text{and} \quad \gamma' = \gamma. \tag{8}$$

From eqn. (7) it readily follows that

$$m(t,h) = -\left(\frac{\partial \psi^{(s)}}{\partial h}\right)_t \approx -FG|t|^{2-\alpha-\Delta} f'(Gh/|t|^{\Delta}) \tag{9}$$

and

$$\chi(t,h) = -\left(\frac{\partial^2 \psi^{(s)}}{\partial h^2}\right)_t \approx -FG^2|t|^{2-\alpha-2\Delta} f''(Gh/|t|^{\Delta}). \tag{10}$$

Letting $h \to 0$, we obtain for the *spontaneous magnetization*

$$m(t,0) \approx B|t|^{\beta} \quad (t \lesssim 0), \tag{11}$$

where

$$B = -FGf'_-(0), \quad \beta = 2-\alpha-\Delta, \tag{12a, b}$$

and for the *low-field susceptibility*

$$\chi(t,0) \approx |t|^{-\gamma} \begin{cases} C_+ & \text{for} \quad t \gtrsim 0 \qquad (13a) \\ C_- & \text{for} \quad t \lesssim 0, \qquad (13b) \end{cases}$$

where

$$C_{\pm} = -FG^2 f''_{\pm}(0), \quad \gamma = \alpha + 2\Delta - 2. \tag{14a, b}$$

Combining (12b) and (14b), we get

$$\Delta = \beta + \gamma = 2 - \alpha - \beta, \tag{15}$$

so that

$$\alpha + 2\beta + \gamma = 2. \tag{16}$$

To recover δ, we write the function $f'(x)$ of eqn. (9) as $x^{\beta/\Delta}g(x)$, so that

$$m(t, h) \approx -FG^{(1+\beta/\Delta)}h^{\beta/\Delta}g(Gh/|t|^{\Delta}). \tag{17}$$

Inverting (17), we can write

$$|t| \approx G^{1/\Delta}h^{1/\Delta} \times \text{a function of } (FG^{(1+\beta/\Delta)}h^{\beta/\Delta}/m). \tag{18}$$

It follows that, along the critical isotherm ($t = 0$), the argument of the function appearing in (18) would have a *universal* value (which makes the function vanish), with the result that

$$m \sim FG^{(1+\beta/\Delta)}h^{\beta/\Delta} \quad (t = 0). \tag{19}$$

Comparing (19) with (11.7.3), we infer that

$$\delta = \Delta/\beta. \tag{20}$$

Combining (20) with the previous relations, viz. (12b) and (15), we get

$$\alpha + \beta(\delta + 1) = 2 \tag{21}$$

and

$$\gamma = \beta(\delta - 1). \tag{22}$$

Finally, combining (21) and (22), we have

$$\gamma = (2 - \alpha)(\delta - 1)/(\delta + 1). \tag{23}$$

For completeness, we write down for the specific heat C_h at $h = 0$

$$C_h^{(s)}(t, 0) = -\frac{\partial^2\psi^{(s)}}{\partial t^2} \approx -(2-\alpha)(1-\alpha)F|t|^{-\alpha}\begin{cases} f_+(0) & \text{for} \quad t \gtrsim 0 \quad \text{(24a)} \\ f_-(0) & \text{for} \quad t \lesssim 0. \quad \text{(24b)} \end{cases}$$

We thus see that the scaling hypothesis (7) leads to a number of relations among the critical exponents of the system, emphasizing the fact that only *two* of them are truly independent. Comparing these relations with the corresponding ones appearing in Sec. 11.8, viz. (16), (21), (22) and (23) with (11.8.1, 8, 23 and 24), we feel satisfied that they are mutually consistent, though the present ones are far more restrictive than the ones in Sec. 11.8. Besides exponent relations, we also obtain here relations among the various amplitudes of the problem; though individually these amplitudes are non-universal, certain combinations thereof turn out to be universal. For instance, the combination $(FC_{\pm}/B^2)$, which consists of coefficients appearing in eqns (7), (11) and (13), is universal; the same is true of the ratio C_+/C_-. For further information on this question, see the original papers by Watson (1969) and a recent review by Privman, Hohenberg and Aharony (1991).

We now pose the question: why do "universality classes" exist in the first place? In other words, what is the reason that a large variety of systems differing widely in their structure should belong to a single universality class and hence have common critical exponents and common scaling functions? The answer lies in the "correlations among the microscopic constituents of the system which, as $T \to T_c$, become large enough to prevail over *macroscopic* distances in the system and in turn make structural details at the *local* level irrelevant". We now turn our attention to this important aspect of the problem.

11.11. The role of correlations and fluctuations

Much can be learned about criticality by scattering radiation–light, x-rays, neutrons, etc.–off the system of interest. In a standard scattering experiment, a well-collimated beam of light, or other radiation, with known wavelength λ is directed at the sample and one measures the intensity, $I(\theta)$, of the light scattered at an angle θ away from the "forward" direction of the beam. The radiation undergoes a shift in wave vector, **k**, which is related to the parameters θ and λ by

$$|k| = \frac{4\pi}{\lambda} \sin \frac{1}{2}\theta. \tag{1}$$

Now, the scattered intensity $I(\theta)$ is determined by the *fluctuations* in the medium. If the medium were perfectly uniform (i.e. spatially homogenous), there would be no scattering at all! If one has in mind light scattering from a fluid, then the relevant fluctuations correspond to regions of different refractive index and, hence, of different particle density $n(r)$. For neutron scattering from a magnet, fluctuations in the spin or magnetization density are the relevant quantities, and so on. We need to study here the *normalized* scattering intensity $I(\theta; T, H)/I^{\text{ideal}}(\theta)$, where $I(\theta; T, H)$ is the actual scattering intensity observed at angle θ, which will normally depend on such factors as temperature, magnetic field, etc., while $I^{\text{ideal}}(\theta)$ is the scattering that would take place if the individual particles (or spins) that cause the scattering could somehow be taken far apart so that they no longer interact and hence are quite uncorrelated with one another. Now, this normalized scattering intensity turns out to be essentially proportional to the quantity

$$\tilde{g}(k) = \int g(r) e^{ik \cdot r}\, dr, \tag{2}$$

which represents the Fourier transform of the appropriate real-space correlation function $g(r)$, which will be defined shortly.

As the critical point of the system (say, a fluid) is approached one observes an enormously enhanced level of scattering, especially at low angles which corresponds, via eqns (1) and (2), to long wavelength density fluctuations in the fluid. In the critical region the scattering is so large that it can be visible to the unaided eye, particularly through the phenomenon of *critical opalescence*. This behavior is, by no means, limited to fluids. Thus if, for example, one scatters neutrons from iron in the vicinity of the Curie point one likewise sees a dramatic growth in the low-angle neutron scattering intensity, as sketched in Fig. 11.13. One sees that for small angle scattering there is a pronounced peak in $I(\theta; T)$ as a function of temperature, and this peak approaches closer and closer to T_c as the angle is decreased. Of course, one could never actually observe zero-angle scattering directly, since that would mean picking up the oncoming beam itself, but one *can* extrapolate to zero angle. When this is done one finds that the zero-angle scattering $I(0; T)$, actually *diverges* at T_c. This is the most dramatic manifestation of

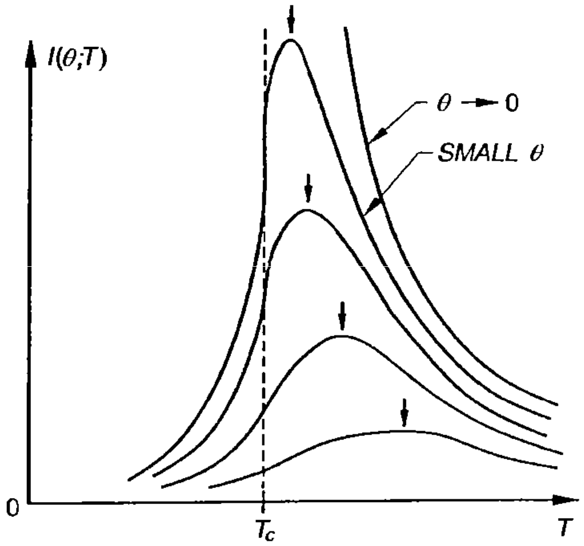


FIG. 11.13. Schematic plot of the elastic scattering intensity of neutrons scattered at an angle θ from a magnetic system, such as iron, in the vicinity of the critical point T_c. The small arrows mark the smoothly rounded maxima (at fixed θ) which occur at a temperature $T_{\max}(\theta)$ that approaches T_c as $\theta \to 0$.

the phenomenon of critical opalescence and is quite general, in that it is observed whenever appropriate scattering experiments can be performed. Empirically, one may write for small-angle scattering

$$I_{\max}(\theta) \sim k^{-\lambda_1}, \quad \{T_{max}(\theta) - T_c\} \sim k^{\lambda_2}, \tag{3}$$

so that

$$I_{\max}(\theta)\{T_{\max}(\theta) - T_c\}^{\lambda_1/\lambda_2} = \text{const.} \tag{4}$$

Here, λ_1 and λ_2 are positive exponents (which, as will be seen later, are determined by the universality class to which the system belongs), while k, for a given θ, is determined by eqn. (1); note that, for small θ, k is essentially proportional to θ.

The first real insight into the problem of critical scattering in fluids was provided by Ornstein and Zernike (1914, 1916) who emphasized the difference between the direct influence of the microscopic interactions among the atoms of the fluid, which are necessarily *short-ranged*, and the indirect (but more crucial) influence of the density–density correlations which become *long-ranged* as the critical temperature is approached; it is the latter that are truly responsible for the propagation of long-range order in the system and for practically everything else that goes with it. Unfortunately, the original work of Ornstein and Zernike makes difficult reading; moreover, it is based on the classical theory of van der Waals. Nevertheless, the subject has been neatly clarified in the review articles by Fisher (1964, 1983) and Domb (1985), to which the reader may turn for further details. Here we shall keep to the language of the magnetic systems and work out the most essential parts of the theory in somewhat general terms.

We define the spin–spin *correlation function* $g(i, j)$, for the pair of spins at sites i and j, by the standard definition

$$g(i, j) = \overline{\sigma_i \sigma_j} - \overline{\sigma}_i \overline{\sigma}_j. \tag{5}$$

For $i = j$, expression (5) denotes the "mean-square fluctuation in the value of the variable σ at site i"; on the other hand, as the separation between the sites i and j increases indefinitely, the spins σ_i and σ_j get uncorrelated, so that $\overline{\sigma_i\sigma_j} \to \bar{\sigma}_i\bar{\sigma}_j$ and the function $g(i, j) \to 0$. In view of the fact that expression (5) can be written as

$$g(i, j) = \overline{(\sigma_i - \bar{\sigma}_i)(\sigma_j - \bar{\sigma}_j)}, \tag{6}$$

the function $g(i, j)$ may also be looked upon as a measure of the "correlation among the fluctuations in the order parameter of the system at sites i and j". This makes sense because σ_i may, quite appropriately, be regarded as the locally fluctuating order parameter linked to site i, just as $\bar{\sigma}$ is the order parameter for the whole system. We shall now establish connections between the function $g(i, j)$ and certain thermodynamic properties of the system.

We start with the partition function of the system, see eqn. (11.3.9),

$$Q_N(H, T) = \sum_{\{\sigma_i\}} \exp\left[\beta J \sum_{\text{n.n.}} \sigma_i\sigma_j + \beta\mu H \sum_i \sigma_i\right], \tag{7}$$

where the various symbols have their usual meanings. It follows that

$$\frac{\partial}{\partial H}(\ln Q_N) = \beta\mu\overline{\left(\sum_i \sigma_i\right)} = \beta\overline{M}, \tag{8}$$

where $M(= \mu \sum_i \sigma_i)$ denotes the net magnetization of the system. Next, since

$$\frac{\partial^2}{\partial H^2}(\ln Q_N) = \frac{\partial}{\partial H}\left(\frac{1}{Q_N}\frac{\partial Q_N}{\partial H}\right) = \frac{1}{Q_N}\frac{\partial^2 Q_N}{\partial H^2} - \frac{1}{Q_N^2}\left(\frac{\partial Q_N}{\partial H}\right)^2 = \beta^2(\overline{M^2} - \overline{M}^2), \tag{9}$$

we obtain for the magnetic susceptibility of the system

$$\chi \equiv \frac{\partial \overline{M}}{\partial H} = \beta(\overline{M^2} - \overline{M}^2) \tag{10a}$$

$$= \beta\mu^2\left\{\overline{\left(\sum_i \sigma_i\right)^2} - \overline{\left(\sum_i \sigma_i\right)}^2\right\} = \beta\mu^2 \sum_i \sum_j g(i, j). \tag{10b}$$

Equation (10a) is generally referred to as a *fluctuation–dissipation* relation; it may be compared with the corresponding relation for fluids, namely (4.5.7), which connects isothermal compressibility κ_T with density fluctuations in the system. Equation (10b), on the other hand, relates χ to a summation of the correlation function $g(i, j)$ over all i and j; assuming homogeneity, this may be written as

$$\chi = N\beta\mu^2 \sum_{\boldsymbol{r}} g(\boldsymbol{r}) \quad (\boldsymbol{r} = \boldsymbol{r}_j - \boldsymbol{r}_i). \tag{11}$$

Treating $\boldsymbol{r}$ as a continuous variable, eqn. (11) may be written as

$$\chi = \frac{N\beta\mu^2}{a^d}\int g(\boldsymbol{r})\, d\boldsymbol{r}, \tag{12}$$

where a is a microscopic length, such as the lattice constant, so defined that $Na^d = V$, the volume of the system; for a similar result appropriate to fluids, see eqn. (14.2.11). Finally, introducing the Fourier transform of the function $g(\boldsymbol{r})$, through eqn. (2), we observe that

$$\chi = \frac{N\beta\mu^2}{a^d}\tilde{g}(0). \tag{13}$$

Our next task consists in determining the mathematical form of the functions $g(\boldsymbol{r})$ and $\tilde{g}(\boldsymbol{k})$. Pending exact calculations, let us see what the mean field theory has to offer in this regard.

Following Kadanoff (1976b), we consider a magnetic system subject to an external field H which is *non-uniform*, i.e. $H = \{H_i\}$, where H_i denotes the field at site i. Using mean field approximation, the thermal average of the variable σ_i is given by, see eqn. (11.5.10),

$$\bar{\sigma}_i = \tanh(\beta\mu H_{\text{eff}}), \tag{14}$$

where

$$H_{\text{eff}} = H_i + (J/\mu)\sum_{\text{n.n.}}\bar{\sigma}_j; \tag{15}$$

note that, for obvious reasons, the product $(q\bar{\sigma})$ of eqn. (11.5.10) has been replaced by a sum of $\bar{\sigma}_j$ over all the nearest neighbors of spin i. If $\bar{\sigma}$ varies slowly in space, which means that the applied field is not too non-uniform, then (15) may be approximated as

$$H_{\text{eff}} \simeq H_i + (qJ/\mu)\bar{\sigma}_i + (cJa^2/\mu)\nabla^2\bar{\sigma}_i, \tag{16}$$

where c is a number of order unity whose actual value depends upon the structure of the lattice, while a is an effective lattice constant; note that the term involving $\nabla\bar{\sigma}_i$ cancels on summation over the q nearest neighbors which are supposed to be positioned in some symmetrical fashion around the site i. At the same time, the function $\tanh x$, for small x, may be approximated by $x - x^3/3$. Retaining only essential terms, we get from (14)

$$\beta\mu H_i = (1 - q\beta J)\bar{\sigma}_i + \tfrac{1}{3}(q\beta J)^3\bar{\sigma}_i^3 - c\beta Ja^2\nabla^2\bar{\sigma}_i. \tag{17}$$

Now, the condition for criticality is: $[\{H_i\} = 0,\ q\beta_c J = 1]$; see eqn. (11.5.13). So, near criticality, we may introduce our familiar variables

$$h_i = \beta\mu H_i, \quad t = (T - T_c)/T_c \simeq (\beta_c - \beta)/\beta_c; \tag{18}$$

eqn. (17) then reduces to

$$\left(t + \tfrac{1}{3}\bar{\sigma}_i^2 - c'a^2\nabla^2\right)\bar{\sigma}_i = h_i, \tag{19}$$

where c' is another number of order unity. Equation (19) generalizes eqn. (11.10.1) of Landau's theory by taking into account the non-uniformity of $\bar{\sigma}$.

Differentiating (19) with respect to h_j, we get

$$(t + \bar{\sigma}_i^2 - c'a^2\nabla^2)\frac{\partial\bar{\sigma}_i}{\partial h_j} = \delta_{i,j}. \tag{20}$$

The "response function" $\partial\bar{\sigma}_i/\partial h_j$ is identical with the correlation function $g(i, j)$;[17] eqn. (20) may, therefore, be written as

$$(t + \bar{\sigma}_i^2 - c'a^2\nabla^2)g(i, j) = \delta_{i,j}. \tag{21}$$

For $t > 0$ and $\{h_i\} \to 0, \bar{\sigma}_i \to 0$; eqn. (21) then becomes

$$(t - c'a^2\nabla^2)g(i, j) = \delta_{i,j}. \tag{22}$$

Assuming homogeneity, so that $g(i, j) = g(\boldsymbol{r})$ where $\boldsymbol{r} = \boldsymbol{r}_j - \boldsymbol{r}_i$, and introducing Fourier transforms, eqn. (22) gives

$$(t + c'a^2k^2)\tilde{g}(\boldsymbol{k}) = \text{const}; \tag{23}$$

it follows that $\tilde{g}(\boldsymbol{k})$ is a function of the magnitude k only (which is not surprising in view of the assumed symmetry of the lattice). Thus

$$\tilde{g}(k) \sim \frac{1}{t + c'a^2k^2}, \tag{24}$$

which is the famous Ornstein–Zernike result derived originally for fluids. Now, taking the inverse Fourier transform of $\tilde{g}(k)$, we obtain (disregarding numerical factors that are not so essential for the present argument)

$$g(r) \sim \int \frac{e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}}{t + c'a^2k^2} \, d^d(ka) \tag{25a}$$

$$\sim \int_0^\infty \frac{a^d}{t + c'a^2k^2} \left(\frac{1}{kr}\right)^{(d-2)/2} J_{(d-2)/2}(kr)\, k^{d-1}\, dk; \tag{25b}$$

see eqns (8) and (11) of Appendix C. The integral in (25b) is tabulated; see Gradshteyn and Ryzhik (1965), p. 686. We get

$$g(r) \sim \left(\frac{a^2}{\xi r}\right)^{(d-2)/2} K_{(d-2)/2}\left(\frac{r}{\xi}\right) \quad \{\xi = a(c'/t)^{1/2}\}, \tag{26}$$

$K_\mu(x)$ being a modified Bessel function. For $x \gg 1$, $K_\mu(x) \sim x^{-1/2}e^{-x}$; eqn. (26) then gives

$$g(r) \sim \frac{a^{d-2}}{\xi^{(d-3)/2} r^{(d-1)/2}} e^{-r/\xi} \quad (r \gg \xi). \tag{27}$$

On the other hand, for $x \ll 1$. $K_\mu(x)$, for $\mu > 0$, $\sim x^{-\mu}$; eqn. (26) then gives

$$g(r) \sim \frac{a^{d-2}}{r^{d-2}} \quad (r \ll \xi; d > 2). \tag{28}$$

In the special case $d = 2$, we obtain instead

$$g(r) \sim \ln(\xi/r) \quad (r \ll \xi; d = 2). \tag{29}$$

It is worth noting that eqn. (26) simplifies considerably when $d = 3$. Since $K_{1/2}(x) = (\pi/2x)^{1/2}e^{-x}$ for all x,

$$g(r)\,|_{d=3} \sim \frac{a}{r} e^{-r/\xi} \tag{30}$$

for all r. Equation (30) is another famous result of Ornstein and Zernike.

Clearly, the quantity ξ appearing here is a measure of the "distances over which the spin–spin (or density–density) correlations in the system extend"–hence the name *correlation length*. So long as T is significantly above T_c, $\xi = O(a)$; see (26). However, as T approaches T_c, ξ increases indefinitely–ultimately diverging at $T = T_c$. The resulting singularity is also of the power-law type:

$$\xi \sim at^{-1/2} \quad (t \gtrsim 0). \tag{31}$$

The divergence of ξ, at $T = T_c$, is perhaps the most important clue we have for our general understanding of the critical phenomena. As $\xi \to \infty$, correlations extend *all over* the system, paving the way for the propagation of a long-range order (even though the microscopic interactions, which are at the root of the phenomenon, are themselves short-ranged). Moreover, since correlations extend over *macroscopic* distances in the system, any structural details that differentiate one system from another at the *microscopic* level lose significance, leading thereby to some sort of a universal behavior!

Going back to eqns (13) and (24), we see that the singularity in χ is indeed of the type expected in a mean field theory, viz.

$$\chi \sim \tilde{g}(0) \sim t^{-1} \quad (t \gtrsim 0). \tag{32}$$

In view of the foregoing results, one may write

$$\frac{1}{\tilde{g}(k)} \sim \frac{1}{\chi}\left(1 + \xi^2 k^2\right). \tag{33}$$

In a so-called Ornstein–Zernike analysis, one plots $1/\tilde{g}(k)$ [or $1/I(k)$, where $I(k)$ is the intensity of the light scattered at an angle θ, see eqn. (1)] in the critical region versus k^2. The data for small k ($ka \lesssim 0.1$), for which the above treatment holds, fall close to a straight line whose intercept with the ($k^2 = 0$)-axis determines $\chi(t)$. As $t \to 0$, this intercept tends to zero but the successive isotherms remain more or less parallel to one another; the reduced slope evidently serves to determine $\xi(t)$. For $t \simeq 0$, these plots show a slight downward curvature, indicating departure from the k^2-law to one in which the power of k is somewhat less than 2. Finally, as regards the plot $I(\theta; T)$ of Fig. 11.13, the maximum in the curve, according to eqn. (24), should lie at $t = 0$ for all θ and the height of the maximum should be $\sim k^{-2}$, i.e. essentially $\sim \theta^{-2}$; thus, according to the mean field expression for $\tilde{g}(k)$, the exponent λ_1 in eqn. (3) should be 2 while λ_2 should be 0.

11.12. The critical exponents ν and η

According to the mean field theory, the divergence of ξ at $T = T_c$ is governed by the power law (11.11.31), with a critical exponent $\frac{1}{2}$. We anticipate that the experimental data on actual systems may not conform to this law. We therefore introduce a new critical exponent, ν, such that

$$\xi \sim t^{-\nu} \quad (h \to 0, t \gtrsim 0). \tag{1}$$

In the spirit of the scaling hypothesis, see Sec. 11.10, the corresponding exponent ν' appropriate to $t \lesssim 0$ would be the same as ν.[18] Table 11.1, of Sec. 11.7, shows experimental results for ν obtained from a variety of systems; we see that the observed values of ν, while varying very little from system to system, differ considerably from the mean field value.

As regards the correlation function, the situation for $t \gtrsim 0$ is described very well by a law of the type (11.11.27), viz.

$$g(r) \sim e^{-r/\xi(t)} \times \text{some power of } r \quad (t \gtrsim 0), \tag{2}$$

where $\xi(t)$ is given by (1). The variation of $g(r)$ with r in this regime is governed primarily by the exponential, so $g(r)$ falls rapidly as r exceeds ξ [which is typically $O(a)$–may be a few times the lattice constant]. As $t \to 0$ and hence $\xi \to \infty$, the behavior of $g(r)$ would be expected to be like eqn. (11.11.28 or 29). A problem now arises: we do have an *exact* expression for $g(r)$ at $T = T_c$ for a two-dimensional Ising model (see Sec. 12.3), according to which

$$g(r) \sim r^{-1/4} \quad (d = 2, n = 1, t = 0), \tag{3}$$

which is quite different from the mean field expression (11.11.29). We therefore generalize our classical result to

$$g(r) \sim r^{-(d-2+\eta)} \quad (t = 0), \tag{4}$$

which introduces another critical exponent, η. Clearly, η for the two-dimensional Ising model is $\frac{1}{4}$, which can even be confirmed by experiments on (effectively) two-dimensional systems. Table 11.1 shows experimental values of η for some systems in three dimensions; typically, η turns out to be a small number, which makes it rather difficult to measure reliably.

We shall now derive some scaling relations involving the exponents ν and η. First of all let us write down the correlation function $g(r; t, h)$ and its Fourier transform $\tilde{g}(k; t, h)$ in a scaled form. For this, we note that, while h scales with t^{Δ}, the only natural variable with which r will scale is ξ; accordingly, r will scale with $t^{-\nu}$. We may, therefore, write

$$g(r; t, h) \approx \frac{G(rt^{\nu}, h/t^{\Delta})}{r^{d-2+\eta}}, \quad \tilde{g}(k; t, h) \approx \frac{\tilde{G}(k/t^{\nu}, h/t^{\Delta})}{k^{2-\eta}}, \tag{5a, b}$$

where the functions $G(x, y)$ and $\tilde{G}(z, y)$, like the exponents Δ, ν and η, are *universal* for a given (universality) class; in expressions (5), for simplicity, we have suppressed non-universal parameters which vary from system to system within a class. In the absence of the field ($h = 0$), expressions (5) reduce to

$$g(r; t, 0) \approx \frac{G_0(rt^{\nu})}{r^{d-2+\eta}}, \quad \tilde{g}(k; t, 0) \approx \frac{\tilde{G}_0(k/t^{\nu})}{k^{2-\eta}}, \tag{6a, b}$$

where $G_0(x)$ and $\tilde{G}_0(z)$ are again universal. At the critical point ($h = 0, t = 0$), we have simply

$$g_c(r) \sim \frac{1}{r^{d-2+\eta}}, \quad \tilde{g}_c(k) \sim \frac{1}{k^{2-\eta}}. \tag{7a, b}$$

We now recall eqn. (11.11.12), which relates χ to an integral of $g(\boldsymbol{r})$ over $d\boldsymbol{r}$, and substitute expression (6a) into it. We get, ignoring non-universal parameters as well as numerical factors,

$$\chi \sim \int \frac{G_0(rt^{\nu})}{r^{d-2+\eta}} r^{d-1}\, dr. \tag{8}$$

By a change of variables, this gives

$$\chi \sim t^{-(2-\eta)\nu}. \tag{9}$$

Invoking the standard behavior of χ, we obtain

$$\gamma = (2-\eta)\nu. \tag{10}$$

Note that the same scaling relation can also be obtained by appealing to eqns (11.11.13) and (6b); the argument is that, in the limit $k \to 0$, the function $\tilde{G}_0(z)$ must be $\sim z^{2-\eta}$ (so that k is eliminated), leaving behind a result identical with (9). In passing, we note that in the critical region

$$\chi \sim \xi^{2-\eta}. \tag{11}$$

Relation (10) is consistent with the *Fisher inequality* (1969)

$$\gamma \le (2-\eta)\nu,$$

and is obviously satisfied by the mean field exponents ($\gamma = 1, \nu = \frac{1}{2}, \eta = 0$); it also checks well with the experimental data given in Table 11.1. In fact, this relation provides a much better method of extracting the elusive exponent η, from a knowledge of γ and ν, than determining it directly from experiment. Incidentally, the presence of η explains the slight downward curvature of the Ornstein–Zernike plot, $1/\tilde{g}(k)$ vs k^2, as $k \to 0$, for the appropriate expression for $1/\tilde{g}(k)$ now is

$$\frac{1}{\tilde{g}(k)} \sim \frac{1}{\chi}(1 + \xi^{2-\eta}k^{2-\eta}), \tag{12}$$

rather than (11.11.33).

We shall now derive another exponent relation involving ν, but first notice that all exponent relations derived so far have no *explicit* dependence on the dimensionality d of the system (though the actual values of the exponents do depend on d). There is, however, one important relationship which *does* involve d explicitly. For this, let us visualize what happens inside the system (say, a magnetic one) as $t \to 0$ from above. At some stage the correlation length ξ becomes significantly larger than the atomic spacings, with the result that magnetic domains, of alternating magnetization, begin to appear. The closer we get to the critical point, the larger the size of these domains; one may, in fact, say that the volume Ω of any such domain is $\sim \xi^d$. Now, the singular part of the free energy density of the system–or, for that matter, of any one of these domains–is given by, see eqn. (11.10.7) with $h = 0$,

$$f^{(s)}(t) \sim t^{2-\alpha}, \tag{13}$$

which vanishes as $t \to 0$. At the same time, the domain volume Ω diverges. It seems natural to expect that $f^{(s)}$, being a density, would vanish as $1/\Omega$, i.e.,

$$f^{(s)}(t) \sim \Omega^{-1} \sim \xi^{-d} \sim t^{d\nu}. \tag{14}$$

Comparing (13) and (14), we obtain the desired relationship

$$d\nu = 2 - \alpha, \tag{15}$$

which is generally referred to as a *hyperscaling relation*—to emphasize the fact that it goes beyond and cannot be derived from the ordinary scaling formulation of Sec. 11.10 without invoking something else, such as the domain volume Ω.

Relation (15) is consistent with the *Josephson inequalities* (1967)

$$d\nu \geq 2 - \alpha, \quad d\nu' \geq 2 - \alpha', \tag{16a, b}$$

proved rigorously by Sokal (1981); of course, the scaling theory does not distinguish between exponents pertaining to $t > 0$ and their counterparts pertaining to $t < 0$. It is important to note that the classical exponents ($\nu = \frac{1}{2}, \alpha = 0$) satisfy (15) *only* for $d = 4$, which shows that the hyperscaling relations, (15) and any others that follow from it, have a rather different status than the other scaling relations (that do not involve d explicitly). The renormalization group theory, to be discussed in Chap. 13, shows why the hyperscaling relations are to be expected fairly generally, why typically they hold for $d < 4$ but break down for $d > 4$; see also Sec. 11.13. The reader may check that relation (15) is satisfied reasonably well by the experimental data of Table 11.1, with $d = 3$; it is also satisfied by the exponents derived theoretically by solving different models exactly, or almost exactly, as in Chap. 12.

Combining (15) with other scaling relations, see Sec. 11.10, we may write

$$d\nu = 2 - \alpha = 2\beta + \gamma = \beta(\delta + 1) = \gamma(\delta + 1)/(\delta - 1). \tag{17}$$

It follows that

$$2 - \eta = \gamma/\nu = d(\delta - 1)/(\delta + 1), \tag{18}$$

which is consistent with the *Buckingham–Gunton inequality* (1969)

$$2 - \eta \leq d(\delta - 1)/(\delta + 1). \tag{19}$$

Notice that the experimental observation that $\delta < 4.8$, for magnetic systems in three dimensions, implies through this inequality that $\eta \geq 0.034$.

11.13. A final look at the mean field theory

We now return to the question: why does the mean field theory fail to represent the true nature of the critical phenomena in real systems? The short answer is–because it neglects fluctuations! As emphasized in Sec. 11.11, correlations among the microscopic constituents of a given system are at the very heart of the phenomenon of phase transitions, for it is through them that the system acquires a long-range order (even when the microscopic interactions are themselves short-ranged). At the same time, there is so direct a relationship between correlations and fluctuations, see eqn. (11.11.6), that they grow together and, as the critical point is approached, become a dominant feature of the system. Neglecting fluctuations is, therefore, a serious drawback of the mean field theory.

The question now arises: is mean field theory *ever* valid? In other words, can fluctuations *ever* be neglected? To answer this question, we recall the fluctuation–dissipation relation (11.11.10a), viz.

$$\chi = (\overline{M^2} - \overline{M}^2)/kT = \overline{(\Delta M)^2}/kT \tag{1}$$

and write it in the form

$$\overline{(\Delta M)^2}/\overline{M}^2 = kT\chi/\overline{M}^2. \tag{2}$$

Now, in order that the neglect of fluctuations be justified, we must have:

$$kT\chi \ll \overline{M}^2. \tag{3}$$

Requirement (3) is generally referred to as the *Ginzburg criterion* (1960); for a more detailed discussion of this criterion, along with physical illustrations, see Als-Nielsen and Birgeneau (1977).

We apply condition (3) to a domain, of volume $\Omega \sim \xi^d$, close to but below the critical point; we are assuming here a system of the Ising type ($n = 1$), so that ξ is finite for $t < 0$. Invoking the power-law behavior of χ and $\overline{M}$, we have

$$kT_c\left(A\xi^d|t|^{-\gamma}\right) \ll \left(B\xi^d|t|^{\beta}\right)^2 \quad (t \lesssim 0), \tag{4}$$

where A and B are positive constants. Since $\xi \sim a|t|^{-\nu}$, we get

$$|t|^{d\nu-2\beta-\gamma} \ll B^2a^d/AkT_c. \tag{5}$$

In view of the scaling relation $\alpha + 2\beta + \gamma = 2$, we may as well write

$$|t|^{d\nu-(2-\alpha)} \ll D, \tag{6}$$

where D is a positive number of order unity.[19] For the mean field theory (with $\nu = \frac{1}{2}, \alpha = 0$) to be valid, condition (6) assumes the form

$$|t|^{(d-4)/2} \ll D. \tag{7}$$

Now, since $|t|$ can be made as small as we like, condition (7) will be violated unless $d > 4$. We therefore conclude that the mean field theory is valid for $d > 4$; by implication, it is inadequate for $d \leq 4$.

The above result has been established for scalar models ($n = 1$) only. In Sec. 12.4, we shall see that in the case of the spherical model, which pertains to the limit $n \to \infty$, the mean field results do apply when $d > 4$. This means that, once again, fluctuations *can* be neglected if $d > 4$. Now fluctuations are supposed to decrease with decreasing n; the validity of the mean field theory for $d > 4$ should, therefore, hold for *all* n.

Ordinarily, when a system is undergoing a phase transition, expression (2), which is a measure of the relative fluctuations in the system, is expected to be of order unity. Condition (6) then suggests that the exponents ν and α obey the hyperscaling relation

$$d\nu = 2 - \alpha. \tag{8}$$

Experience shows that this relation is indeed obeyed when $d < 4$. At $d = 4$, the mean field theory begins to take over and thereafter, for all $d > 4$, the critical exponents are stuck at the mean field values (which are independent of both d and n). The dimensionality $d = 4$ is often referred to as the *upper critical dimension* for the kind of systems under consideration.

An alternative way of looking into the question posed at the beginning of this section is to examine the specific heat of the system which, according to the mean field theory, undergoes a jump discontinuity at the critical point whereas in real systems it shows a weak divergence. The question now arises: what is the source of this divergence that is missed by the mean field theory? The answer again lies in the "neglect of fluctuations". To see it more explicitly, let us look at the internal energy of the system which, in the absence of the field, is given by

$$U = -J\overline{\left(\sum_{\text{n.n.}} \sigma_i \sigma_j\right)} = -J \sum_{\text{n.n.}} \overline{\sigma_i \sigma_j}. \tag{9}$$

In the mean field theory, one replaces $\overline{\sigma_i \sigma_j}$ by $\overline{\sigma_i}\ \overline{\sigma_j}(=\overline{\sigma}^2)$, see eqn. (11.5.5), which leads to the jump discontinuity in the specific heat of magnitude $\frac{3}{2}/Nk$; see eqn. (11.5.18). The *fluctuational* part of U, which is neglected in the mean field theory, may be written as

$$U_f = -J \sum_{\text{n.n.}} \left(\overline{\sigma_i \sigma_j} - \overline{\sigma}_i \overline{\sigma}_j\right) = -J \sum_{\text{n.n.}} g(r_{ij}), \tag{10}$$

where $g(r)$ is the spin–spin correlation function, for which we may use the mean field expression (11.11.26); thus, we will be using the mean field theory itself to predict its own shortcomings! Since the nearest-neighbor distances r_{ij} in (10) are all much smaller than ξ, one may be tempted to use for $g(r_{ij})$ the zeroth-order approximation (11.11.28). This, however, produces a temperature-independent term, which does not contribute to the specific heat of the system. We must, therefore, go to the next approximation, which can be obtained by using the asymptotic formulae

$$K_\mu(x)|_{x\ll 1} \approx \begin{cases} \frac{1}{2}\Gamma(\mu)\left(\frac{1}{2}x\right)^{-\mu} + \frac{1}{2}\Gamma(-\mu)\left(\frac{1}{2}x\right)^{\mu} & \text{for} \quad 0 < \mu < 1 \quad (11a) \\ x^{-1} + \left(\frac{1}{2}x\right)\ln\left(\frac{1}{2}x\right) & \text{for} \quad \mu = 1 \quad (11b) \\ \frac{1}{2}\Gamma(\mu)\left(\frac{1}{2}x\right)^{-\mu} - \frac{1}{2}\Gamma(\mu-1)\left(\frac{1}{2}x\right)^{2-\mu} & \text{for} \quad \mu > 1. \quad (11c) \end{cases}$$

with $\mu = (d-2)/2$ and $x = r_{ij}/\xi$. The temperature-dependent part of U_f comes from the second term(s) in (11); remembering that ξ here is $\sim at^{-1/2}$, we get[20]

$$(U_f/NJ)_{\text{thermal}} \sim \begin{cases} t^{(d-2)/2} & \text{for} \quad 2 < d < 4 \quad (12a) \\ t\ln(1/t) & \text{for} \quad d = 4 \quad (12b) \\ t & \text{for} \quad d > 4. \quad (12c) \end{cases}$$

The *fluctuational* part of the specific heat thus turns out to be

$$C_f/Nk \sim \begin{cases} t^{(d-4)/2} & \text{for} \quad 2 < d < 4 \quad (13a) \\ \ln(1/t) & \text{for} \quad d = 4 \quad (13b) \\ \text{const.} & \text{for} \quad d > 4. \quad (13c) \end{cases}$$

It follows that the specific-heat singularity for $d > 4$ is indeed a "jump discontinuity", and hence the mean field theory remains applicable in this case. For $d = 4$, on the other hand, C_f shows a logarithmic divergence and for $d < 4$ a power-law divergence, making mean field theory invalid for $d \leq 4$.

It is rather instructive to see what part of the fluctuation–correlation spectrum, $\tilde{g}(k)$, contributes significantly to the divergence of the specific heat at $t = 0$. For this, we examine the quantity $-\partial g(r_{ij})/\partial t$ which essentially determines the behavior of C_f near the critical point; see eqn. (10). Using eqn. (11.11.25a), we have

$$-\frac{\partial g(r_{ij})}{\partial t} \sim a^d \int \frac{e^{-ik\cdot r_{ij}}}{t^2(1+\xi^2k^2)^2} k^{d-1}\, dk. \tag{14}$$

Now, the values of k that are much larger than ξ^{-1} contribute little to this integral: the only significant contributions come from the range $(0, k_{\max})$, where $k_{\max} = O(\xi^{-1})$; moreover, since $r_{ij} \ll \xi$, the exponential for these k is essentially equal to 1. Expression (14) may, therefore, be written as

$$-\frac{\partial g(r_{ij})}{\partial t} \sim a^d \int \frac{k^{d-1}\, dk}{t^2(1+\xi^2k^2)^2} \tag{15}$$

which, for $d < 4$, *is* $\sim (a/\xi)^d t^{-2} \sim t^{(d-4)/2}$, as in (13a). We thus see that the most significant contribution to the criticality of the problem arises from fluctuations whose length scale, k^{-1}, is of order ξ or longer and hardly any contribution comes from fluctuations whose length scale is shorter. Now, it is only the latter that are likely to pick up the structural details of the system at the atomic level; since they do not play any significant role in bringing about the phenomenon, the precise nature of criticality remains independent of the structural details. This explains why a large variety of systems, differing so much in their structures, may, insofar as critical behavior is concerned, fall into a single universality class.

Problems

11.1. Assume that in the virial expansion

$$\frac{Pv}{kT} = 1 - \sum_{j=1}^{\infty} \frac{j}{j+1}\beta_j \left(\frac{\lambda^3}{v}\right)^j \tag{9.4.22}$$

where the β_j are the irreducible cluster integrals of the system, only terms with $j = 1$ and $j = 2$ are appreciable in the critical region. Determine the relationship between β_1 and β_2 at the critical point, and show that $kT_c/P_cv_c = 3$.

11.2. Assuming the *Dietrici* equation of state,

$$P(v-b) = kT\exp(-a/kTv),$$

evaluate the critical constants P_c, v_c and T_c of the given system in terms of the parameters a and b, and show that the quantity $kT_c/P_cv_c = e^2/2 \simeq 3.695$.

Further show that the following statements hold in regard to the Dietrici equation of state:

(a) It yields the same expression for the second virial coefficient B_2 as the van der Waals equation does.

(b) For all values of P and for $T \geq T_c$, it yields a *unique* value of v.

(c) For $T < T_c$, there are *three* possible values of v for certain values of P and the critical volume v_c is always intermediate between the largest and the smallest of the three volumes.

(d) The Dietrici equation of state yields the same critical exponents as the van der Waals equation does.

11.3. Consider a non-ideal gas obeying a modified van der Waals equation of state

$$(P + a/v^n)(v-b) = RT \quad (n > 1).$$

Examine how the critical constants P_c, v_c and T_c, and the critic[illegible]onents β, γ, γ' and δ, of this system depend on the number n.

11.4. Following expressions (11.5.2), define

$$p = (1+L)/2 \quad \text{and} \quad q = (1-L)/2 \quad (-1 \le L \le 1) \tag{1}$$

as the probabilities that a spin chosen at random in a lattice composed of N spins is either "up" or "down". The partition function of the system may then be written as

$$Q(B,T) = \sum_L g(L) e^{\beta N\left(\frac{1}{2}qJL^2 + \mu BL\right)}, \tag{2}$$

where $g(L)$ is the "degree of degeneracy" associated with a particular value of L, i.e.

$$g(L) = N!/(Np)!(Nq)!; \tag{3}$$

note that in writing the Hamiltonian here we have made the assumption of *random mixing*, according to which

$$(N_{++} + N_{--} - N_{+-}) = \tfrac{1}{2}qN(p^2 + q^2 - 2pq) = \tfrac{1}{2}qNL^2.$$

(a) Determine the value, L^*, of L that *maximizes* the summand in (2). Check that L^* is identical with the mean value, $\bar{L}$, given by eqn. (11.5.10).

(b) Write down the free energy A and the internal energy U of the system, and show that the entropy S conforms to the relation

$$S(B,T) = -Nk(p^* \ln p^* + q^* \ln q^*),$$

where $p^* = p(L^*)$ and $q^* = q(L^*)$.

11.5. Using the correspondence established in Sec. 11.4, apply the results of the preceding problem to the case of a lattice gas. Show, in particular, that the pressure, P, and the volume per particle, v, are given by

$$P = \mu B - \frac{1}{8}q\varepsilon_0(1+\bar{L}^2) - \frac{1}{2}kT \ln\left(\frac{1-\bar{L}^2}{4}\right)$$

and

$$v^{-1} = \tfrac{1}{2}(1 \pm \bar{L}).$$

Check that the critical constants of this system are: $T_c = q\varepsilon_0/4k$, $P_c = kT_c(\ln 2 - \frac{1}{2})$ and $v_c = 2$, so that the quantity $kT_c/P_c v_c = 1/(\ln 4 - 1) \simeq 2.589$.

11.6. Consider an Ising model with an *infinite-range* interaction such that each spin interacts equally strongly with all other spins:

$$H = -c\sum_{i<j}\sigma_i\sigma_j - \mu B\sum_i \sigma_i.$$

Express this Hamiltonian in terms of the parameter $L(= N^{-1}\Sigma_i\sigma_i)$ and show that, in the limit $N \to \infty$ and $c \to 0$, the mean field theory, with $J = Nc/q$, is exact for this model.

11.7. Study the Heisenberg model of a ferromagnet, based on the interaction (11.3.6), in the *mean field approximation* and show that this also leads to a phase transition of the kind met with in the Ising model. Show, in particular, that the transition temperature T_c and the Curie–Weiss constant C are given by

$$T_c = \frac{qJ}{k}\frac{2S(S+1)}{3} \quad \text{and} \quad C = \frac{N(g\mu_B)^2}{Vk}\frac{S(S+1)}{3}.$$

Note that the ratio $T_c/CV = 2qJ/N(g\mu_B)^2$ is the *molecular field constant* of the problem; cf. eqn. (11.5.8).

11.8. Study the spontaneous magnetization of the Heisenberg model in the *mean field approximation* and examine the dependence of $\bar{L}_0$ on T (i) in the neighborhood of the critical temperature where $(1 - T/T_c) \ll 1$, and (ii) at sufficiently low temperatures where $T/T_c \ll 1$. Compare these results with the corresponding ones, viz. (11.5.14 and 15), for the Ising model.

[In this connection, it may be pointed out that, at very low temperatures, the experimental data do not agree with the theoretical formula derived here. We find instead a much better agreement with the formula $\bar{L}_0 = \{1 - A(kT/J)^{3/2}\}$, where A is a numerical constant (equal to 0.1174 in the case of a simple cubic [illegible]). This formula is known as Bloch's $T^{3/2}$-law and is derivable from the spin-wave theory of ferromagnetism; see Wannier (1966), sec. 15.5.]

11.9. An antiferromagnet is characterized by the fact that the exchange integral J is negative, which tends to align neighboring spins antiparallel to one another. Assume that a given lattice structure is such that the whole lattice can be divided into two interpenetrating sub-lattices, a and b say, so that the spins belonging to each sub-lattice, a as well as b, tend to align themselves in the same direction, while the directions of alignment in the two sub-lattices are opposite to one another. Using the Ising as well as Heisenberg type of interaction, and working in the mean field approximation, evaluate the paramagnetic susceptibility of such a lattice at high temperatures.

11.10. The *Néel temperature* T_N of an antiferromagnet is defined as that temperature below which the sub-lattices a and b possess *nonzero* spontaneous magnetizations M_a and M_b. Determine T_N for the model described in the preceding problem.

11.11. Suppose that each atom of a crystal lattice can be in one of several *internal states* (which may be denoted by the symbol σ) and the interaction energy between an atom in state σ' and its nearest neighbor in state σ'' is denoted by $u(\sigma', \sigma'')\{= u(\sigma'', \sigma')\}$. Let $f(\sigma)$ be the probability of an atom being in the particular state σ, *independently* of the states in which its nearest neighbors are. The interaction energy and the entropy of the lattice may then be written as

$$E = \tfrac{1}{2}qN \sum_{\sigma', \sigma''} u(\sigma', \sigma'') f(\sigma') f(\sigma'')$$

and

$$S/Nk = -\sum_{\sigma} f(\sigma) \ln f(\sigma),$$

respectively. Minimizing the free energy $(E - TS)$, show that the equilibrium value of the function $f(\sigma)$ is determined by the equation

$$f(\sigma) = C \exp\{-(q/kT)\Sigma_{\sigma'} u(\sigma, \sigma') f(\sigma')\},$$

where C is the constant of normalization. Further show that, for the special case $u(\sigma', \sigma'') = -J\sigma'\sigma''$, where the σ can be either +1 or −1, this equation reduces to the Weiss eqn. (11.5.11), with $f(\sigma) = \frac{1}{2}(1 + \bar{L}_0 \sigma)$.

11.12. Consider a binary alloy containing N_A atoms of type A and N_B atoms of type B, so that the relative concentrations of the two components are: $x_A = N_A/(N_A + N_B) \leq \frac{1}{2}$ and $x_B = N_B/(N_A + N_B) \geq \frac{1}{2}$. The degree of *long-range order*, X, is such that

$$\begin{bmatrix} A \\ a \end{bmatrix} = \tfrac{1}{2}Nx_A(1 + X), \quad \begin{bmatrix} A \\ b \end{bmatrix} = \tfrac{1}{2}Nx_A(1 - X),$$

$$\begin{bmatrix} B \\ a \end{bmatrix} = \tfrac{1}{2}N(x_B - x_A X), \quad \begin{bmatrix} B \\ b \end{bmatrix} = \tfrac{1}{2}N(x_B + x_A X),$$

where $N = N_A + N_B$, while the symbol $\begin{bmatrix} A \\ a \end{bmatrix}$ denotes the number of atoms of type A occupying sites of the sub-lattice a, and so on. In the *Bragg–Williams approximation*, the number of nearest-neighbor pairs of different kinds can be written down straightaway; for instance,

$$\begin{bmatrix} AA \\ ab \end{bmatrix} = \tfrac{1}{2}qN \cdot x_A(1 + X) \cdot x_A(1 - X),$$

and so on. The configurational energy of the lattice then follows from eqn. (11.4.9). In the same approximation, the entropy of the lattice is given by $S = k \ln W$, where

$$W = \frac{\left(\frac{1}{2}N\right)!}{\begin{bmatrix} A \\ a \end{bmatrix}! \begin{bmatrix} B \\ a \end{bmatrix}!} \cdot \frac{\left(\frac{1}{2}N\right)!}{\begin{bmatrix} A \\ b \end{bmatrix}! \begin{bmatrix} B \\ b \end{bmatrix}!}.$$

Minimizing the free energy of the lattice, show that the equilibrium value of X is determined by the equation

$$\frac{X}{x_B + x_A X^2} = \tanh\left(\frac{2qx_A\varepsilon}{kT} X\right); \quad \varepsilon = \tfrac{1}{2}(\varepsilon_{11} + \varepsilon_{22}) - \varepsilon_{12} > 0.$$

Note that, in the special case of equal concentrations ($x_A = x_B = \frac{1}{2}$), this equation assumes the more familiar form

$$X = \tanh\left(\frac{q\varepsilon}{2kT}X\right)$$

Further show that the transition temperature of the system is given by

$$T_c = 4x_A(1 - x_A)T_c^0,$$

where $T_c^0(= q\varepsilon/2k)$ is the transition temperature in the case of equal concentrations.

[*Note*: In the Kirkwood approximation [see Kubo (1965), problem 5.19], T_c^0 turns out to be $(\varepsilon/k)\{1 - \surd[1 - (4/q)]\}^{-1}$, which may be written as $(q\varepsilon/2k)(1 - 1/q + \cdots)$. To this order, the Bethe approximation also yields the same result.]

11.13. Consider a two-component solution of N_A atoms of type A and N_B atoms of type B, which are supposed to be randomly distributed over $N(= N_A + N_B)$ sites of a single lattice. Denoting the energies of the nearest-neighbor pairs AA, BB and AB by ε_{11}, ε_{22} and ε_{12}, respectively, write down the free energy of the system in the Bragg–Williams approximation and evaluate the chemical potentials μ_A and μ_B of the two components. Next, show that if $\varepsilon = (\varepsilon_{11} + \varepsilon_{22} - 2\varepsilon_{12}) < 0$, i.e. if the atoms of the same species display greater affinity to be neighborly, then for temperatures below a critical temperature T_c, which is given by the expression $q|\varepsilon|/2k$, the solution separates out into two phases of unequal relative concentrations.

[*Note*: For a study of phase separation in an isotopic mixture of hard-sphere bosons and fermions, and for the relevance of this study to the actual behavior of He^3–He^4 solutions, see Cohen and van Leeuwen (1960, 1961).

11.14. Modify the Bragg–Williams approximation (11.5.29) to include a *short-range order* parameter s, such that

$$\overline{N}_{++} = \frac{1}{2}qN\gamma\left(\frac{1+\overline{L}}{2}\right)^2(1+s), \quad \overline{N}_{--} = \frac{1}{2}qN\gamma\left(\frac{1-\overline{L}}{2}\right)^2(1+s),$$

$$\overline{N}_{+-} = 2\cdot\frac{1}{2}qN\gamma\left(\frac{1+\overline{L}}{2}\right)\left(\frac{1-\overline{L}}{2}\right)(1-s).$$

(a) Evaluate γ from the condition that the total number of nearest-neighbor pairs is $\frac{1}{2}qN$.

(b) Show that the critical temperature T_c of this model is $(1 - s^2)qJ/k$.

(c) Determine the nature of the specific-heat singularity at $T = T_c$, and compare your result with both the Bragg–Williams approximation of Sec. 11.5 and the Bethe approximation of Sec. 11.6.

11.15. Show that in the *Bethe approximation* the entropy of the Ising lattice at $T = T_c$ is given by the expression

$$\frac{S_c}{Nk} = \ln 2 + \frac{q}{2}\ln\left(1 - \frac{1}{q}\right) - \frac{q(q-2)}{4(q-1)}\ln\left(1 - \frac{2}{q}\right).$$

Compare this result with the one following from the *Bragg–Williams approximation*, viz. (11.5.20).

11.16. Examine the critical behavior of the low-field susceptibility, χ_0, of an Ising model in the Bethe approximation of Sec. 11.6, and compare your results with eqns (11.5.22) of the Bragg–Williams approximation.

11.17. A function $f(x)$ is said to be *concave* over an interval (a, b) if it satisfies the property

$$f\{\lambda x_1 + (1-\lambda)x_2\} \geq \lambda f(x_1) + (1-\lambda)f(x_2),$$

where x_1 and x_2 are two arbitrary points in the interval (a, b) while λ is a positive number in the interval (0, 1). This means that the chord joining the points x_1 and x_2 lies *below* the curve $f(x)$. Show that this also means that the tangent to the curve $f(x)$ at any point x in the interval (a, b) lies *above* the curve $f(x)$ or, equivalently, that the second derivative $\partial^2 f/\partial x^2$ throughout this interval ≤ 0.

11.18. In view of the thermodynamic relationship

$$C_v = TV(\partial^2 P/\partial T^2)_V - TN(\partial^2\mu/\partial T^2)_V$$

for a fluid, μ being the chemical potential of the system, Yang and Yang (1964) pointed out that, if C_V is singular at $T = T_c$, then either $(\partial^2 P/\partial T^2)_V$ or $(\partial^2\mu/\partial T^2)_V$ or both will be singular. Define an exponent Θ by writing

$$(\partial^2 P/\partial T^2)_V \sim (T_c - T)^{-\Theta} \quad (T \lesssim T_c).$$

and show that (Griffiths 1965b)

$$\Theta \leq \alpha' + \beta \quad \text{and} \quad \Theta \leq (2 + \alpha'\delta)/(\delta + 1).$$

11.19. Determine the numerical values of the coefficients r_1 and s_0 of eqn. (11.9.5) in (i) the Bragg–Williams approximation of Sec. 11.5 and (ii) the Bethe approximation of Sec. 11.6. Using these values of r_1 and s_0, verify that eqns (11.9.4, 9–11, and 14) reproduce correctly the results obtained in the zeroth and the first approximation, respectively.

11.20. Consider a system with a modified expression for the Landau free energy, viz.

$$\psi_h(t, m) = -hm + q(t) + r(t)m^2 + s(t)m^4 + u(t)m^6,$$

with $u(t)$ a fixed positive constant. Minimize ψ with respect to the variable m and examine the spontaneous magnetizaiion m_0 as a function of the parameters r and s. In particular, show the following:[21]

(a) For $r > 0$ and $s > -(3ur)^{1/2}$, $m_0 = 0$ is the only real solution.

(b) For $r > 0$ and $-(4ur)^{1/2} < s \leq -(3ur)^{1/2}$, $m_0 = 0$ or $\pm m_1$, where $m_1^2 = \dfrac{\sqrt{(s^2 - 3ur)} - s}{3u}$. However, the minimum of ψ at $m_0 = 0$ is lower than the minima at $m_0 = \pm m_1$, so the ultimate equilibrium value of m_0 is 0.

(c) For $r > 0$ and $s = -(4ur)^{1/2}$, $m_0 = 0$ or $\pm(r/u)^{1/4}$. Now, the minimum of ψ at $m_0 = 0$ is of the same height as the ones at $m_0 = \pm(r/u)^{1/4}$, so a nonzero spontaneous magnetization is as likely to occur as the zero one.

(d) For $r > 0$ and $s < -(4ur)^{1/2}$, $m_0 = \pm m_1$–which implies a *first-order* phase transition (because the two possible states available here differ by a *finite* amount in m). The line $s = -(4ur)^{1/2}$, with r positive, is generally referred to as a "line of first-order phase transitions".

(e) For $r = 0$ and $s < 0$, $m_0 = \pm(2|s|/3u)^{1/2}$.

(f) For $r < 0$, $m_0 = \pm m_1$ for all s. As $r \to 0$, $m_1 \to 0$ if s is positive.

(g) For $r = 0$ and $s > 0$, $m_0 = 0$ is only solution. Combining this result with (f), we conclude that the line $r = 0$, with s positive, is a "line of second-order phase transitions", for the two states available here differ by a *vanishing* amount in m.

The lines of first-order phase transitions and second-order phase transitions meet at the point ($r = 0, s = 0$), which is commonly referred to as a *tricritical point* (Griffiths, 1970).

11.21. In the preceding problem, put $s = 0$ and approach the tricritical point along the r-axis, setting $r \approx r_1 t$. Show that the critical exponents pertaining to the tricritical point in this model are

$$\alpha = \tfrac{1}{2}, \beta = \tfrac{1}{4}, \gamma = 1 \textit{ and } \delta = 5.$$

11.22. Consider a fluid near its critical point, with isotherms as sketched in Fig. 11.3. Assume that the singular part of the Gibbs free energy of the fluid is of the form

$$G^{(s)}(T, P) \sim |t|^{2-\alpha} g(\pi/|t|^{\Delta}).$$

where $\pi = (P - P_c)/P_c$, $t = (T - T_c)/T_c$ while $g(x)$ is a *universal* function, with branches g_+ for $t > 0$ and g_- for $t < 0$; in the latter case, the function g_- has a point of *infinite* curvature at a value of π which varies smoothly with t, such that $\pi(0) = 0$ and $(\partial\pi/\partial t)_{t\to 0} = \text{const.}$

(a) Using the above expression for $G^{(s)}$, determine the manner in which the densities, ρ_l and ρ_g, of the two phases approach one another as $t \to 0$ from below.

(b) Also determine how $(P - P_c)$ varies with $(\rho - \rho_c)$ as the critical point is approached along the critical isotherm ($t = 0$).

(c) Examine as well the critical behavior of the isothermal compressibility κ_T, the adiabatic compressibility κ_S, the specific heats C_P and C_V, the coefficient of volume expansion α_P and the latent heat of vaporization l.

11.23. Consider a model equation of state which, near the critical point, can be written as

$$h \approx am(t + bm^2)^{\Theta} \quad (1 < \Theta < 2; a, b > 0).$$

Determine the critical exponents β, γ and δ of this model, and check that they obey the scaling relation (11.10.22).

11.24. Assuming that the correlation function $g(\boldsymbol{r}_i, \boldsymbol{r}_j)$ is a function only of the distance $r = |\boldsymbol{r}_j - \boldsymbol{r}_i|$, show that $g(r)$ for $r \neq 0$ satisfies the differential equation

$$\frac{d^2g}{dr^2} + \frac{d-1}{r}\frac{dg}{dr} - \frac{1}{\xi^2}g = 0.$$

Check that expression (11.11.27) for $g(r)$ satisfies this equation in the regime $r \gg \xi$, while expression (11.11.28) does so in the regime $r \ll \xi$.

11.25. Consider the correlation function $g(r; t, h)$ of Sec. 11.11 with $h > 0$.[22] Assume that this function has the following behavior:

$$g(r) \sim e^{-r/\xi(t,h)} \times \text{some power of } r (t \gtrsim 0),$$

such that $\xi(0, h) \sim h^{-\nu^c}$. Show that $\nu^c = \nu/\Delta$.

Next, assume that the susceptibility $\chi(0, h) \sim h^{-\gamma^c}$. Show that $\gamma^c = \gamma/\Delta = (\delta - 1)/\delta$.

11.26. Liquid He^4 undergoes a superfluid transition at $T \simeq 2.17$ K. The order parameter in this case is a complex number Ψ, which is related to the Bose condensate density ρ_0 as

$$\rho_0 \sim |\Psi|^2 \sim |t|^{2\beta} \quad (t \lesssim 0).$$

The superfluid density ρ_s, on the other hand, behaves as

$$\rho_s \sim |t|^{\nu} \quad (t \lesssim 0).$$

Show that the ratio[23]

$$(\rho_0/\rho_s) \sim |t|^{\eta\nu} \quad (t \lesssim 0).$$

11.27. The surface tension, σ, of a liquid approaches zero as $T \to T_c$ from below. Define an exponent μ by writing

$$\sigma \sim |t|^{\mu} \quad (t \lesssim 0).$$

Identifying σ with the "free energy associated with a unit area of the liquid–vapor interface", argue that $\mu = (d-1)\nu = (2-\alpha)(d-1)/d$.

[*Note*: Analysis of the experimental data on surface tension yields: $\mu = 1.27 \pm 0.02$, which agrees with the fact that for most fluids $\alpha \simeq 0.1$.]

Notes

[1] In this connection, one should note that the mathematical schemes developed in Chapters 9 and 10 give reliable results only if the interactions among the microscopic constituents of the given system are sufficiently weak–in fact, too weak to bring about *cooperative transitions*.

[2] For historical details, see Griffiths (1972), p. 12.

[3] The physical basis of the Maxwell construction can be seen with the help of the Gibbs free energy density $g(T, P)$. Since $dg = -s\,dT + v\,dP$ and along the "corrected" isotherm $dP = dT = 0$, it follows that $g_1 = g_3$; see Fig. 11.2. To achieve the same result from the theoretical isotherm (along which $dT = 0$ but $dP \neq 0$), we clearly require that the quantity $v dP$, integrated along the isotherm from state 1 to state 3, must vanish; this leads to the "theorem of equal areas".

[4] Compare eqn. (11), which has been derived here *non-thermodynamically*, with eqn. (4.5.7) derived earlier.

[5] The presence of such a region entails that $(\overline{N^2} - \overline{N}^2)$ be $O(\overline{N}^2)$. This implies that the fluctuations in the variable N be *macroscopically* large, which in turn implies equally large fluctuations in the variable v within the system and hence the co-existence of two or more phases with different values of $\overline{v}$. In a single-phase state, $(\overline{N^2} - \overline{N}^2)$ is $O(\overline{N})$; the slope $(\partial P'/\partial \overline{v})$ is then $O(\overline{N}^0)$, as an intensive quantity should be.

[6] A more precise formulation of the van der Waals theory, as the limit of a theory with an infinite-range potential, has been formulated by Kac, Uhlenbeck and Hemmer (1963). They considered the potential

$$u(r) = \begin{cases} +\infty & \text{for } r \leq \sigma \\ -\kappa e^{-\kappa r} & \text{for } r > \sigma, \end{cases}$$

so that the integral $\int_{\sigma}^{\infty} u(r) dr$ is simply $-\exp(-\kappa\sigma)$; when $\kappa \to 0$ the potential becomes infinite in range but infinitesimally weak. Kac *et al.* showed that in this limit the model becomes essentially the same as van der Waals'–with one noteworthy improvement, i.e. no unphysical regions in the (P, v)-diagram appear and hence no need for the Maxwell construction arises.

[7] For an historical account of the origin and development of the Lenz–Ising model, see the review article by Brush (1967). This review gives a large number of other references as well.

[8] Henceforth, [illegible] use the symbol μ instead of μ_B.

[9] The coordination number q for a linear chain is obviously 2; for two-dimensional lattices, viz. honeycomb, square and triangular, it is 3, 4 and 6, respectively; for three-dimensional lattices, viz. simple cubic, body-centered cubic and face-centered cubic, it is 6, 8 and 12, respectively.

[10] In connection with the present approximation, we may as well mention that early attempts to construct a theory of binary solutions were based on the assumption that the atoms in the solution mix randomly. One finds that the results following from this assumption of *random mixing* are mathematically equivalent to the ones following from the *mean field approximation*; see Problems 11.12 and 11.13.

[11] The concept of "broken symmetry" plays a vital role in this and many other phenomena in physics; for details, see Fisher (1972) and Anderson (1984).

[12] Recall eqn. (3.3.14), whereby $S = k \ln \Omega$.

[13] The special case $n = 2$ with $d = 2$ is qualitatively different from others; for details, see Sec. 12.6.

[14] Recalling the correspondence between a gas–liquid system and a magnet, one might wonder why we have employed C_H, rather than C_M, in place of C_V. The reason is that, since we are letting $H \to 0$ and $T \to T_c-$, $M \to 0$ as well. So, as argued by Fisher (1967), in the limit considered here, C_H and C_M display the *same* singular behavior. In fact, it can be shown that if the ratio $C_M/C_H \to 1$ as $T \to T_c$, then $(\alpha' + 2\beta + \gamma')$ must be greater than 2; on the other hand, if this ratio tends to a value less than 1, then $(\alpha' + 2\beta + \gamma') = 2$. For details, see Stanley (1971), Sec. 4.1.

[15] It may be mentioned here that the passage from eqn. (1) to (6) is equivalent to effecting a Legendre transformation from the Helmholtz free energy A to the Gibbs free energy $G(= A - HM)$, and eqn. (7) is analogous to the relation $(\partial A/\partial M)_T = H$.

[16] In certain systems such as superconductors, the effective interactions (which, for instance, lead to the formation of Cooper pairs of electrons) are, in fact, long-ranged. The critical exponents pertaining to such systems turn out to be the same as one gets from the mean field theory. For details, see Tilley and Tilley (1990).

[17] Remembering that $\overline{M} = \mu \Sigma_i \overline{\sigma}_i$, we change the field $\{H_i\}$ to $\{H_i + \delta H_i\}$, with the result that

$$\delta\overline{M} = \mu \sum_i \left[\sum_j (\partial\overline{\sigma}_i/\partial H_j)\delta H_j \right].$$

Now, for simplicity, we let all δH_j be the same; we then get

$$(\delta\overline{M}/\delta H) = \mu \sum_i \sum_j (\partial\overline{\sigma}_i/\partial H_j).$$

Comparing this with (10), we infer that $(\partial\overline{\sigma}_i/\partial H_j) = \beta\mu g(i, j)$ and hence $(\partial\overline{\sigma}_i/\partial h_j) = g(i, j)$.

[18] It turns out that the exponent ν' is relevant only for scalar models, for which $n = 1$; for vector models ($n \geq 2$), ξ is infinite at all $T \leq T_c$ and hence ν' is irrelevant.

[19] To see this, we note that $A \sim N\mu^2/\Omega k T_c$ while $B \sim N\mu/\Omega$, with the result that $D \sim Na^d/\Omega = O(1)$.

[20] Note that the negative sign in (10) cancels the implicit negative sign of $\Gamma(-\mu)$ in (11a), that of $\ln\left(\frac{1}{2}x\right)$ in (11b) and the explicit negative sign in (11c).

[21] To fix ideas, it is helpful to use (r, s)-plane as our "parameter space".

[22] For more details, see Tarko and Fisher (1975).

[23] The corresponding ratio for a magnetic system is M_0^2/Γ, where M_0 is the spontaneous magnetization and Γ the helicity modulus of the system; for details, see Fisher, Barber and Jasnow (1973).

CHAPTER 12

PHASE TRANSITIONS: EXACT (OR ALMOST EXACT) RESULTS FOR THE VARIOUS MODELS

IN THE preceding chapter we saw that the onset of a phase transition in a given physico-chemical system is characterized by (singular) features whose qualitative nature is determined by the universality class to which the system belongs. In this chapter we propose to consider a variety of model systems belonging to different universality classes and analyze them theoretically to find out how these features arise and how they vary from class to class. In this context we recall that the parameters distinguishing one universality class from another are: (i) the space dimensionality d, (ii) the dimensionality of the order parameter, often referred to as the spin dimensionality, n and (iii) the range of the microscopic interactions. As regards the latter, unless a statement is made to the contrary, we shall assume a short-range interaction which, in most cases, will be of the nearest-neighbor type; the only parameters open for selection will then be d and n.

We shall start our analysis with the Ising model ($n = 1$) in one dimension, followed by a study of the general n-vector models (again in one dimension). We shall then return to the Ising model–this time in two dimensions–and follow it with a study of two models in general d but with $n \to \infty$. These studies will give us a fairly good idea as to what to expect in the most practical, three-dimensional situations for which, unfortunately, we have no exact solutions, though a variety of mathematical techniques have been developed to obtain *almost exact* results in many cases of interest. For completeness, these and other results of physical importance will be visited in the last section of this chapter.

12.1. The Ising model in one dimension

In this section we present an exact treatment of the Ising model in one dimension. This is important for several reasons. First of all, there do exist phenomena, such as adsorption on a linear polymer or on a protein chain, the elastic properties of fibrous proteins, etc., that can be looked upon as one-dimensional nearest-neighbor problems. Secondly, it helps us evolve mathematical techniques for treating lattices in higher dimensions, which is essential for understanding the critical behavior of a variety of physical systems met with in nature. Thirdly, it enables us to estimate the status of the Bethe approximation as a "possible" theory of the Ising model, for it demonstrates mathematically that at least in one dimension this approximation leads to exact results.

In a short paper published in 1925, Ising himself gave an exact solution to this problem in one dimension. He employed a combinatorial approach which has by now been superseded by other approaches. Here we shall follow the *transfer matrix* method, first introduced by Kramers and Wannier (1941). In the one-dimensional case this method worked with immediate success. Three years later, in 1944, it became, through Onsager's ingenuity, the first method to treat successfully the *field-free* Ising model in two dimensions. To apply this method, we replace the actual lattice by one having the topology of a closed, endless structure; thus, in the one-dimensional case we replace the straight, open chain by a curved one such that the Nth spin becomes a neighbor of the first (see Fig. 12.1). This replacement eliminates the inconvenient end effects; it does not, however, alter the thermodynamic properties of the (infinitely long) chain. The important advantage of this replacement is that it enables us to write the Hamiltonian of the system,

$$H_N\{\sigma_i\} = -J\sum_{\text{n.n.}}\sigma_i\sigma_j - \mu B\sum_{i=1}^{N}\sigma_i, \tag{1}$$

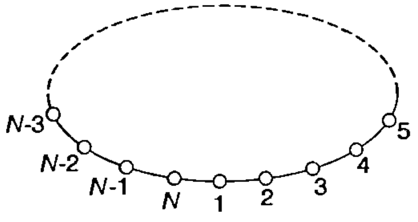


FIG. 12.1. An Ising chain with a closed, endless structure.

in a symmetrical form, namely

$$H_N\{\sigma_i\} = -J\sum_{i=1}^{N}\sigma_i\sigma_{i+1} - \tfrac{1}{2}\mu B\sum_{i=1}^{N}(\sigma_i + \sigma_{i+1}), \tag{2}$$

because $\sigma_{N+1} \equiv \sigma_1$. The partition function of the system is then given by

$$Q_N(B,T) = \sum_{\sigma_1=\pm1}\cdots\sum_{\sigma_N=\pm1}\exp\left[\beta\sum_{i=1}^{N}\left\{J\sigma_i\sigma_{i+1} + \tfrac{1}{2}\mu B\,(\sigma_i + \sigma_{i+1})\right\}\right] \tag{3a}$$

$$= \sum_{\sigma_1=\pm1}\cdots\sum_{\sigma_N=\pm1}\langle\sigma_1|P|\sigma_2\rangle\langle\sigma_2|P|\sigma_3\rangle\cdots\langle\sigma_{N-1}|P|\sigma_N\rangle\langle\sigma_N|P|\sigma_1\rangle, \tag{3b}$$

where P denotes an operator with matrix elements

$$\langle\sigma_i|P|\sigma_{i+1}\rangle = \exp\left[\beta\left\{J\sigma_i\sigma_{i+1} + \tfrac{1}{2}\mu B(\sigma_i + \sigma_{i+1})\right\}\right],$$

that is,

$$(P) = \begin{pmatrix} e^{\beta(J+\mu B)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-\mu B)} \end{pmatrix}. \tag{4}$$

According to the rules of matrix algebra, the summations over the various σ_i in eqn. (3b) lead to the simple result

$$Q_N(B, T) = \sum_{\sigma_1 = \pm 1} \langle \sigma_1 | \boldsymbol{P}^N | \sigma_1 \rangle = \text{Trace } (\boldsymbol{P}^N) = \lambda_1^N + \lambda_2^N, \tag{5}$$

where λ_1 and λ_2 are the eigenvalues of the matrix $\boldsymbol{P}$. These eigenvalues are given by the equation

$$\begin{vmatrix} e^{\beta(J+\mu B)} - \lambda & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-\mu B)} - \lambda \end{vmatrix} = 0, \tag{6}$$

that is, by

$$\lambda^2 - 2\lambda e^{\beta J} \cosh(\beta\mu B) + 2\sinh(2\beta J) = 0. \tag{7}$$

One readily obtains

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = e^{\beta J} \cosh(\beta\mu B) \pm \{e^{-2\beta J} + e^{2\beta J} \sinh^2(\beta\mu B)\}^{1/2}. \tag{8}$$

Quite generally, $\lambda_2 < \lambda_1$; so, $(\lambda_2/\lambda_1)^N \to 0$ as $N \to \infty$. Thus, it is only the *larger* eigenvalue, λ_1, that determines the major physical properties of the system in the thermodynamic limit; see eqn. (5). It follows that

$$\frac{1}{N} \ln Q_N(B, T) \approx \ln \lambda_1 \tag{9}$$

$$= \ln [e^{\beta J} \cosh(\beta\mu B) + \{e^{-2\beta J} + e^{2\beta J} \sinh^2(\beta\mu B)\}^{1/2}]. \tag{10}$$

The Helmholtz free energy then turns out to be

$$A(B, T) = -NJ - NkT \ln [\cosh(\beta\mu B) + \{e^{-4\beta J} + \sinh^2(\beta\mu B)\}^{1/2}]. \tag{11}$$

The various other properties of the system follow readily from eqn. (11). Thus,

$$U(B, T) \equiv -T^2 \frac{\partial}{\partial T}\left(\frac{A}{T}\right) = -NJ - \frac{N\mu B \sinh(\beta\mu B)}{\{e^{-4\beta J} + \sinh^2(\beta\mu B)\}^{1/2}} + \frac{2NJe^{-4\beta J}}{[\cosh(\beta\mu B) + \{e^{-4\beta J} + \sinh^2(\beta\mu B)\}^{1/2}]\{e^{-4\beta J} + \sinh^2(\beta\mu B)\}^{1/2}}, \tag{12}$$

from which the specific heat can be derived, and

$$\overline{M}(B, T) \equiv -\left(\frac{\partial A}{\partial B}\right)_T = \frac{N\mu \sinh(\beta\mu B)}{\{e^{-4\beta J} + \sinh^2(\beta\mu B)\}^{1/2}}, \tag{13}$$

from which the susceptibility can be derived.

Right away we note that, as $B \to 0$, $\overline{M}$ (for all finite β) $\to 0$. This rules out the possibility of spontaneous magnetization, and hence of a phase transition, at any finite temperature T. Of course, at $T = 0$, $\overline{M}$ (for any value of B) is equal to the saturation value $N\mu$, which implies perfect order in the system. This means that there is, after all, a phase transition at a critical temperature T_c, which coincides with absolute zero!

Figure 12.2 shows the degree of magnetization, $\overline{M}$, of the lattice as a function of the parameter $(\beta\mu B)$ for different values of (βJ). For $J = 0$, we have the *paramagnetic* result $\overline{M} = N\mu \tanh(\beta\mu B)$; cf. eqn. (3.9.27). A positive J enhances magnetization and, in turn, leads to a faster approach towards saturation. As $\beta J \to \infty$, the magnetization curve becomes a step function–indicative of a singularity at $T = 0$. The low-field susceptibility of the system is given by the initial slope of the magnetization curve; one obtains

$$\chi_0(T) = \frac{N\mu^2}{kT} e^{2J/kT}. \tag{14}$$

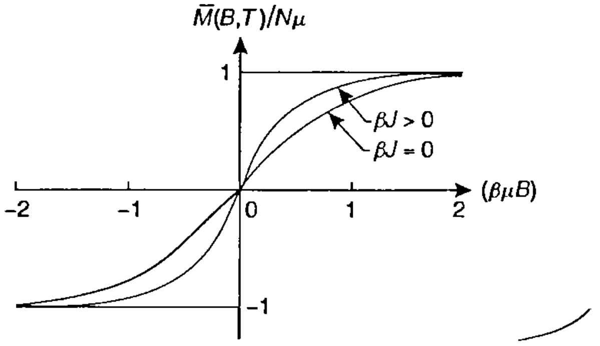


FIG. 12.2. The degree of magnetization of an Ising chain as a function of the parameter $(\beta\mu B)$.

which diverges as $T \to 0$. It should be noted that the singularity here is *not* of the power-law type; it is exponential instead.

The zero-field energy and the zero-field specific heat of the system follow from eqn. (12); one gets

$$U_0(T) = -NJ \tanh(\beta J) \tag{15}$$

and

$$C_0(T) = Nk(\beta J)^2 \operatorname{sech}^2(\beta J). \tag{16}$$

Figure 12.3 shows the variation of the specific heat C_0 as a function of temperature. Although it passes through a maximum, C_0 is a smooth function of T, vanishing as $T \to 0$. Note that eqns (15) and (16) are identical with the corresponding equations, (11.6.29) and (11.6.30), of the Bethe approximation, with coordination number 2 for which $T_c = 0$. It turns out that for a one-dimensional chain the Bethe approximation, in fact, yields *exact* results; for a fuller demonstration of this, see Problem 12.3.

At this stage it seems instructive to express the free energy of the system, near its critical point, in a *scaled* form, as in Sec. 11.10. Unfortunately, there is a problem here. Since $T_c = 0$, the conventional definition, $t = (T - T_c)/T_c$, does not work. A closer look at eqns (11) and (14), however, suggests that we may instead adopt the definition

$$t = e^{-pJ/kT} \quad (p > 0) \tag{17}$$

so that, as $T \to T_c$, $t \to 0$ as desired while for temperatures close to T_c, t is much less than unity. The definition of h remains the same, viz. $\mu B/kT$. The free energy

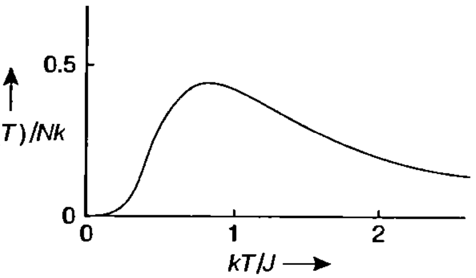


FIG. 12.3. The zero-field specific heat of an Ising chain as a function of temperature.

function $(A + NJ)/NkT$ then takes the form

$$\psi^{(s)}(t, h) = -\ln\left[\cosh h + (t^{4/p} + \sinh^2 h)^{1/2}\right] \tag{18a}$$

$$\approx -(t^{4/p} + h^2)^{1/2} \quad (t, h \ll 1), \tag{18b}$$

which may be written in the scaled form

$$\psi^{(s)}(t, h) \approx t^{2/p} f(h/t^{2/p}). \tag{19}$$

At the same time, eqn. (14) becomes

$$\frac{\chi_0(T)}{(N\mu^2/kT)} \approx t^{-2/p}. \tag{20}$$

Comparing these results with the scaling formulation of Sec. 11.10, we infer that for this model

$$\alpha = 2 - 2/p, \quad \Delta = 2/p \quad \text{and} \quad \gamma = 2/p, \tag{21}$$

in conformity with the exponent relation (11.10.14b). Note that the exponents β and δ for this model cannot be defined in the normal, conventional sense. One may, however, write eqn. (13) in the form

$$m = \sinh h/(t^{4/p} + \sinh^2 h)^{1/2} \tag{22a}$$

$$\approx h/(t^{4/p} + h^2)^{1/2} \quad (t, h \ll 1) \tag{22b}$$

$$= -t^0 f'(h/t^{2/p}), \tag{22c}$$

suggesting that β may formally be taken as zero. At the same time, since

$$m|_{t=0} = 1, \tag{23}$$

which is $\sim h^0$, the exponent δ may formally be taken as infinite.

We shall now study spin–spin correlations in the Ising chain. For this we set $B = 0$ but at the same time generalize the interaction parameter J to become site-dependent (the reason for which will become clear soon). The partition function of the system is then given by

$$Q_N(T) = \sum_{\sigma_1=\pm 1} \quad \sum_{\sigma_N=\pm 1} \prod_i e^{\beta J_i \sigma_i \sigma_{i+1}}; \tag{24}$$

cf. eqn. (3a). With $B = 0$, it is simpler to work with an *open* chain, which has only $(N - 1)$ nearest-neighbor pairs; the advantage of this choice is that in the summand of (24) the variables σ_1 and σ_N appear only once! A summation over either of these can be carried out easily; doing this over σ_N, we have

$$\sum_{\sigma_N=\pm 1} e^{\beta J_{N-1}\sigma_{N-1}\sigma_N} = 2\cosh(\beta J_{N-1}\sigma_{N-1}) = 2\cosh(\beta J_{N-1}), \tag{25}$$

regardless of the sign of σ_{N-1}. We thus obtain the recurrence relation

$$Q_N(T; J_1, \ldots, J_{N-1}) = 2\cosh(\beta J_{N-1}) Q_{N-1}(T; J_1, \ldots, J_{N-2}). \tag{26}$$

By iteration, we get

$$Q_N(T) = \prod_{i=1}^{N-1}\{2\cosh(\beta J_i)\} \sum_{\sigma_1=\pm 1} 1 = 2^N \prod_{i=1}^{N-1} \cosh(\beta J_i), \tag{27}$$

so that

$$\frac{1}{N}\ln Q_N(T) = \ln 2 + \frac{1}{N}\sum_{i=1}^{N-1} \ln\cosh(\beta J_i), \tag{28}$$

which may be compared with eqn. (9)–remembering that, in the absence of the field, $\lambda_1 = 2\cosh(\beta J)$. We are now ready to calculate the correlation function, $g(r)$, of the Ising chain.

It is straightforward to see from eqn. (24) that

$$\overline{\sigma_k \sigma_{k+1}} = \frac{1}{Q_N}\left(\frac{1}{\beta}\frac{\partial}{\partial J_k}\right) Q_N = \left(\frac{1}{\beta}\frac{\partial}{\partial J_k}\right)\ln Q_N. \tag{29}$$

Substituting from eqn. (28), and remembering that $\overline{\sigma}_k = 0$ at all finite temperatures, we obtain for the *nearest-neighbor correlation function*

$$g_k(n.n.) = \overline{\sigma_k \sigma_{k+1}} = \tanh(\beta J_k). \tag{30}$$

For a pair of spins separated by r lattice constants, we get

$$\begin{aligned} g_k(r) &= \overline{\sigma_k \sigma_{k+r}} = \overline{(\sigma_k \sigma_{k+1})(\sigma_{k+1}\sigma_{k+2})\ldots(\sigma_{k+r-1}\sigma_{k+r})} \quad (\text{since all } \sigma_i^2 = 1) \\ &= \frac{1}{Q_N}\left(\frac{1}{\beta}\frac{\partial}{\partial J_k}\right)\left(\frac{1}{\beta}\frac{\partial}{\partial J_{k+1}}\right)\cdots\left(\frac{1}{\beta}\frac{\partial}{\partial J_{k+r-1}}\right) Q_N \\ &= \prod_{i=k}^{k+r-1} \tanh(\beta J_i). \end{aligned} \tag{31}$$

Reverting to a common J, we obtain the desired result

$$g(r) = \tanh^r(\beta J), \tag{32}$$

which may be written in the standard form

$$g(r) = e^{-r/\xi}, \quad \text{with} \quad \xi = [\ln\coth(\beta J)]^{-1}. \tag{33a, b}$$

For $\beta J \gg 1$,

$$\xi \approx \tfrac{1}{2} e^{2\beta J}, \tag{34}$$

which diverges as $T \to 0$. In terms of the variable t, defined in eqn. (17),

$$\xi \sim t^{-2/p} \quad (t \ll 1), \tag{35}$$

giving $\nu = 2/p$. And since our $g(r)$ does not contain any power of r, we infer that $(d - 2 + \eta) = 0$ – giving $\eta = 1$; one may check that the same result follows from eqn. (11.12.10) or (11.12.11). In passing, we note that, regardless of the choice of the number p in defining t, we have for this model

$$\gamma = \nu = 2 - \alpha. \tag{36}$$

We further note that, since $d = 1$ here, the hyperscaling relation, $d\nu = 2 - \alpha$, is also obeyed.

Finally, we observe that expression (33b) for ξ is in conformity with the general result

$$\xi^{-1} = \ln(\lambda_1/\lambda_2), \tag{37}$$

where λ_1 is the largest eigenvalue of the transfer matrix P of the problem and λ_2 the next largest; for a derivation of this result, see Yeomans (1992), Sec. 5.3. In our case, $\lambda_1 = 2\cosh(\beta J)$ and $\lambda_2 = 2\sinh(\beta J)$, see (8), and hence expression (33b) for ξ.

12.2. The n-vector models in one dimension

We now consider a generalization, of the Ising chain, in which the spin variable σ_i is an n-dimensional vector of magnitude unity, whose components can vary *continuously* over the range -1 to $+1$; in contrast, the Ising spin σ_i could have only a *discrete* value, $+1$ or -1. We shall see that the vector models (with $n \geq 2$), while differing quantitatively from one another, differ rather qualitatively from the scalar models (for which $n = 1$). While some of these qualitative differences will show up in the present study, more will become evident in higher dimensions. Here we follow a treatment due to Stanley (1969a,b) who first solved this problem for general n.

Once again we employ an *open* chain composed of N spins constituting $(N - 1)$ nearest-neighbor pairs. The Hamiltonian of the system, in zero field, is given by

$$H_N\{\sigma_i\} = -\sum_{i=1}^{N-1} J_i \sigma_i \cdot \sigma_{i+1} \tag{1}$$

We assume our spins to be classical, so we do not have to worry about the commutation properties of their components. And since the components $\sigma_{i\alpha} (\alpha = 1, \ldots, n)$ of each spin vector σ_i are now continuous variables, the partition function of the system will involve integrations, rather than summations, over these variables. Associating equal *a priori* probabilities with solid angles of equal magnitude in the n-dimensional spin-vector space, we may write

$$Q_N = \int \frac{d\Omega_1}{\Omega(n)} \cdots \frac{d\Omega_N}{\Omega(n)} \prod_{i=1}^{N-1} e^{\beta J_i \sigma_i \cdot \sigma_{i+1}} \quad . \tag{2}$$

where $\Omega(n)$ is the total solid angle in an n-dimensional space; see eqn. (7b) of Appendix C, which gives

$$\Omega(n) = 2\pi^{n/2}/\Gamma(n/2). \tag{3}$$

We first carry out integration over σ_N, keeping the other σ_i fixed. The relevant integral to do is

$$\frac{1}{\Omega(n)}\int e^{\beta J_{N-1}\sigma_{N-1}\cdot\sigma_N}\,d\Omega_N. \tag{4}$$

For σ_N we employ spherical polar coordinates, with polar axis in the direction of σ_{N-1}, while for $d\Omega_N$ we use expression (9) of Appendix C. The integration over angles other than the polar angle θ yields a factor of

$$2\pi^{(n-1)/2}/\Gamma\{(n-1)/2\}. \tag{5}$$

The integral over the polar angle is

$$\int_0^\pi e^{\beta J_{N-1}\cos\theta}\sin^{n-2}\theta\,d\theta = \frac{\pi^{1/2}\Gamma\{(n-1)/2\}}{\left(\frac{1}{2}\beta J_{N-1}\right)^{(n-2)/2}}I_{(n-2)/2}(\beta J_{N-1}), \tag{6}$$

where $I_\mu(x)$ is a modified Bessel function; see Abramowitz and Stegun (1964), formula 9.6.18. Combining (3), (5) and (6), we obtain for (4) the expression

$$\frac{\Gamma(n/2)}{\left(\frac{1}{2}\beta J_{N-1}\right)^{(n-2)/2}}I_{(n-2)/2}(\beta J_{N-1}), \tag{7}$$

regardless of the direction of σ_{N-1}. By iteration, we get

$$Q_N = \prod_{i=1}^{N-1}\frac{\Gamma(n/2)}{\left(\frac{1}{2}\beta J_i\right)^{(n-2)/2}}I_{(n-2)/2)}(\beta J_i); \tag{8}$$

the last integration, over $d\Omega_1$, gave simply a factor of unity.

Expression (8) is valid for *all* n–including $n = 1$, for which it gives: $Q_N = \prod_i \cosh(\beta J_i)$. This last result differs from expression (27) of the preceding section by a factor of 2^N; the reason for this difference lies in the fact that the Q_N of the present study is *normalized* to go to unity as the βJ_i go to zero [see eqn. (2)] whereas the Q_N of the preceding section, being a sum over 2^N discrete states [see eqn. (12.1.24)] goes to 2^N instead. This difference is important in the evaluation of the entropy of the system; it is of no consequence for the calculations that follow.

First of all we observe that the partition function Q_N is analytic at all β–except possibly at $\beta = \infty$ where the singularity of the problem is expected to lie. Thus, no long-range order is expected to appear at any finite temperature T–except at $T = 0$ where, of course, perfect order is supposed to prevail. In view of this, the correlation function for the nearest-neighbor pair (σ_k, σ_{k+1}) is simply $\overline{\sigma_k\cdot\sigma_{k+1}}$ and is given by, see eqns (2) and (8),

$$g_k(n.n.) = \frac{1}{Q_N}\left(\frac{1}{\beta}\frac{\partial}{\partial J_k}\right)Q_N = \frac{I_{n/2}(\beta J_k)}{I_{(n-2)/2}(\beta J_k)}. \tag{9}$$

The internal energy of the system turns out to be

$$U_0 \equiv -\frac{\partial}{\partial\beta}(\ln Q_N) = -\sum_{i=1}^{N-1} J_i \frac{I_{n/2}(\beta J_i)}{I_{(n-2)/2}(\beta J_i)}; \tag{10}$$

not surprisingly, U_0 is simply a sum of the expectation values of the nearest-neighbor interaction terms $-J_i\sigma_i \cdot \sigma_{i+1}$, which is identical with a sum of the quantities $-J_i g_i(n.n.)$ over all nearest-neighbor pairs in the system.

The calculation of $g_k(r)$ is somewhat tricky because of the vector character of the spins, but things are simplified by the fact that we are dealing with a one-dimensional system only. Let us consider the trio of spins σ_k, σ_{k+1} and σ_{k+2}, and suppose for a moment that our spins are 3-dimensional vectors; our aim is to evaluate $\overline{\sigma_k \cdot \sigma_{k+2}}$. We choose spherical polar coordinates with polar axis in the direction of σ_{k+1}; let the direction of σ_k be defined by the angles (θ_0, ϕ_0) and that of σ_{k+2} by (θ_2, ϕ_2). Then

$$\overline{\sigma_k \cdot \sigma_{k+2}} = \overline{\cos\theta(k, k+2)} = \overline{\cos\theta_0\cos\theta_2 + \sin\theta_0\sin\theta_2\cos(\phi_0 - \phi_2)}. \tag{11}$$

Now, with σ_{k+1} fixed, spins σ_k and σ_{k+2} will orient themselves *independently of one another* because, apart from σ_{k+1}, there is no other channel of interaction between them. Thus, the pair of angles (θ_0, ϕ_0) and the pair (θ_2, ϕ_2) vary independently of one another; this makes $\overline{\cos(\phi_0 - \phi_2)} = 0$ and $\overline{\cos\theta_0\cos\theta_2} = \overline{\cos\theta_0}\,\overline{\cos\theta_2}$. It follows that

$$\overline{\sigma_k \cdot \sigma_{k+2}} = \overline{\sigma_k \cdot \sigma_{k+1}}\;\overline{\sigma_{k+1} \cdot \sigma_{k+2}}. \tag{12}$$

Extending this argument to general n and to a segment of length r, we get

$$g_k(r) = \prod_{i=k}^{k+r-1} g_i(\text{n.n.}) = \prod_{i=k}^{k+r-1} I_{n/2}(\beta J_i)/I_{(n-2)/2}(\beta J_i). \tag{13}$$

With a common J, eqns (9), (10) and (13) take the form

$$g(\text{n.n.}) = I_{n/2}(\beta J)/I_{(n-2)/2}(\beta J), \tag{14}$$

$$U_0 = -(N-1)JI_{n/2}(\beta J)/I_{(n-2)/2}(\beta J) \tag{15}$$

and

$$g(r) = \{I_{n/2}(\beta J)/I_{(n-2)/2}(\beta J)\}^r. \tag{16}$$

The last result here may be written in the standard form $e^{-r/\xi}$, with

$$\xi = [\ln\{I_{(n-2)/2}(\beta J)/I_{n/2}(\beta J)\}]^{-1} \tag{17}$$

For $n = 1$, we have: $I_{1/2}(x)/I_{-1/2}(x) = \tanh x$; the results of the preceding section are then correctly recovered.

For a study of the low-temperature behavior, where $\beta J \gg 1$, we invoke the asymptotic expansion

$$I_\mu(x) = \frac{e^x}{\sqrt{(2\pi x)}}\left[1 - \frac{4\mu^2 - 1}{8x} + \cdots\right] \qquad \gg 1), \tag{18}$$

with the result that

$$g(\text{n.n.}) \approx 1 - \frac{n-1}{2\beta J}, \tag{14a}$$

$$U_0 \approx -(N-1)J\left[1 - \frac{n-1}{2\beta J}\right] \tag{15a}$$

and

$$\xi \approx \frac{2\beta J}{n-1} \sim T^{-1} \tag{17a}$$

Clearly, the foregoing results hold only for $n \geq 2$; for $n = 1$, the asymptotic expansion (18) is no good because it yields the same result for $\mu = \frac{1}{2}$ as for $\mu = -\frac{1}{2}$. In that case one is obliged to use the closed form result, $g(n.n.) = \tanh(\beta J)$, which for $\beta J \gg 1$ gives

$$g(\text{n.n.}) \approx 1 - 2e^{-2\beta J} \tag{14b}$$

$$U_0 \approx -(N-1)J[1 - 2e^{-2\beta J}] \tag{15b}$$

and

$$\xi \approx \tfrac{1}{2}e^{2\beta J}, \tag{17b}$$

in complete agreement with the results of the preceding section. For completeness, we write down for the low-temperature specific heat of the system

$$C_0 \approx (N-1)\begin{cases} \frac{1}{2}(n-1)k & \text{for} \quad n \geq 2 & (19a) \\ 4k\,(\beta J)^2 e^{-2\beta J} & \text{for} \quad n = 1. & (19b) \end{cases}$$

The most obvious distinction between one-dimensional models with *continuous* symmetry ($n \geq 2$) and those with *discrete* symmetry ($n = 1$) is in relation to the nature of the singularity at $T = 0$. While in the case of the former it is a power-law singularity, with critical exponents[1]

$$\alpha = 1, \quad \nu = 1, \quad \eta = 1, \quad \text{and hence} \quad \gamma = 1, \tag{20}$$

in the case of the latter it is an exponential singularity. Nevertheless, by introducing the temperature parameter $t = e^{-p\beta J}$, see eqn. (12.1.17), we converted this exponential singularity in T into a power-law singularity in t, with

$$\alpha = 2 - 2/p, \quad \gamma = \nu = 2/p, \quad \eta = 1. \tag{21}$$

However, the inherent arbitrariness in the choice of the number p left an ambiguity in the values of these exponents; we now see that by choosing $p = 2$ we can bring exponents (21) in line with (20).

Next we observe that the critical exponents (20) for $n \geq 2$ turn out to be independent of n—a feature that seems peculiar to situations where $T_c = 0$. In higher dimensions, where T_c is finite, the critical exponents do vary with n; for details, see Sec. 12.6. In any case, the amplitudes always depend on n. In this connection we note that, since each of the N spins comprising the system has n components, the total number of degrees of freedom in this problem is Nn. It seems appropriate that the extensive quantities, such as U_0 and C_0, be divided by Nn, so that they are expressed as *per degree of freedom*. A look at eqn. (15a) now tells us that our parameter J must be of the form nJ', so that in the thermodynamic limit

$$\frac{U_0}{Nn} \approx -J' + \frac{n-1}{2n}kT \tag{22}$$

and, accordingly,

$$\frac{C_0}{Nn} \approx \frac{n-1}{2n}k. \tag{23}$$

Equation (17a) then becomes

$$\xi \approx \frac{2n}{n-1}\frac{J'}{kT}. \tag{24}$$

Note that the amplitudes appearing in eqns (22)–(24) are such that the limit $n \to \infty$ exists; this limit pertains to the so-called *spherical model*, which will be studied in Sec. 12.4.

Figure 12.4 shows the normalized energy $u_0 (= U_0/NnJ')$ as a function of temperature for several values of n, including the limiting case $n = 1$. We note that u_0 [which, in the thermodynamic limit, is equal and opposite to the nearest-neighbor correlation function $g(\text{n.n.})$] increases monotonically with n–implying that $g(\text{n.n.})$, and hence $g(r)$, decrease monotonically as n increases. This is consistent with the fact that the correlation length ξ also decreases as n increases; see eqn. (24). The physical reason for this behavior is that an increase in the number of degrees of freedom available to each spin in the system effectively diminishes the degree of correlation among any two of them.

Another feature emerges here that distinguishes vector models ($n \geq 2$) from the scalar model ($n = 1$); this is related to the manner in which the quantity u_0 approaches its ground-state value -1. While for $n = 1$, the approach is quite slow–leading to a vanishing specific heat, see eqns (15b) and (19b)–for $n \geq 2$, the approach is essentially linear in T, leading to a *finite* specific heat; see eqns (15a) and (19a). This last result violates the third law of thermodynamics, according to which the specific heat of a real system must go to zero as $T \to 0$. The resolution of this dilemma lies in the fact that the low-lying states of a system with *continuous* symmetry ($n \geq 2$) are dominated by long-wavelength excitations, known as *Goldstone modes*, which in the case of a magnetic system assume the form of "spin waves", characterized by a particle-like spectrum: $\omega(k) \sim k^2$. The very low-temperature behavior of the system is primarily governed by these modes, and the thermal energy associated with them is given by

$$U_{\text{thermal}} \sim \int \frac{\hbar\omega}{\exp(\hbar\omega/k_B T) - 1} k^{d-1}\,dk \sim T^{(d+2)/2}; \tag{25}$$

this results in a specific heat $\sim T^{d/2}$, which indeed is consistent with the third law. For a general account of the Goldstone excitations, see Huang (1987); for their role as "spin waves" in a magnetic system, see Plischke and Bergersen (1989).

For further information on one-dimensional models, see Lieb and Mattis (1966), and Thompson (1972a, b).

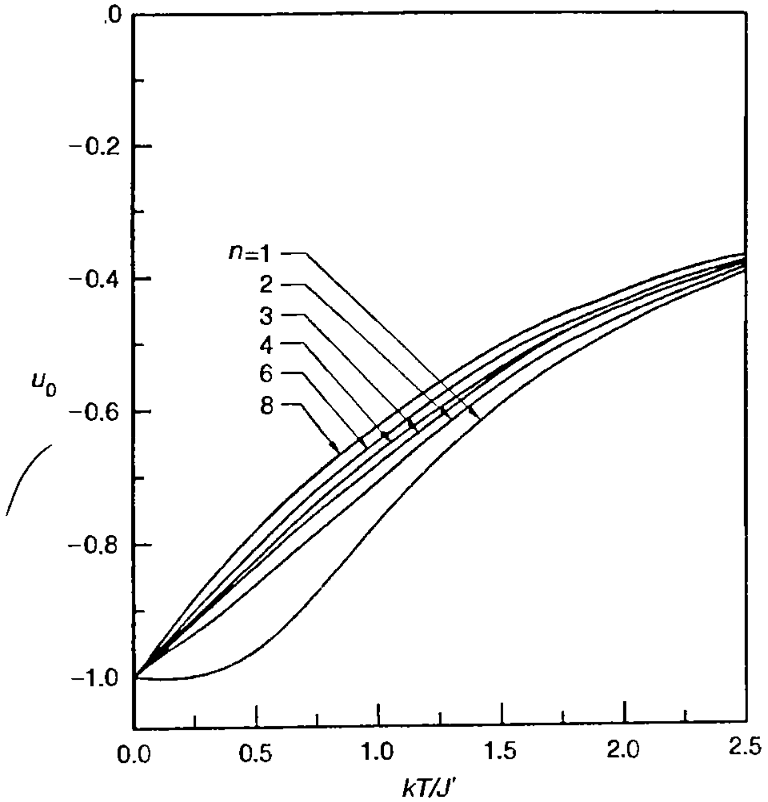


FIG. 12.4. The normalized energv $u_0(= U_0/NnJ')$ of a one-dimensional chain as a function of the temperature parameter kT/J' for several values of n (after Stanley, 1969a,b).

12.3. The Ising model in two dimensions

As stated earlier, Ising (1925) himself carried out a combinatorial analysis of the one-dimensional model and found that there was no phase transition at a finite temperature T. This led him to conclude, erroneously though, that his model would not exhibit a phase transition in higher dimensions either. In fact, it was this "supposed failure" of the Ising model that motivated Heisenberg to develop, in 1928, the theory of ferromagnetism based on a more sophisticated interaction among the spins; compare the Heisenberg interaction (11.3.6) with the Ising interaction (11.3.7). It was only after some exploitation of the Heisenberg model that people returned to investigate the properties of the Ising model.

The first exact, quantitative result for the two-dimensional Ising model was obtained by Kramers and Wannier (1941) who successfully located the critical temperature of the system. They were followed by Onsager (1944) who derived an explicit expression for the free energy in zero field and thereby established the precise nature of the specific-heat singularity. These authors employed the *transfer matrix method* which was introduced in Sec. 12.1 to solve the corresponding one-dimensional problem; its application to the two-dimensional model, even in the absence of the field, turned out to be an extremely difficult task. Although some of these difficulties were softened by subsequent treatments due to Kaufman (1949) and to Kaufman and Onsager (1949), it seemed very natural to look for simpler approaches.

One such approach was developed by Kac and Ward (1952), later refined by Potts and Ward (1955), in which combinatorial arguments were used to express the partition function of the system as the determinant of a certain matrix A. This method throws special light on the "topological conditions" which give rise to an exact solution in two dimensions and which are clearly absent in three dimensions; a particularly lucid account of this method has been given by Baker (1990). In 1960 Hurst and Green introduced yet another approach to investigate the Ising problem; this involved the use of "triangular arrays of quantities closely related to antisymmetric determinants" and became rightly known as the method of Pfaffians. This method applies rather naturally to the study of "configurations of dimer molecules on a given lattice" which, in turn, is closely related to the Ising problem; for details, see Kasteleyn (1963), Montroll (1964), and Green and Hurst (1964). A pedagogical account of the approach through Pfaffians is given in Thompson (1972b), where a comprehensive treatment of the original, algebraic approach can also be found. Another combinatorial solution, which is generally regarded as the simplest, was obtained by Vdovichenko (1965) and by Glasser (1970), and is readily accessible in Stanley (1971). For an exhaustive account of the two-dimensional Ising model, see McCoy and Wu (1973).

We analyze this problem with the help of a combinatorial approach assisted, from time to time, by a graphical representation. The zero-field partition function of the system is given by the familiar expression

$$Q(N, T) = \sum_{\{\sigma_i\}} \prod_{\text{n.n.}} e^{K\sigma_i\sigma_j} \quad (K = J/kT). \tag{1}$$

Our first step consists in carrying out a high-temperature and a low-temperature expansion of the partition function and establishing an intimate relation between the two.

(i) *High-temperature expansion:* Since the product $(\sigma_i\sigma_j)$ can only be $+1$ or -1, we may write

$$e^{K\sigma_i\sigma_j} = \cosh K + \sigma_i\sigma_j \sinh K = \cosh K(1 + \sigma_i\sigma_j v), \qquad v = \tanh K. \tag{2}$$

The product over all nearest-neighbor pairs then takes the form

$$\prod_{\text{n.n.}} e^{K\sigma_i\sigma_j} = (\cosh K)^{\mathcal{N}} \prod_{\text{n.n.}} (1 + \sigma_i\sigma_j v), \tag{3}$$

$\mathcal{N}$ being the total number of nearest-neighbor pairs on the lattice; for a lattice with periodic boundary conditions, $\mathcal{N} = \frac{1}{2}qN$ where q is the coordination number. The partition function may then be written as

$$Q(N, T) = (\cosh K)^{\mathcal{N}} \sum_{\sigma_1=\pm1} \cdots \sum_{\sigma_N=\pm1} \left[1 + v \sum_{(i,j)} \sigma_i\sigma_j \right.$$
$$\left. + v^2 \sum_{(i,j)} \sum_{\substack{(k,l) \\ (i,j)\neq(k,l)}} \sigma_i\sigma_j\sigma_k\sigma_l + \ldots \right]. \tag{4}$$

Now we represent each product $(\sigma_i\sigma_j)$ appearing in (4) by a "bond connecting sites i and j on the given lattice"; then, each coefficient of v^r in the expansion would be represented by a "graph consisting of r different bonds on the lattice". Figure 12.5 shows all possible graphs, with $r = 1$ and 2, on a square lattice. Notice that in each case we have some of the σ_i appearing *only once* in the term, which makes all these terms vanish on summation over $\{\sigma_i\}$. The same is true for $r = 3$. Only when we reach $r = 4$ do we receive a nonvanishing contribution from terms of the type $(\sigma_i\sigma_j\sigma_j\sigma_k\sigma_k\sigma_l\sigma_l\sigma_i) \equiv 1$ which, on summation over $\{\sigma_i\}$, yield a contribution of 2^N each. It is obvious that a nonvanishing term corresponds to a graph in which each vertex is met by an *even* number of bonds–making the graph necessarily a closed one; see Fig. 12.6, where some other closed graphs are also shown. In view of these observations, expression (4) may be written as

$$Q(N,T) = 2^N(\cosh K)^{\text{[illegible]}} \sum_r n(r) v^r \quad [n(0) = 1], \tag{5}$$

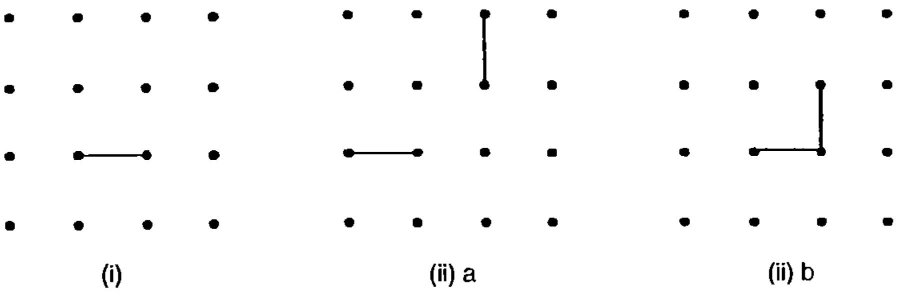

(i) (ii) a (ii) b

FIG. 12.5. Graphs with $r = 1$ and $r = 2$ on a square lattice

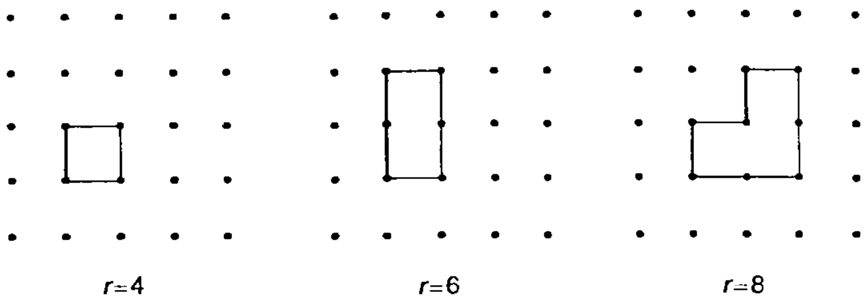

$r=4$ $r=6$ $r=8$

FIG. 12.6. Examples of closed graphs with r bonds on a square lattice.

where $n(r)$ is the number of graphs that can be drawn on the given lattice using r bonds such that each vertex of the graph is met by an even number of bonds. For simplicity, we shall refer to these graphs as *closed graphs*. Our problem thus reduces to one of enumerating such graphs on the given lattice.

Since $v = \tanh(J/kT)$, the higher the temperature the smaller the v. Expansion (5) is, therefore, particularly useful at higher temperatures (even though it is exact for all T). As illustration, we apply this result to a one-dimensional Ising chain. In the case of an *open* chain, no closed graphs are possible, so all we get from

(5) is the term with $r = 0$; with $\mathcal{N} = N - 1$, this g

$$Q(N,T) = 2^N (\cosh K)^{N-1}, \tag{6}$$

which agrees with our previous result (12.1.27). In the case of a *closed* chain, we do have a closed graph–the one with $r = N$; we now get (with $\mathcal{N} = N$)

$$Q(N,T) = 2^N (\cosh K)^N [1 + v^N] = 2^N [(\cosh K)^N + (\sinh K)^N], \tag{7}$$

which agrees with expression (12.1.5), with $(\lambda_1)_{B=0} = 2\cosh K$ and $(\lambda_2)_{B=0} = 2\sinh K$.

(ii) *Low-temperature expansion*: We start with the observation that the ground state of the system consists of all spins aligned in the same direction, with the total energy $E_0 = -J\mathcal{N}$. As one spin is flipped, q *unlike* nearest-neighbor pairs are created at the expense of *like* ones, and the energy of the system increases by an amount $2qJ$. It seems appropriate, therefore, that the Hamiltonian of the system be written in terms of the number, N_{+-}, of unlike nearest-neighbor pairs in the lattice, i.e.

$$H(N_{+-}) = -J(N_{++} + N_{--} - N_{+-}) = -J(\mathcal{N} - 2N_{+-}). \tag{8}$$

The partition function of the system may then be written as

$$Q(N,T) = e^{K\mathcal{N}} \sum_r m(r) e^{-2Kr} \quad [m(0) = 1], \tag{9}$$

where $m(r)$ denotes the "number of distinct ways in which the N spins of the lattice can be so arranged as to yield r unlike nearest-neighbor pairs". It is obvious that the first nonzero term in (9), after the one with $r = 0$, would be the one with $r = q$.

A graphical representation of the number $m(r)$ is straightforward. Referring to Fig. 12.7, which pertains to a square lattice, we see that each term in expansion (9) can be associated with a closed graph that cordons off region(s) of "up" spins from those of "down" spins, the perimeter of the graph being precisely the number of unlike nearest-neighbor pairs in the lattice for that particular configuration. Our problem then reduces to one of enumerating closed graphs, of appropriate perimeters, that can be drawn on the given lattice.

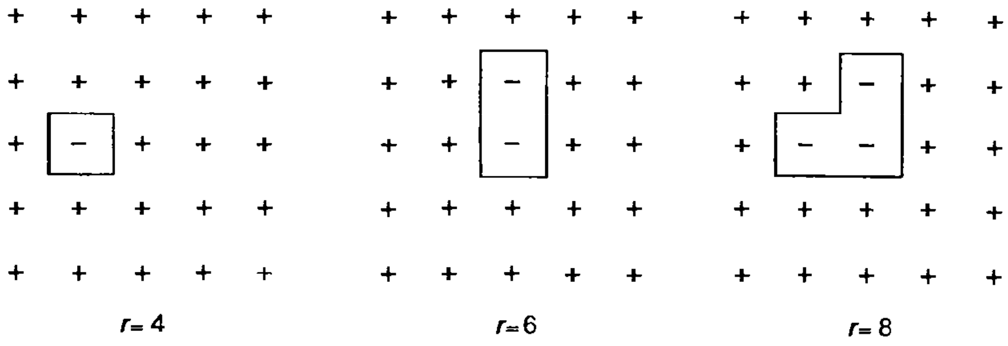


FIG. 12.7. Graphs cordoning off regions of "up" spins from those of "down" spins, with r *unlike* nearest-neighbor pairs.

Now, since expansion (9) is a power series in the variable e^{-2K} which increases as T increases, this expansion is particularly useful at lower temperatures (even though it is exact for all T). We shall now establish an important relation between the coefficients appearing in expansion (5) and the ones appearing in (9).

(iii) *The duality transformation*: To establish the desired relation we construct a lattice *dual* to the one given. By definition, we draw right bisectors of all the bonds in the lattice, so that the points of intersection of these bisectors become the sites of the new lattice. The resulting lattice may not be similar in structure to the one we started with; for instance, while the dual of a square lattice is itself a square lattice, the dual of a triangular lattice ($q = 6$) is, in fact, a honeycomb lattice ($q = 3$), and vice versa; see Figs 12.8 and 12.9. The argument now runs as follows:

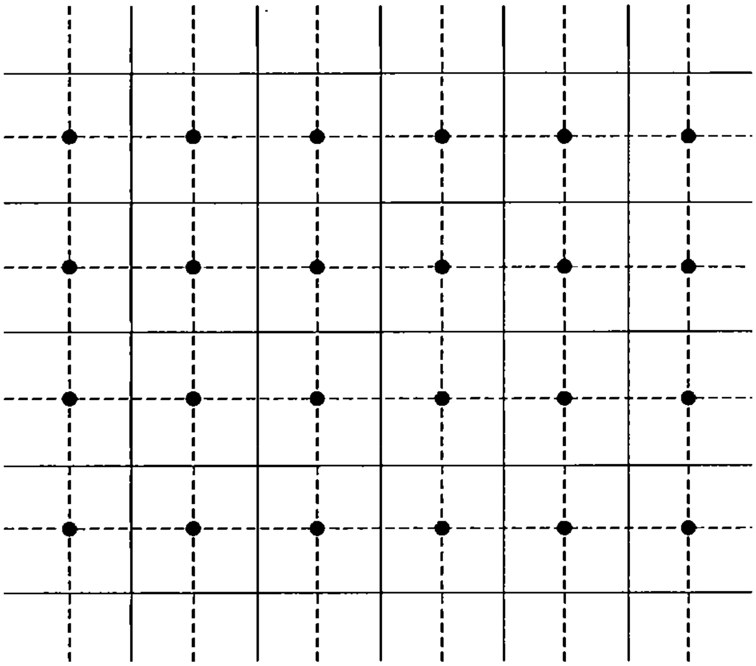


FIG. 12.8. A square lattice and its dual (which is also square).

We start with the given lattice on which one of the $n(r)$ closed graphs, with r bonds, is drawn and construct the lattice dual to this one, placing spins of one sign on the sites inside this graph and spins of opposite sign on the sites outside. Then this graph represents precisely a configuration with r unlike nearest-neighbor pairs in the dual lattice and hence qualifies to be counted as one of the $m(r)$ graphs on the dual lattice. Conversely, if we start with one of the $m(r)$ graphs, of perimeter r, representing a configuration with r unlike nearest-neighbor pairs in the original lattice and go through the process of constructing the dual lattice, then this graph will qualify to be one of the $n(r)$ closed graphs, with r bonds, on the dual lattice. In fact, there is a *one-to-one* correspondence between graphs of one kind on the given lattice and graphs of the other kind on the dual lattice; cf. Figs 12.6 and 12.7. It follows that

$$n(r) = m_D(r) \quad \text{and} \quad m(r) = n_D(r), \tag{10a, b}$$

where the suffix D refers to the dual lattice.

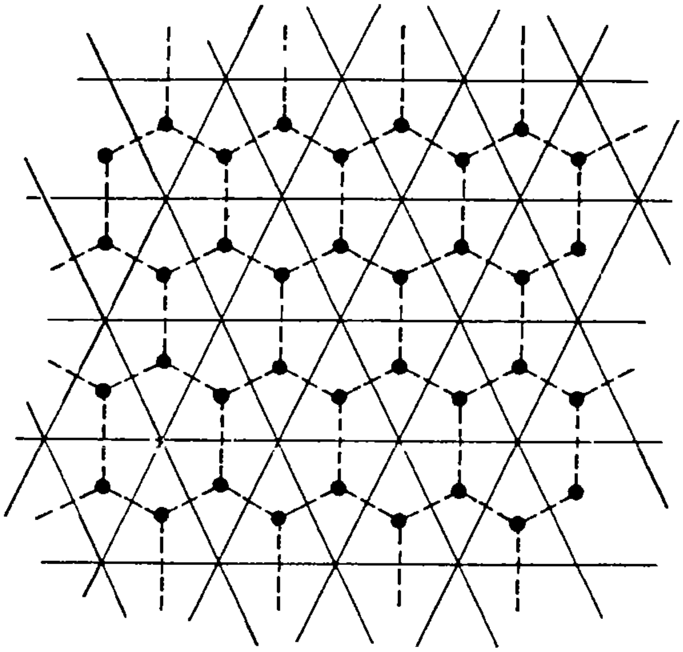


FIG. 12.9. A honeycomb lattice ($q = 3$) and its dual, which is triangular ($q = 6$).

With relations (10) established, we go back to eqn. (9) and introduce another temperature variable $K^*(= J/kT^*)$ such that

$$\tanh K^* = e^{-2K}; \tag{11}$$

note that eqn. (11) can also be written in the symmetrical form

$$\sinh(2K)\sinh(2K^*) = 1. \tag{12}$$

Substituting (10b) and (11) into (9), we get

$$Q(N, T) = e^{K\mathcal{N}} \sum_r n_D(r) v^{*r}, \quad v^* = \tanh K^*. \tag{13}$$

At the same time we apply eqn. (5) to the *dual* lattice at temperature T^*, to get

$$Q_D(N_D, T^*) = 2^{N_D}(\cosh K^*)^{\mathcal{N}} \sum_r n_D(r) v^{*r}, \tag{14}$$

where $N_D = qN/q_D$; see again Fig. 12.9. Comparing (13) and (14), we arrive at the desired relation

$$Q(N, T) = 2^{-N_D}(\sinh K^* \cosh K^*)^{-\mathcal{N}/2} Q_D(N_D, T^*) \tag{15}$$

which relates the partition function of the given lattice at temperature T to that of the dual lattice at temperature T^*. Equation (15) constitutes the so-called *duality transformation*.

(iv) *Location of the critical point*: For a square lattice, which is self-dual, there should be no distinction between Q and Q_D. With $q = 4$, and hence $\mathcal{N} = 2N$, eqn. (15) becomes

$$Q(N, T) = [\sinh(2K^*)]^{-N} Q(N, T^*), \tag{16}$$

which may also be written as

$$[\sinh(2K)]^{-N/2}Q(N,T) = [\sinh(2K^*)]^{-N/2}Q(N,T^*). \tag{17}$$

It will be noted from eqn. (11) or (12) that as $T \to \infty$, $T^* \to 0$ and as $T \to 0$, $T^* \to \infty$; eqn. (17), therefore. establishes a one-to-one correspondence between the high-temperature and the low-temperature values of the partition function of the lattice. It then follows that if there exists a singularity in the partition function at a particular temperature T_c, there must exist an equivalent singularity at the corresponding temperature T_c^*. And in case we have only one singularity, as indeed follows from one of the theorems of Yang and Lee (1952), it must exist at a temperature T_c such that $T_c^* = T_c$. The critical temperature of the square lattice is, therefore, given by the equation, see formula (12),

$$\sinh(2K_c) = 1, \tag{18}$$

whence

$$K_c = \tfrac{1}{2}\sinh^{-1} 1 = \tfrac{1}{2}\ln(\surd 2 + 1) = \tfrac{1}{2}\ln\cot(\pi/8) \simeq 0.4407. \tag{19}$$

For comparison, we note that for the same lattice the Bragg–Williams approximation gave $K_c = 0.25$ while the Bethe approximation gave $K_c = \frac{1}{2}\ln 2 \simeq 0.3466$.

The situation for other lattices such as the triangular or the honeycomb, which are *not* self-dual, is complicated by the fact that the functions Q and Q_D in eqn. (15) are not the same. One then needs another trick–the so-called *star–triangle transformation*–which was first alluded to by Onsager (1944) in his famous paper on the solution of the square lattice problem but was written down explicitly by Wannier (1945); for details, see Baxter (1982). Unlike the duality transformation, this one establishes a relation between a *high*-temperature model on the triangular lattice and again a *high*-temperature model on the honeycomb lattice, and so on. Combining the two transformations, one can eliminate the dual lattice altogether and obtain a relation between a high-temperature and a low-temperature model on the same lattice. The location of the critical point is then straightforward; one obtains for the triangular lattice ($q = 6$)

$$K_c = \frac{1}{2}\sinh^{-1}\frac{1}{\surd 3} \simeq 0.2747, \tag{20}$$

and for the honeycomb lattice ($q = 3$)

$$K_c = \tfrac{1}{2}\sinh^{-1}\surd 3 \simeq 0.6585. \tag{21}$$

The numerical values of K_c, given by eqns (19)–(21), reinforce the fact that higher coordination numbers help propagate long-range order in the system more effectively and hence raise the critical temperature T_c.

(v) *The partition function and the specific-heat singularity*: The partition function of the Ising model on a square lattice is given by, see references cited at the beginning of this section,

$$\frac{1}{N}\ln Q(T) = \ln\{2^{1/2}\cosh(2K)\} + \frac{1}{\pi}\int_0^{\pi/2} d\phi \ln\{1 + \surd(1-\kappa^2\sin^2\phi)\}, \tag{22}$$

where

$$\kappa = 2\sinh(2K)/\cosh^2(2K). \tag{23}$$

Differentiating (22) with respect to $-\beta$, one obtains for the internal energy per spin

$$\frac{1}{N}U_0(T) = -2J\tanh(2K) + \frac{1}{\pi}\left(\kappa\frac{d\kappa}{d\beta}\right) \times$$

$$\int_0^{\pi/2} d\phi \frac{\sin^2\phi}{\{1+\surd(1-\kappa^2\sin^2\phi)\}\surd(1-\kappa^2\sin^2\phi)}. \tag{24}$$

Rationalizing the integrand, the integral in (24) can be written as

$$\frac{1}{\kappa^2}\left\{-\frac{\pi}{2}+K_1(\kappa)\right\}, \tag{25}$$

where $K_1(\kappa)$ is the complete elliptic integral of the *first* kind, κ being the modulus of the integral:

$$K_1(\kappa) = \int_0^{\pi/2} \frac{d\phi}{\surd(1-\kappa^2\sin^2\phi)}. \tag{26}$$

Now, a logarithmic differentiation of (23) with respect to β gives

$$\frac{1}{\kappa}\frac{d\kappa}{d\beta} = 2J\{\coth(2K) - 2\tanh(2K)\}. \tag{27}$$

Substituting these results into (24), we obtain

$$\frac{1}{N}U_0(T) = -J\coth(2K)\left\{1+\frac{2\kappa'}{\pi}K_1(\kappa)\right\}, \tag{28}$$

where κ' is the *complementary* modulus:

$$\kappa' = 2\tanh^2(2K) - 1 \quad (\kappa^2+\kappa'^2=1). \tag{29}$$

Figure 12.10 shows the variation of the moduli κ and κ' with the temperature parameter $(kT/J) = K^{-1}$. We note that, while κ is always positive, κ' can be positive or negative; actually, κ lies between 0 and 1 while κ' lies between -1 and 1. At the critical point, where $\sinh(2K_c) = 1$ and hence $K_c^{-1} \simeq 2.269$, the moduli κ and κ' are equal to 1 and 0, respectively.

To determine the specific heat of the lattice, we differentiate (28) with respect to temperature. In doing so, we make use of the following results:

$$\frac{d\kappa}{d\beta} = -\frac{\kappa'}{\kappa}\frac{d\kappa'}{d\beta}, \quad \frac{d\kappa'}{d\beta} = 8J\tanh(2K)\{1-\tanh^2(2K)\} \tag{30}$$

and

$$\frac{dK_1(\kappa)}{d\kappa} = \frac{1}{\kappa'^2\kappa}\{E_1(\kappa) - \kappa'^2K_1(\kappa)\}, \tag{31}$$

where $E_1(\kappa)$ is the complete elliptic integral of the *second* kind:

$$E_1(\kappa) = \int_0^{\pi/2} \surd(1-\kappa^2\sin^2\phi)\,d\phi. \tag{32}$$

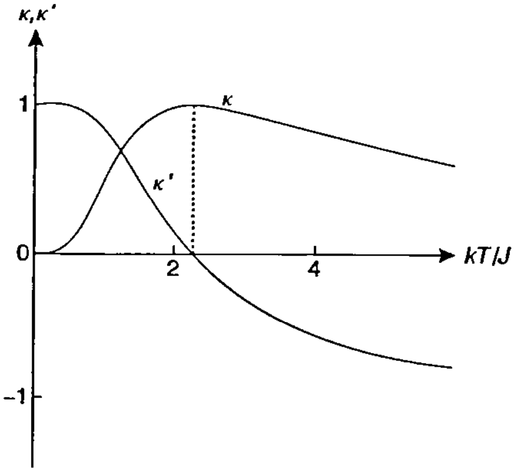


FIG. 12.10. Variation of the moduli κ and κ' with (kT/J).

We finally obtain

$$\frac{1}{Nk}C_0(T) = \frac{2}{\pi}\{K\coth(2K)\}^2\left[2\{K_1(\kappa)-E_1(\kappa)\} - (1-\kappa')\left\{\frac{\pi}{2}+\kappa' K_1(\kappa)\right\}\right]. \tag{33}$$

Now, the elliptic integral $K_1(\kappa)$ has a singularity at $\kappa = 1$ (i.e. at $\kappa' = 0$), in the neighborhood of which

$$K_1(\kappa) \approx \ln\{4/|\kappa'|\} \quad \text{and} \quad E_1(\kappa) \approx 1. \tag{34}$$

Accordingly, the specific heat of the lattice displays a *logarithmic singularity* at a temperature T_c, given by the condition: $\kappa_c = 1$ (or $\kappa'_c = 0$), which is identical with (18). In the vicinity of the critical point, eqn. (33) reduces to

$$\frac{1}{Nk}C_0(T) \simeq \frac{8}{\pi}K_c^2\left[\ln\{4/|\kappa'|\} - \left(1+\frac{\pi}{4}\right)\right]; \tag{35}$$

at the same time, the parameter κ' reduces to, see eqn. (30),

$$\kappa' \simeq 2\sqrt{2}K_c\left(1-\frac{T}{T_c}\right). \tag{36}$$

The specific heat singularity is, therefore, given by

$$\begin{aligned}\frac{1}{Nk}C_0(T) &\simeq \frac{8}{\pi}K_c^2\left[-\ln\left|1-\frac{T}{T_c}\right| + \left\{\ln\left(\frac{\sqrt{2}}{K_c}\right) - \left(1+\frac{\pi}{4}\right)\right\}\right] \\ &\simeq -0.4945\ln\left|1-\frac{T}{T_c}\right| + \text{const.},\end{aligned} \tag{37}$$

signalling a logarithmic divergence at $T = T_c$.

Figures 12.11 and 12.12 show the temperature dependence of the internal energy and the specific heat of the square lattice, as given by the Onsager expressions (28) and (33); for comparison, the results of the Bragg–Williams approximation and of the Bethe approximation (with $q = 4$) are also included. The specific-heat singularity, given correctly by the Onsager expression (37), is seen as a (logarithmic)

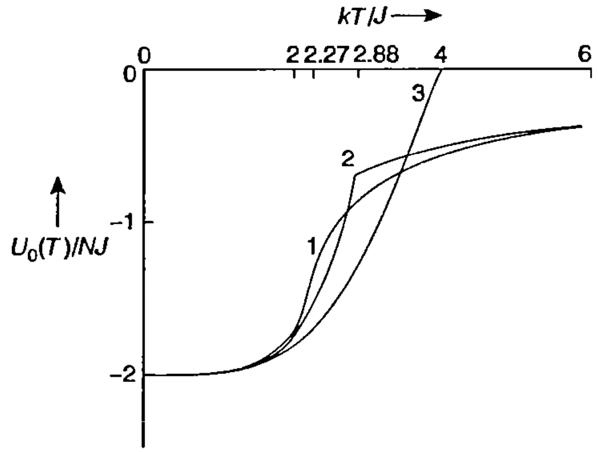


FIG. 12.11. The internal energy of a square lattice ($q = 4$) according to (1) the Onsager solution, (2) the Bethe approximation and (3) the Bragg–Williams approximation.

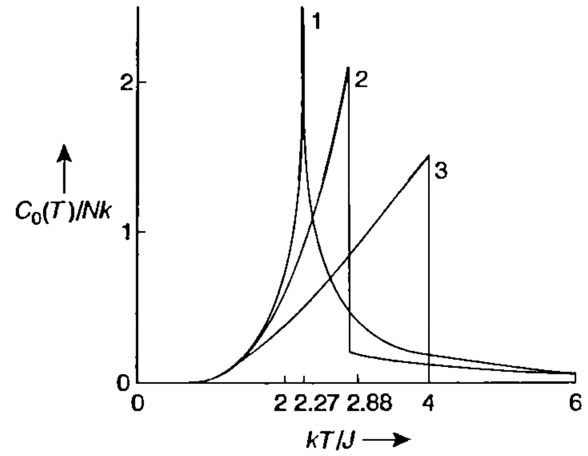


FIG. 12.12. The specific heat of a square lattice ($q = 4$) according to (1) the Onsager solution, (2) the Bethe approximation and (3) the Bragg–Williams approximation.

peak in Fig. 12.12, which differs markedly from the jump discontinuity predicted by the mean field theory. We conclude that the critical exponent $\alpha = \alpha' = 0$ (log).

In passing, we note that the internal energy of the lattice is *continuous* at the critical point, having a value of $-\sqrt{2}J$ per spin and an infinite, positive slope; needless to say, the continuity of the internal energy implies that the transition takes place without any latent heat.

(vi) *Other properties:* We now consider the temperature dependence of the *order parameter*, i.e. the spontaneous magnetization, of the lattice. An exact expression for this quantity was first derived by Onsager (1949), though he never published the details of his derivation. The first published derivation is due to Yang (1952), who showed that

$$\overline{L}_0(T) \equiv \frac{1}{N\mu}\overline{M}(0,T) = \begin{cases} [1 - \{\sinh(2K)\}^{-4}]^{1/8} & \text{for } T \leq T_c \quad (38a) \\ 0 & \text{for } T \geq T_c, \quad (38b) \end{cases}$$

where K, as usual, is J/kT. In the limit $T \to 0$,

$$\bar{L}_0(T) \simeq 1 - 2\exp(-8J/kT), \tag{39}$$

which implies a very slow variation with T. On the other hand, in the limit $T \to T_c-$,

$$\bar{L}_0(T) \approx \left\{8\sqrt{2}K_c\left(1 - \frac{T}{T_c}\right)\right\}^{1/8} \simeq 1.2224\left(1 - \frac{T}{T_c}\right)^{1/8} \tag{40}$$

which indicates a very fast variation with T. The detailed dependence of $\bar{L}_0$ on T is shown in Fig. 12.13; for comparison, the results of the Bragg–Williams approximation and the Bethe approximation are also included. We infer that the critical exponent β for this model is $\frac{1}{8}$, which is so different from the mean field value of $\frac{1}{2}$.

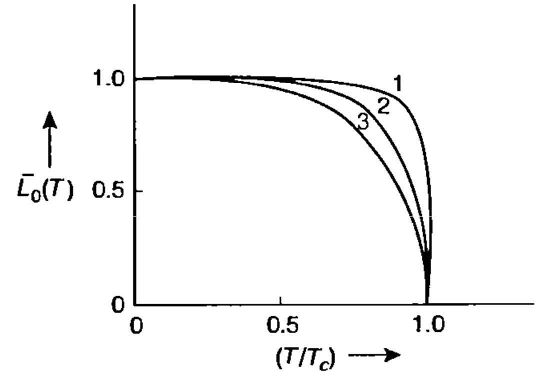


FIG. 12.13. The spontaneous magnetization of a square lattice ($q = 4$) according to (1) the Onsager solution, (2) the Bethe approximation and (3) the Bragg–Williams approximation.

Onsager also calculated the correlation length ξ of the lattice which showed that the critical exponent $\nu = \nu' = 1$ — in sharp contrast with the classical value of $\frac{1}{2}$. Finally, he set up calculations for the correlation function $g(r)$ from which one could infer that the exponent $\eta = \frac{1}{4}$, again in disagreement with the classical value of zero. Precise asymptotic expressions for the correlation function in different regimes of temperature were derived by later authors (Fisher, 1959; Kadanoff, 1966a; Au-Yang and Perk, 1984):

$$g(r) \approx \frac{\{4(K_c - K)\}^{1/4}}{2^{3/8}(\pi r/\xi)^{1/2}} e^{-r/\xi}, \qquad \xi = \{4(K_c - K)\}^{-1} \tag{41}$$

for $T > T_c$,

$$g(r) \approx \frac{\{4(K - K_c)\}^{1/4}}{2^{21/8}\pi(r/\xi)^2} e^{-2r/\xi}, \qquad \xi = \{4(K - K_c)\}^{-1} \tag{42}$$

for $T < T_c$, and

$$g(r) \approx \frac{2^{1/12}\exp\{3\zeta'(-1)\}}{r^{1/4}} \tag{43}$$

at $T = T_c$; in the last expression, $\zeta'(x)$ denotes the derivative of the Riemann zeta function. We note from expressions (41) and (42) that the correlation length ξ, while diverging at $T = T_c$, is finite on *both* sides of the critical point. This feature is peculiar to the scalar model ($n = 1$) only, for in the case of vector models ($n \geq 2$), ξ turns out to be infinite at all $T \leq T_c$.

The zero-field susceptibility of this system has also been worked out (see Barouch *et al.*, 1973; Tracy and McCoy, 1973; Wu *et al.*, 1976); asymptotically, one finds that

$$\chi_0 \approx \frac{N\mu^2}{kT_c} \times \begin{cases} C_+ t^{-7/4} & \text{for} \quad t \gtrsim 0 \qquad (44a) \\ C_- |t|^{-7/4} & \text{for} \quad t \lesssim 0, \qquad (44b) \end{cases}$$

where t, as usual, is $(T - T_c)/T_c$ while the constants C_+ and C_- are about 0.96258 and 0.02554, respectively. We see that the critical exponent $\gamma = \gamma' = \frac{7}{4}$, as opposed to the mean field value of 1, and the ratio $C_+/C_- \simeq 37.69$, as opposed to the mean field value of 2. Assembling all the exponents at one place, we have for the two-dimensional Ising model

$$\alpha = \alpha' = 0\,(\log), \quad \beta = \tfrac{1}{8}, \quad \gamma = \gamma' = \tfrac{7}{4}, \quad \nu = \nu' = 1, \quad \eta = \tfrac{1}{4}. \qquad (45)$$

Since this model has not yet been solved in the presence of a field, a direct evaluation of the exponent δ has not been carried out. Assuming the validity of the scaling relations, however, we can safely conclude that $\delta = 15$–again very different from the classical value of 3. All in all, the results of this section tell us very clearly, and loudly, how inadequate the mean field theory can be.

Before we close this section a few remarks seem to be in order. First of all, it may be mentioned that for the model under consideration one can also calculate the *interfacial tension* s, which may be defined as the "free energy associated, per unit area, with the interfaces between the domains of *up* spins and those of *down* spins"; in our analogy with the gas–liquid systems, this corresponds to the conventional surface tension σ. The corresponding exponent μ, that determines the manner in which $s \to 0$ as $T \to T_c-$, turns out to be 1 for this model; see Baxter (1982). This indeed obeys the scaling relation $\mu = (d - 1)\nu$, as stated in Problem 11.26. Secondly, we would like to point out that, while the solution of the two-dimensional Ising model was the first exact treatment that exposed the inadequacy of the mean field theory, it was also the first to disclose the underlying universality of the problem. As discovered by Onsager himself, if the spin–spin interactions were allowed to have different strengths, J and J', in the horizontal and vertical directions of the lattice, the specific-heat divergence at $T = T_c$ continued to be logarithmic–independently of the ratio J'/J–even though the value of T_c itself and of the various amplitudes appearing in the final expressions were modified. A similar result for the spontaneous magnetization was obtained by Chang (1952) who showed that, regardless of the value of J'/J, the exponent β continued to be $\frac{1}{8}$. Further corroborative evidence for universality came from the analysis of two-dimensional lattices other than the square one which, despite structural differences, led to the same critical exponents as the ones listed in (45).

2.4. The spherical model in arbitrary dimensions

In the wake of Onsager's solution to the two-dimensional Ising problem in zero field, several attempts were made to go beyond Onsager–by solving either the three-dimensional problem in zero field or the two-dimensional problem with field. None of these attempts succeeded; the best one could accomplish was to rederive the Onsager solution by newer means. This led to the suggestion that one may instead consider certain "adaptations" of the Ising model, which may turn out to be mathematically tractable in more than two dimensions and hopefully throw some light on the problem of phase transitions in more realistic situations (where d is generally 3). One such adaptation was devised by Kac who, in 1947, considered a model in which the spin variable σ_i, instead of being restricted to the *discrete* choices -1 or $+1$, could vary *continuously*, from $-\infty$ to $+\infty$, subject to a Gaussian probability distribution law,

$$p(\sigma_i)\,d\sigma_i = (A/\pi)^{1/2} e^{-A\sigma_i^2}\,d\sigma_i \quad (i = 1, \ldots, N), \tag{1}$$

so that σ_i^2, on an average, $= 1/(2A)$. Clearly, for conformity with the standard practice, viz. $\sigma_i^2 = 1$, the constant A here should be equal to $\frac{1}{2}$; we may, however, leave it arbitrary for the time being. The resulting model is generally referred to as the *Gaussian model*, and its partition function in the presence of the field is given by the multiple integral

$$Q_N = \int_{-\infty}^{+\infty} \cdots \int \left(\frac{A}{\pi}\right)^{N/2} e^{-A\sum\limits_i \sigma_i^2 + K \sum\limits_{\text{n.n.}} \sigma_i\sigma_j + h\sum\limits_i \sigma_i} \prod_i d\sigma_i$$

$$(K = \beta J, h = \beta\mu B). \tag{2}$$

The exponent in the integrand is a symmetric, quadratic function in the σ_i; using standard techniques, it can be diagonalized. Integrations over the (transformed) σ_j can then be carried out straightforwardly–and in any number of dimensions; for details, see Berlin and Kac (1952) or Baker (1990).

One finds that for $d > 2$ the Gaussian model undergoes a phase transition at a *finite* temperature T_c which, for a simple hypercubic lattice, is determined by the condition $K_c = A/d$; note that, with $A = \frac{1}{2}$, this result is precisely the one predicted by the mean field theory (with $q = 2d$). There are differences, though. First of all, the present model does not exhibit a phase transition at a finite temperature for $d \leq 2$. Secondly, the critical exponents for $2 < d < 4$ are nonclassical, in the sense that some of them are d-dependent, though for $d > 4$ they do become classical. More importantly, at temperatures below T_c, where K exceeds A/d, the integral in (2) diverges and the model breaks down! This led Kac to abandon this model and invent a new one in which the spins were again *continuous* variables but subject to an *overall* constraint,

$$\sum_i \sigma_i^2 = N, \tag{3}$$

rather than to individual constraints, $\sigma_i^2 = 1$ for each i, or to an arbitrary probability distribution law. Constraint (3) allows individual spins to vary over a rather wide range, $-N^{1/2}$ to $+N^{1/2}$, but restricts the super spin vector $\{\sigma_i\}$ to the "surface of an N-dimensional hypersphere of radius $N^{1/2}$"; in the Ising model, the same vector is restricted to the "corners of a hypercube inscribed within the above hypersphere". The resulting model is generally referred to as the *spherical model.*

Constraint (3) can be taken care of by inserting an appropriate delta function in the integrand of the partition function. Using the representation

$$\delta\left(N-\sum_i \sigma_i^2\right) = \frac{1}{2\pi i}\int\limits_{x-i\infty}^{x+i\infty} e^{z(N-\sum_i \sigma_i^2)}\,dz, \tag{4}$$

the partition function of the spherical model is given by

$$Q_N = \frac{1}{2\pi i}\int\limits_{x-i\infty}^{x+i\infty} dz\, e^{zN}\int\limits_{-\infty}^{+\infty}\!\!\dots\int e^{-z\sum_i \sigma_i^2+K\sum_{\text{n.n.}}\sigma_i\sigma_j+h\sum_i \sigma_i}\prod_i d\sigma_i. \tag{5}$$

For a fixed z, the integral over the σ_i can be carried out in the same manner as in the Gaussian model; see eqn. (2). Let the result of that calculation be denoted by the symbol $Z_N(K,h;z)$. The partition function Q_N of the spherical model is then given by the complex integral

$$Q_N = \frac{1}{2\pi i}\int\limits_{x-i\infty}^{x+i\infty} dz\, e^{zN} Z_N(K,h;z), \tag{6}$$

which can be evaluated by the *saddle-point* method–also known as the method of *steepest descent*; see Sec. 3.2. One finds that the saddle point of the integrand in (6) lies at the point $z = x_0$, where x_0 is determined by the condition

$$\frac{\partial}{\partial z}\{zN + \ln Z_N(K,h;z)\}\bigg|_{z=x_0} = 0, \tag{7}$$

with the result that, asymptotically,

$$\frac{1}{N}\ln Q_N \approx x_0 + \frac{1}{N}\ln Z_N(K,h;x_0). \tag{8}$$

The thermodynamic properties of the system can then be worked out in detail.

It turned out that many a physicist felt uncomfortable at the necessity of using the method of steepest descent, so a search for an alternative approach seemed desirable. In this connection Lewis and Wannier (1952) pointed out that while the ensemble underlying the model of Berlin and Kac was canonical in the variable E it was microcanonical in the variable $\sum_i \sigma_i^2$. They proposed that one consider instead an ensemble which is canonical in both E and $\sum_i \sigma_i^2$; the method of steepest descent could then be avoided. All one requires now is that the constraint (3) be obeyed only in the sense of an ensemble average,

$$\left\langle \sum_i \sigma_i^2 \right\rangle = N, \tag{9}$$

rather than in the original sense that was comparatively more rigid. The resulting model is generally referred to as the *mean spherical model.*

Constraint (9) can easily be taken care of by modifying the Hamiltonian of the system by including a term proportional to $\sum_i \sigma_i^2$, i.e. writing

$$H = -J \sum_{\text{n.n.}} \sigma_i \sigma_j - \mu B \sum_i \sigma_i + \lambda \sum_i \sigma_i^2, \tag{10}$$

where λ is the so-called *spherical field*, and requiring that

$$\langle (\partial H / \partial \lambda) \rangle = N. \tag{11}$$

The partition function of the revised model is thus given by

$$Q_N = \int_{-\infty}^{-\infty} \cdots \int e^{\beta\lambda \sum_i \sigma_i^2 + K \sum_{\text{n.n.}} \sigma_i \sigma_j + h \sum_i \sigma_i} \prod_i d\sigma_i \tag{12a}$$

$$= Z_N(K, h; \beta\lambda), \tag{12b}$$

with the constraint

$$\frac{1}{\beta} \frac{\partial \ln Z_N(K, h; \beta\lambda)}{\partial \lambda} = -N. \tag{13}$$

Comparing (13) with (7), we readily see that the parameter x_0 of the spherical model is precisely equal to the parameter $\beta\lambda$ of the mean spherical model. The free energy resulting from (12), however, differs a little from the one given by (8), which is not surprising because the transition from a model that was microcanonical in the variable $\mathscr{S}^2 (\equiv \sum_i \sigma_i^2)$ to one that is canonical modifies the nature of the free energy–it goes from "being at constant $\mathscr{S}$" to "being at constant λ". The two free energies are connected by the Legendre transformation

$$A_\lambda = A_{\mathscr{S}} + \lambda \langle \mathscr{S}^2 \rangle = A_{\mathscr{S}} + \lambda N, \tag{14}$$

so that

$$\frac{1}{N} A_{\mathscr{S}} = \frac{1}{N} A_\lambda - \lambda. \tag{15}$$

This is precisely the difference arising from the use of expression (8) or (12).

We shall now proceed to examine the thermodynamic properties of the (mean) spherical model, especially the nature of its critical behavior in arbitrary dimensions. The importance of these results will be discussed towards the end of this section.

(i) *The thermodynamic functions:* We consider a simple hypercubic lattice, of dimensions $L_1 \times \cdots \times L_d$, subject to periodic boundary conditions. The partition function of the system, as given by eqn. (12a), then turns out to be (see Joyce, 1972; Barber and Fisher, 1973)

$$Z_N(K, h; \beta\lambda) = \prod_k \left[\frac{\pi}{\beta(\lambda - \mu_k)} \right]^{1/2} e^{Nh^2/4\beta(\lambda - \mu_0)}, \tag{16}$$

where μ_k are the eigenvalues of the problem,

$$\mu_k = J \sum_{j=1}^{d} \cos(k_j a), \quad k_j = \frac{2\pi n_j}{L_j} \quad \{n_j = 0, 1, \ldots, (N_j - 1)\} \tag{17a}$$

$$N_j = L_j / a, \quad N = \prod_j N_j, \tag{17b}$$

and a the lattice constant of the system. The free en[illegible] A_λ is then given by

$$A_\lambda = \frac{1}{2\beta}\sum_k \ln\frac{\beta(\lambda-\mu_k)}{\pi} - \frac{N\mu^2B^2}{4(\lambda-\mu_0)}, \tag{18}$$

while the parameter λ is determined by the *constraint equation*, see (13),

$$\frac{1}{2\beta}\sum_k \frac{1}{(\lambda-\mu_k)} + \frac{N\mu^2B^2}{4(\lambda-\mu_0)^2} = N. \tag{19}$$

The magnetization $\overline{M}$ and the low-field susceptibility χ_0 follow readily from (18):[2]

$$\overline{M} = \frac{N\mu^2B}{2(\lambda-\mu_0)}, \quad \chi_0 = \frac{N\mu^2}{2(\lambda-\mu_0)}. \tag{20a, b}$$

Introducing the variable $m(\equiv \overline{M}/N\mu)$, the constraint equation (19) may be written in the form

$$\sum_k \frac{1}{(\lambda-\mu_k)} = 2N\beta(1-m^2). \tag{21}$$

Next, the entropy of the system in zero field is given by[3]

$$S_0 = -(\partial A/\partial T)_{\mu_k,B=0} = \tfrac{1}{2}k_B\sum_k[1-\ln\{\beta(\lambda-\mu_k)\}] \tag{22}$$

and the corresponding specific heat by

$$C_0 = T\left(\frac{\partial S_0}{\partial T}\right) = \frac{1}{2}k_B\sum_k\left[1-\frac{T(\partial\lambda/\partial T)_0}{(\lambda-\mu_k)}\right] = N\left[\frac{1}{2}k_B - \left(\frac{\partial\lambda}{\partial T}\right)_0\right]; \tag{23}$$

here, use has been made of eqn. (19), with $B = 0$.

To make further progress we need to determine λ, as a function of B and T, from the constraint eqn. (19). But first note, from eqn. (18), that for the free energy of the system to be well-behaved, λ must be larger than the largest eigenvalue μ_0–which, by (17a), is equal to Jd. At the same time, eqn. (20b) tells us that the singularity of the problem presumably lies at $\lambda = \mu_0$. We may thus infer that, as T decreases from higher values downward, λ also decreases and eventually reaches its lowest possible value, μ_0, at some critical temperature, T_c, where the system undergoes a phase transition. The condition for criticality, therefore, is

$$\lambda_c = \mu_0 = Jd, \tag{24}$$

which suggests that we may introduce a "reduced field", ϕ, by the definition

$$\phi = (\lambda-\lambda_c)/J = (\lambda/J) - d; \tag{25}$$

the condition for criticality then becomes

$$\phi_c = 0. \tag{26}$$

It follows that, as we approach the critical point from above, the parameter ϕ becomes much smaller than unity; ultimately, it becomes zero as T_c is reached and stays so for *all* $T < T_c$.

Now, substituting for the eigenvalues μ_k into the sum appearing in eqns (19) and (21), and making use of the representation

$$\frac{1}{z} = \int_0^\infty e^{-zx}\, dx, \tag{27}$$

we have

$$\sum_k \frac{1}{(\lambda - \mu_k)} = \frac{1}{J} \sum_{\{n_j\}} \int_0^\infty \exp\left\{ - \left[\phi + \sum_{j=1}^{d} \left\{ 1 - \cos\left(\frac{2\pi n_j}{N_j} \right) \right\} \right] x \right\} dx$$

$$= \frac{1}{J} \int_0^\infty e^{-\phi x} \prod_j \left[\sum_{n_j=0}^{N_j-1} \exp\left\{ -x + x\cos\left(\frac{2\pi n_j}{N_j} \right) \right\} \right] dx. \tag{28}$$

For $N_j \gg 1$, the summation over n_j may be replaced by an integration; writing $\theta_j = 2\pi n_j / N_j$, one gets

$$\sum_{n_j} \exp\{\cdots\} \approx \int_0^{2\pi} e^{-x + x\cos\theta_j} \frac{N_j}{2\pi}\, d\theta_j = N_j e^{-x} I_0(x), \tag{29}$$

where $I_0(x)$ is a modified Bessel function. Multiplying over j, one finally gets

$$\sum_k \frac{1}{\lambda - \mu_k} = \frac{N}{J} W_d(\phi), \tag{30}$$

where $W_d(\phi)$ is the so-called *Watson function*, defined by[4]

$$W_d(\phi) = \int_0^\infty e^{-\phi x} [e^{-x} I_0(x)]^d\, dx. \tag{31}$$

Equations (19) and (21) now take the form

$$W_d(\phi) = 2K - \frac{(\beta\mu B)^2}{2K\phi^2} \tag{32a}$$

$$= 2K(1 - m^2). \tag{32b}$$

The asymptotic behavior of the function $W_d(\phi)$, for $\phi \ll 1$, is examined in Appendix F.

(ii) *The critical behavior:* We now analyze the various physical properties of the mean spherical model in different regimes of d.

(a) $d < 2$. For this regime we take expression (7a) of Appendix F and substitute it into eqn. (32a), with $B = 0$. We obtain

$$\phi|_{B=0} \approx \left[\frac{\Gamma\{(2-d)/2\}}{2(2\pi)^{d/2} K} \right]^{2/(2-d)} \sim \left(\frac{k_B T}{J} \right)^{2/(2-d)} \tag{33}$$

We see that ϕ in this case goes to zero only as $T \to 0$. The phase transition, therefore, takes place at $T_c = 0$. Equations (20b), (23), (24) and (25) then give for the low-temperature susceptibility

$$\chi_0 = \frac{N\mu^2}{2J\phi} \sim \frac{N\mu^2}{J}\left(\frac{k_B T}{J}\right)^{-2/(2-d)} \tag{34}$$

and for the low-temperature specific heat

$$C_0 - \frac{1}{2}Nk_B = -NJ\left(\frac{\partial\phi}{\partial T}\right)_0 \sim -Nk_B\left(\frac{k_B T}{J}\right)^{d/(2-d)} \tag{35}$$

(b) $d = 2$. We now use expression (7b) of Appendix F and obtain

$$\phi|_{B=0} \sim \exp(-4\pi J/k_B T), \tag{36}$$

so that once again $T_c = 0$ but now at low temperatures

$$\chi_0 \sim (N\mu^2/J)\exp(4\pi J/k_B T) \tag{37}$$

and

$$C_0 - \tfrac{1}{2}Nk_B \sim -Nk_B(J/k_B T)^2 \exp(-4\pi J/k_B T). \tag{38}$$

(c) $2 < d < 4$. We now substitute expression (7c) of Appendix F into eqn. (32a), with the result

$$W_d(0) - \frac{|\Gamma\{(2-d)/2\}|}{(2\pi)^{d/2}}\phi^{(d-2)/2} = 2K - \frac{(\beta\mu B)^2}{2K\phi^2}. \tag{39}$$

The critical point is now determined by setting $B = 0$ and letting $\phi \to 0$; the condition for criticality then reads

$$K_c = \tfrac{1}{2}W_d(0). \tag{40}$$

The variation of ϕ with T as one approaches the critical point is given by

$$\phi|_{B=0} \approx \left[\frac{2(2\pi)^{d/2}(K_c - K)}{|\Gamma\{(2-d)/2\}|}\right]^{2/(d-2)} \quad (K \lesssim K_c). \tag{41a}$$

We also note that once ϕ becomes zero it stays so for all temperatures below, i.e.

$$\phi|_{B=0} = 0 \quad (K \geq K_c). \tag{41b}$$

It now follows that

$$\chi_0 \sim (K_c - K)^{-2/(d-2)} \sim (T - T_c)^{-2/(d-2)} \quad (T \gtrsim T_c) \tag{42}$$

and is infinite for $T \leq T_c$. At the same time

$$C_0 - \tfrac{1}{2}Nk_B \sim (T - T_c)^{(4-d)/(d-2)} \quad (T \gtrsim T_c) \tag{43}$$

and it vanishes for $T \leq T_c$. The spontaneous magnetization is determined by eqns (32b), (40) and (41b); we obtain a remarkably simple result

$$m_0 = (1 - K_c/K)^{1/2} = (1 - T/T_c)^{1/2} \quad (T \leq T_c). \tag{44}$$

Finally, if in eqn. (39) we retain B but set $T = T_c$, we get

$$\phi_c \sim B^{4/(d+2)} \quad (T = T_c); \tag{45}$$

eqn. (20a) then gives

$$m_c = \frac{\mu B}{2J\phi_c} \sim B^{(d-2)/(d+2)} \quad (T = T_c). \tag{46}$$

The foregoing results give, for the critical exponents of the system,

$$\alpha = \frac{d-4}{d-2}, \quad \beta = \frac{1}{2}, \quad \gamma = \frac{2}{d-2}, \quad \delta = \frac{d+2}{d-2} \quad (2 < d < 4). \tag{47}$$

(d) $d > 4$. In this regime we employ expression (8) of Appendix F. The condition for criticality remains the same as in (40); the variation of ϕ with T as we approach the critical point is, however, different. We now have:

$$\phi|_{B=0} \sim (K_c - K)^1 \quad (K \lesssim K_c). \tag{48}$$

The subsequent results are modified accordingly:

$$\chi_0 \sim (T - T_c)^{-1}, \quad C_0 - \tfrac{1}{2}Nk_B \sim (T - T_c)^0 \qquad (T \gtrsim T_c) \tag{49}$$

$$\phi_c \sim B^{2/3}, \quad m_c \sim B^{1/3} \qquad (T = T_c). \tag{50}$$

Equations (41b) and (44) continue to apply as such. We therefore conclude that

$$\alpha = 0, \quad \beta = \tfrac{1}{2}, \quad \gamma = 1, \quad \delta = 3 \quad (d > 4). \tag{51}$$

(e) $d = 4$. For this borderline case, we use expression (12) of Appendix F. Once again, the condition for criticality remains the same; however, the variation of ϕ with T as one approaches the critical point is now determined by the implicit relation

$$\phi \ln(1/\phi) \approx 8\pi^2(K_c - K) \quad (K \lesssim K_c). \tag{52}$$

Introducing the conventional parameter

$$t = (T - T_c)/T_c = (K_c - K)/K, \tag{53}$$

we get, to leading order in t,

$$\phi|_{B=0} \sim t/\ln(1/t) \quad (0 < t \ll 1). \tag{54}$$

It follows that

$$\chi_0 \sim t^{-1}\ln(1/t), \quad C_0 \sim 1/\ln(1/t) \quad (0 < t \ll 1) \tag{55}$$

$$\phi_c \sim B^{2/3}/\{\ln(1/B)\}^{1/3}, \quad m_c \sim \{B\ln(1/B)\}^{1/3} \quad (t = 0). \tag{56}$$

(iii) *Spin–spin correlations:* Following the procedure that led to eqns (16)–(19), we obtain *in the absence of the field*

$$G(\boldsymbol{r}, \boldsymbol{r}') \equiv \overline{\sigma(\boldsymbol{r})\sigma(\boldsymbol{r}')} = \frac{1}{2N\beta}\sum_{\boldsymbol{k}} \frac{\exp\{i(\boldsymbol{k}\cdot\boldsymbol{R})\}}{\lambda - \mu_{\boldsymbol{k}}} \quad (\boldsymbol{R} = \boldsymbol{r} - \boldsymbol{r}'); \tag{57}$$

cf. eqn. (19) with $B = 0$. The summation over $\boldsymbol{k}$ in (57) can be handled in the same manner as was done in (28); however, the resulting summation over n_j now turns out to be

$$\begin{aligned}\sum_{n_j}\{\cdots\} &\approx \int_0^{2\pi} \exp\{iR_j\theta_j/a\} e^{-x+x\cos\theta_j}\frac{N_j}{2\pi}\, d\theta_j \\ &= N_j e^{-x} I_{R_j/a}(x);\end{aligned} \tag{58}$$

cf. (29). This leads to the result

$$G(R) = \frac{1}{2K}\int_0^{\infty} e^{-\phi x}\prod_j [e^{-x} I_{R_j/a}(x)]\, dx; \tag{59}$$

cf. eqn. (32a), with $B = 0$. For the functions $I_n(x)$ we may use the asymptotic expression (see Singh and Pathria, 1985a)

$$I_n(x) \approx \frac{e^{x-n^2/2x}}{\surd(2\pi x)} \quad (x \gg 1), \tag{60}$$

so that, for $\psi \ll 1$,

$$\begin{aligned}G(R) &\approx \frac{1}{2(2\pi)^{d/2}K}\int_0^{\infty} e^{-\phi x - R^2/2a^2x} x^{-d/2}\, dx \\ &= \frac{1}{(2\pi)^{d/2}K}\left(\frac{a^2}{\xi R}\right)^{(d-2)/2} K_{(d-2)/2}\left(\frac{R}{\xi}\right),\end{aligned} \tag{61}$$

where $K_\mu(x)$ is the other modified Bessel function while

$$\xi = a/(2\phi)^{1/2}. \tag{62}$$

For $R \gg \xi$, we may use the asymptotic result $K_\mu(x) \approx (\pi/2x)^{1/2}e^{-x}$; eqn. (61) then becomes

$$G(R) \approx \frac{a^{d-2}}{2K\xi^{(d-3)/2}(2\pi R)^{(d-1)/2}} e^{-R/\xi}, \tag{63}$$

which identifies ξ as the *correlation length* of the system.

Now, comparing (62) with (20b), we find that

$$\chi_0 = \frac{N\mu^2}{2J\phi} = \frac{N\mu^2}{Ja^2}\xi^2. \tag{64}$$

In view of the fact that $\xi \sim \chi_0^{1/2}$, we infer that, in *all* regimes of d, the exponent $\nu = \frac{1}{2}\gamma$ and hence, by relations (11.12.10 and 11), the exponent $\eta = 0$. To obtain this last result directly from (61), we observe that, as $T \to T_c$ from above, the parameter $\phi \to 0$ and hence $\xi \to \infty$. We may then use the approximation $R/\xi \ll 1$ and employ the formula

$$K_\mu(x) \approx \tfrac{1}{2}\Gamma(\mu)\left(\tfrac{1}{2}x\right)^{-\mu} \quad (\mu > 0), \tag{65}$$

to obtain

$$G(R)|_{T=T_c} \approx \frac{\Gamma\{(d-2)/2\}}{4\pi^{d/2}K_c}\frac{a^{d-2}}{R^{d-2}} \quad (d > 2). \tag{66}$$

The power R appearing here clearly shows that $\eta = 0$. Finally, substituting (41a) into (6_), we get

$$\xi \approx \frac{1}{2}a\left[\frac{|\Gamma\{(2-d)/2\}|}{4\pi^{d/2}(K_c - K)}\right]^{1/(d-2)} \quad (K \lesssim K_c), \tag{67}$$

which shows that for $2 < d < 4$ the critical exponent $\nu = 1/(d-2)$.

For $T < T_c$ we expect the function $G(\boldsymbol{R})$ to affirm the presence of *long range order* in the system, i.e. in the limit $R \to \infty$, it should tend to a limit, $\overline{\sigma}^2$, that is *nonzero*. To demonstrate this property of $G(\boldsymbol{R})$, we need to take a closer look at the derivations of this sub-section which were carried out with the express purpose of obtaining results valid in the thermodynamic limit ($N \to \infty$). This resulted in "errors" that were negligible in the region $T \gtrsim T_c$ but are not so when $T < T_c$. The first such error crept in when we replaced the summations over $\{n_j\}$ in eqn. (28) by integrations; that suppressed contribution from the term with $n = 0$. Equation (30), therefore, accounts for *only* the ($k \neq 0$)-terms of the original sum in (28), and the missing term, $1/J\phi$, may be added to it *ad hoc*.[5] Equation (19), with $B = 0$, then becomes

$$\frac{1}{2\beta}\left[\frac{1}{J\phi} + \frac{N}{J}W_d(\phi)\right] = N. \tag{68}$$

Now, when ϕ becomes very very small, $W_d(\phi)$ may be approximated by $W_d(0)$ which is precisely equal to $2K_c$; eqn. (68) then gives

$$\phi \approx [2N(K - K_c)]^{-1} \quad (K > K_c), \tag{69}$$

rather than zero! The correlation length then turns out to be

$$\xi = a/(2\phi)^{1/2} \approx a[N(K - K_c)]^{1/2} \quad (K > K_c), \tag{70}$$

rather than infinite! Now, the same error was committed once again in going from eqn. (57) to (59); so, the primary result for $G(\boldsymbol{R})$, as given in eqn. (61), may be similarly amended by adding the missing term $1/(2N\beta J\phi)$ which, by (69), is exactly equal to $(1 - K_c/K)$. Thus, for $R \ll \xi$, we obtain, instead of (66),

$$G(\boldsymbol{R}) \approx \left(1 - \frac{K_c}{K}\right) + \frac{\Gamma\{(d-2)/2\}}{4\pi^{d/2}K}\frac{a^{d-2}}{R^{d-2}} \quad (K > K_c). \tag{71}$$

Now if we let $R \to \infty$, $G(\boldsymbol{R})$ does approach a *nonzero* value $\overline{\sigma}^2$, which is precisely the same as m_0^2 given by eqn. (44). It is remarkable, though, that in the present derivation the magnetic field B has not been introduced at any stage of the calculation, which underscores the fundamental role played by correlations in bringing about long range order in the system.

In the preceding paragraph we have outlined the essential argument that led to the desired expression, (71), for $G(\boldsymbol{R})$. For a more rigorous analysis of this problem, see Singh and Pathria (1985b, 1987a).

(iv) *Physical significance of the spherical model*: With a constraint as relaxed as (3), or even more so (9), one wonders how meaningful the spherical model is from a physical point of view. Relief comes from the fact, first established by

Stanley (1968, 1969a,b), that the spherical model provides a correct representation of the ($n \to \infty$)-limit of an n-vector model with nearest-neighbor interactions; see also Kac and Thompson (1971). This connection arises from the very nature of the constraint imposed on the model, which introduces a super spin vector $\mathscr{S}$ with N degrees of freedom; it is not surprising that, in the limit $N \to \infty$, the model in some sense acquires the same sort of freedom that an n-vector model has in the limit $n \to \infty$. In any case, this connection brings the spherical model in line with, and actually makes it a good guide for, all models with *continuous* symmetry, viz. the ones with $n \geq 2$. And since it can be solved exactly in arbitrary dimensions, this model gives us some idea as to what to expect of others for which n is finite. For instance, we have seen that, for $d > 4$, the critical exponents of the spherical model are the same as the ones obtained from the mean field theory. Now, fluctuations are neglected in the mean field theory but, among the variety of models we are considering, fluctuations should be largest in the spherical model, for it has the largest number of degrees of freedom. If, for $d > 4$, fluctuations turn out to be negligible in the spherical model, they would be even more so in models with finite n. It thus follows that, regardless of the actual value of n, mean field theory should be valid for *all* these models when $d > 4$. See, in this connection, Sec. 13.4 as well.

For $d < 4$, the final results depend significantly on n. The spherical model now provides a starting point from which one may carry out the so-called ($1/n$)-expansions to determine how models with finite n would behave in this regime. Such an approach was initiated by Abe and collaborators (1972, 1973) and independently by Ma (1973); for a detailed account of this approach, along with the results following from it, see Ma (1976c).

Finally, for a comprehensive discussion of the spherical model, including the one with long-range interactions, see the review article by Joyce (1972).

12.5. The ideal Bose gas in arbitrary dimensions

In this section we propose to examine the problem of Bose–Einstein condensation in an ideal Bose gas in *arbitrary* dimensions. As was first shown by Gunton and Buckingham (1968), the phenomenon of Bose–Einstein condensation falls in the same universality class as the phase transition in the spherical model; accordingly, the ideal Bose gas too corresponds to the ($n \to \infty$)-limit of an n-vector model. It must, however, be borne in mind that liquid He^4, whose transition from a normal to a superfluid state is often regarded as a manifestation of the "Bose–Einstein condensation in an interacting Bose liquid", actually pertains to the case $n = 2$. Now, just as the spherical model turns out to be a good guide for all models with continuous symmetry [including the X–Y model (for which $n = 2$)], in the same way the ideal Bose gas has also been a good guide for liquid He^4.

We consider a Bose gas composed of N non-interacting particles confined to a box of volume $V (= L_i \times \cdots \times L_d)$ at temperature T. Following the procedure of Sec. 7.1, we obtain for the pressure P of the gas

$$P = -\frac{k_B T}{V} \sum_{\varepsilon} \ln\left(1 - z e^{-\beta\varepsilon}\right) = \frac{k_B T}{\lambda^d} g_{(d+2)/2}(z), \qquad (1)$$

where $\lambda[= h/\sqrt{(2\pi m k_B T)}]$ is the mean thermal wavelength of the particles, z is the fugacity of the gas which is related to the chemical potential μ through the formula

$$z = \exp(\beta\mu) < 1 \quad (\beta = 1/k_B T), \tag{2}$$

while $g_\nu(z)$ are Bose–Einstein functions whose main properties are discussed in Appendix D. The quantity z is determined by the equation

$$N = \sum_\varepsilon (z^{-1}e^{\beta\varepsilon} - 1)^{-1} = N_0 + N_e, \tag{3}$$

where N_0 is the mean number of particles in the ground state ($\varepsilon = 0$),

$$N_0 = z/(1-z), \tag{4}$$

while N_e is the mean number of particles in the excited states ($\varepsilon > 0$):

$$N_e = \frac{V}{\lambda^d} g_{d/2}(z). \tag{5}$$

At high temperatures, where z is significantly below the limiting value 1, N_0 is negligible in comparison with N; the quantity z is then determined by the simplified equation

$$N = \frac{V}{\lambda^d} g_{d/2}(z), \tag{6}$$

and the pressure P in turn is given by the expression

$$P = \frac{Nk_B T}{V} \frac{g_{(d+2)/2}(z)}{g_{d/2}(z)}. \tag{7}$$

The internal energy of the gas may be obtained from the relationship

$$U = \tfrac{1}{2} d(PV); \tag{8}$$

see the corresponding derivation of eqn. (7.1.12) as well as of eqn. (6.4.4). Now, making use of the recurrence relation (D. 10) and remembering that the mean thermal wavelength $\lambda \propto T^{-1/2}$, we get from eqn. (6)

$$\frac{1}{z}\left(\frac{\partial z}{\partial T}\right)_v = -\frac{d}{2T}\frac{g_{d/2}(z)}{g_{(d-2)/2}(z)} \quad \left(v = \frac{V}{N}\right), \tag{9}$$

and from eqn. (1)

$$\frac{1}{z}\left(\frac{\partial z}{\partial T}\right)_P = -\frac{d+2}{2T}\frac{g_{(d+2)/2}(z)}{g_{d/2}(z)}. \tag{10}$$

It is now straightforward to show that the specific heats C_V and C_P of the gas are given by the formulae

$$\frac{C_V}{Nk_B} = \frac{d(d+2)}{4}\frac{g_{(d+2)/2}(z)}{g_{d/2}(z)} - \frac{d^2}{4}\frac{g_{d/2}(z)}{g_{(d-2)/2}(z)} \tag{11}$$

and

$$\frac{C_P}{Nk_B} = \frac{(d+2)^2}{4}\frac{\{g_{(d+2)/2}(z)\}^2 g_{(d-2)/2}(z)}{\{g_{d/2}(z)\}^3} - \frac{d(d+2)}{4}\frac{g_{(d+2)/2}(z)}{g_{d/2}(z)}, \tag{12}$$

respectively; it follows that the ratio

$$\frac{C_P}{C_V} = \frac{(d+2)}{d} \frac{g_{(d+2)/2}(z) g_{(d-2)/2}(z)}{\{g_{d/2}(z)\}^2}. \tag{13}$$

The isothermal compressibility κ_T and the adiabatic compressibility κ_S turn out to be

$$\kappa_T = -\frac{1}{v}\left(\frac{\partial v}{\partial P}\right)_T = -\frac{1}{v}\frac{(\partial v/\partial z)_T}{(\partial P/\partial z)_T} = \frac{g_{(d+2)/2}(z) g_{(d-2)/2}(z)}{\{g_{d/2}(z)\}^2}\frac{1}{P} \tag{14}$$

and

$$\kappa_S = -\frac{1}{v}\left(\frac{\partial v}{\partial P}\right)_S = -\frac{1}{v}\frac{(\partial v/\partial T)_z}{(\partial P/\partial T)_z} = \frac{d}{d+2}\frac{1}{P}, \tag{15}$$

respectively; note that the ratio κ_T/κ_S is precisely equal to the ratio C_P/C_V, as is expected thermodynamically.

As the temperature of the gas is reduced, keeping v constant, the fugacity z increases and ultimately approaches its limiting value 1–marking the end of the regime where N_0 was negligible in comparison with N. Whether this limit is reached at a finite T or at $T = 0$ depends entirely on the value of d; see eqn. (6), which tells us that if the function $g_{d/2}(z)$, as $z \to 1-$, is *bounded* then the limit in question will be reached at a finite T. On the other hand, if $g_{d/2}(z)$, as $z \to 1-$, is *unbounded* then the desired limit will be reached at $T = 0$ instead. To settle this question, we refer to eqns (D. 9) and (D. 11) which summarize the behavior of the function $g_\nu(z)$ as $z \to 1-$ (or as $\alpha \to 0+$, where $\alpha = -\ln z$); we thus have

$$g_{d/2}(e^{-\alpha}) \approx \begin{cases} \Gamma\left(\frac{2-d}{2}\right)\alpha^{-(2-d)/2} + \text{const.} & \text{for } d < 2 \quad (16a) \\ \ln(1/\alpha) + \frac{1}{2}\alpha & \text{for } d = 2 \quad (16b) \\ \zeta\left(\frac{d}{2}\right) - \left|\Gamma\left(\frac{2-d}{2}\right)\right|\alpha^{(d-2)/2} & \text{for } 2 < d < 4 \quad (16c) \\ \zeta(2) - \{\ln(1/\alpha) + 1\}\alpha & \text{for } d = 4 \quad (16d) \\ \zeta\left(\frac{d}{2}\right) - \zeta\left(\frac{d-2}{2}\right)\alpha & \text{for } d > 4, \quad (16e) \end{cases}$$

$\zeta(\nu)$ being the Riemann zeta function. Similarity with the spherical model is quite transparent. We readily see that, for $d > 2$, $\alpha \to 0$ at a *finite* temperature, T_c, given by the equation

$$\lambda_c^d = v\zeta(d/2), \tag{17}$$

with the result that

$$T_c = \frac{h^2}{2\pi m k_B}\left[\frac{1}{v\zeta(d/2)}\right]^{2/d}; \tag{18}$$

for $d \leq 2$, $\alpha \to 0$ only as $\lambda \to \infty$, so $T_c = 0$. For brevity, we shall confine further discussion only to $d > 2$.

(i) *The critical behavior*: As T approaches T_c from above, the manner in which $\alpha \to 0$ is determined by substituting the appropriate eqn. (16) into (6) and utilizing the criticality condition (17). For $2 < d < 4$, one gets asymptotically

$$\left|\Gamma\left(\frac{2-d}{2}\right)\right| \alpha^{(d-2)/2} \approx \frac{1}{v}(\lambda_c^d - \lambda^d) \simeq \frac{d}{2}\zeta\left(\frac{d}{2}\right)\left[\frac{T}{T_c} - 1\right]. \tag{19}$$

For $T \gtrsim T_c$, this gives

$$\alpha \sim t^{2/(d-2)} \quad [t = (T - T_c)/T_c, 0 < t \ll 1]. \tag{20}$$

As $T \to T_c$, the specific heat C_P and the isothermal compressibility κ_T diverge because the function $g_{(d-2)/2}(z)$ appearing in eqns (12) and (14), being $\sim \alpha^{-(4-d)/2}$ [see eqn. (D.9)], becomes divergent; for small t,

$$C_P \sim \kappa_T \sim t^{-(4-d)/(d-2)}. \tag{21}$$

The specific heat C_V, on the other hand, approaches a finite value,

$$\left(\frac{C_V}{Nk_B}\right)_{T \to T_C+} = \frac{d(d+2)}{4}\frac{\zeta\{(d+2)/2\}}{\zeta(d/2)}, \tag{22}$$

with a derivative that, depending on the actual value of d, might diverge:

$$\frac{1}{Nk_B}\left(\frac{\partial C_V}{\partial T}\right) = \frac{1}{T}\left[\frac{d^2(d+2)}{8}\frac{g_{(d+2)/2}(z)}{g_{d/2}(z)} - \frac{d^2}{4}\frac{g_{d/2}(z)}{g_{(d-2)/2}(z)} - \frac{d^3}{8}\frac{\{g_{d/2}(z)\}^2 g_{(d-4)/2}(z)}{\{g_{(d-2)/2}(z)\}^3}\right] \tag{23a}$$

$$\sim -\alpha^{-(d-3)} \sim -t^{-2(d-3)/(d-2)} \quad (3 < d < 4). \tag{23b}$$

Equating the exponent appearing here with $(1 + \alpha)$, we conclude[6] that the critical exponent α for this system is $(d-4)/(d-2)$; cf. eqn. (12.4.47). For a proper appreciation of the critical behavior of C_V, we must as well examine the region $T < T_c$, along with the limit $T \to T_c-$.

For $T < T_c$, the fugacity z is essentially equal to 1; eqns (5) and (17) then give

$$N_e = \frac{V}{\lambda^d}\zeta\left(\frac{d}{2}\right) = N\left(\frac{\lambda_c}{\lambda}\right)^d = N\left(\frac{T}{T_c}\right)^{d/2} \tag{24}$$

It follows that

$$N_0 = N - N_e = N\left[1 - \left(\frac{T}{T_c}\right)^{d/2}\right] \tag{25}$$

Equation (4) then tells us that the precise value of z in this region is given by

$$z = N_0/(N_0 + 1) \simeq 1 - 1/N_0, \tag{26}$$

whence

$$\alpha = -\ln z \simeq 1/N_0, \tag{27}$$

rather than zero. Disregarding this subtlety, eqn. (1) gives

$$P = \frac{k_B T}{\lambda^d}\zeta\left(\frac{d+2}{2}\right) \propto T^{(d+2)/2}. \tag{28}$$

Since P here is a function of T only, the quantities κ_T and C_P in this region are infinite; see, however, Problem 12.26. From eqns (8) and (28), we get

$$U = \frac{1}{2} d \frac{k_B T V}{\lambda^d} \zeta\left(\frac{d+2}{2}\right), \tag{29}$$

whence

$$\frac{C_V}{Nk_B} = \frac{d(d+2)}{4} \frac{v}{\lambda^d} \zeta\left(\frac{d+2}{2}\right) = \frac{d(d+2)}{4} \frac{\zeta\{(d+2)/2\}}{\zeta(d/2)} \left(\frac{T}{T_c}\right)^{d/2}. \tag{30}$$

As $T \to T_c-$, we obtain precisely the same limit as in (22)–showing that C_V is *continuous* at the critical point. Its derivative, however, turns out to be different from the one given in (23), for now

$$\frac{1}{Nk_B}\left(\frac{\partial C_V}{\partial T}\right) = \frac{1}{T_c} \frac{d^2(d+2)}{8} \frac{\zeta\{(d+2)/2\}}{\zeta(d/2)} \left(\frac{T}{T_c}\right)^{(d-2)/2} \tag{31a}$$

$$\to \frac{1}{T_c} \frac{d^2(d+2)}{8} \frac{\zeta\{(d+2)/2\}}{\zeta(d/2)} \tag{31b}$$

as $T \to T_c-$.

As for the condensate fraction, N_0/N, eqn. (25) gives

$$\frac{N_0}{N} = 1 - \left(\frac{T}{T_c}\right)^{d/2} \approx \frac{d}{2}|t| \quad [t < 0, |t| \ll 1]. \tag{32}$$

Now the order parameter in the present problem is a complex number, Ψ_0, such that $|\Psi_0|^2 = N_0/V$, the condensate particle density in the system; see Gunton and Buckingham (1968). We therefore expect that, for $|t| \ll 1, N_0$ would be $\sim t^{2\beta}$; eqn. (32) then tells us that the critical exponent β in this case has the classical value $\frac{1}{2}$ for *all* $d > 2$. To determine the exponents γ and δ, we must introduce a "complex Bose field, conjugate to the order parameter Ψ_0" and examine quantities such as the "Bose susceptibility" χ as well as the variation of Ψ_0 with the Bose field at $T = T_c$. One obtains: $\gamma = 2/(d-2)$ and $\delta = (d+2)/(d-2)$, just as for the spherical model.

(ii) *The correlations*: We now examine the correlation function of the ideal Bose gas

$$G(\boldsymbol{R}) = \frac{1}{V} \sum_{\boldsymbol{k}} \frac{e^{i\boldsymbol{k}\cdot\boldsymbol{R}}}{e^{\alpha+\beta\varepsilon(\boldsymbol{k})} - 1}. \tag{33}$$

As usual, we replace the summation over $\boldsymbol{k}$ by an integration (remembering that this replacement suppresses the ($\boldsymbol{k} = 0$)-term which may, therefore, be kept aside). Making use of eqn. (C.11), we get

$$\begin{aligned} G(\boldsymbol{R}) &= \frac{N_0}{V} + \frac{1}{(2\pi)^d} \int \frac{e^{i\boldsymbol{k}\cdot\boldsymbol{R}}}{e^{\alpha+\beta\hbar^2k^2/2m} - 1} d^d k \\ &= \frac{N_0}{V} + \frac{1}{(2\pi)^{d/2} R^{(d-2)/2}} \int_0^\infty \left[\sum_{j=1}^{\infty} e^{-j\alpha - j\beta\hbar^2k^2/2m}\right] J_{(d-2)/2}(kR) k^{d/2}\, dk \end{aligned}$$

$$= \frac{N_0}{V} + \frac{1}{\lambda^d} \sum_{j=1}^{\infty} e^{-j\alpha - \pi R^2/j\lambda^2} j^{-d/2} \qquad \left[\lambda = \hbar \left(\frac{2\pi\beta}{m}\right)^{1/2}\right]; \tag{34}$$

cf. eqns (3)–(5) which pertain to the case $R = 0$. For $R > 0$, one may extend the summation over j from $j = 0$ to $j = \infty$, for the term so added is identically zero. At the same time, the summation over j may be replaced by an integration–committing errors $O(e^{-R/\lambda})$, which are negligible so long as $R \gg \lambda$; for details, see Zasada and Pathria (1976). We thus obtain

$$G(R) = \frac{N_0}{V} + \frac{1}{\lambda^d} \int_0^{\infty} e^{-j\alpha - \pi R^2/j\lambda^2} j^{-d/2}\, dj$$

$$= \frac{N_0}{V} + \frac{2}{\lambda^2 (2\pi\xi R)^{(d-2)/2}} K_{(d-2)/2}\left(\frac{R}{\xi}\right), \tag{35}$$

where $K_\mu(x)$ is a modified Bessel function while

$$\xi = \lambda/(2\pi^{1/2}\alpha^{1/2}). \tag{36}$$

For $T \gtrsim T_c$, we may use expression (20) for α; eqn. (36) then gives

$$\xi \sim \lambda t^{-1/(d-2)} \quad (0 < t \ll 1), \tag{37}$$

which means that $\xi \gg \lambda$. Now, if $R \gg \xi$, eqn. (35) reduces to

$$G(R) \approx \frac{1}{\lambda^2 (2\pi\xi)^{(d-3)/2} R^{(d-1)/2}} e^{-R/\xi}, \tag{38}$$

which identifies ξ as the *correlation length* of the system. Equation (37) then tells us that for $2 < d < 4$ the critical exponent ν of the ideal Bose gas is $1/(d-2)$. At $T = T_c$, ξ is infinite; eqn. (35) now gives

$$G(R) \approx \frac{\Gamma\{(d-2)/2\}}{\pi^{(d-2)/2} \lambda_c^2 R^{d-2}}, \tag{39}$$

which shows that the critical exponent $\eta = 0$. For $T < T_c$, ξ continues to be infinite but now the condensate fraction, which is a measure of the long-range order in the system, is nonzero. The correlation function then assumes the form

$$G(R) = |\Psi_0|^2 + \frac{A(T)}{R^{d-2}}, \tag{40}$$

where

$$A(T) = \frac{\Gamma\{(d-2)/2\}}{\pi^{(d-2)/2} \lambda^2} \propto T; \tag{41}$$

cf. the corresponding eqn. (12.4.71) of the spherical model.

In the paper quoted earlier, Gunton and Buckingham also generalized the study of Bose–Einstein condensation to the single-particle energy spectrum $\varepsilon \sim k^\sigma$, where σ is a positive number not necessarily equal to 2. They found that the phase transition, at a finite temperature T_c, now took place for all $d > \sigma$, and the critical exponents in the regime $\sigma < d < 2\sigma$ turned out to be

$$\left.\begin{array}{c} \alpha = \dfrac{d-2\sigma}{d-\sigma}, \quad \beta = \dfrac{1}{2}, \quad \gamma = \dfrac{\sigma}{d-\sigma}, \quad \delta = \dfrac{d+\sigma}{d-\sigma} \\ \nu = \dfrac{1}{d-\sigma}, \quad \eta = 2-\sigma. \end{array}\right\} \tag{42}$$

While mathematically correct, these results left one with the awkward conclusion that the Bose gas in its extreme relativistic state ($\sigma = 1$) was in a *different* universality class than the one in the nonrelativistic state ($\sigma = 2$). It was shown later by Singh and Pandita (1983) that, if one employs the appropriate energy spectrum $\varepsilon = c\sqrt{(m_0^2c^2 + \hbar^2k^2)}$ and at the same time allows for the possibility of particle–antiparticle pair production in the system, as had been suggested earlier by Haber and Weldon (1981, 1982), then the relativistic Bose gas falls in the *same* universality class as the nonrelativistic one.

12.6. Other models

In Sec. 12.3 we saw that a two-dimensional lattice model characterized by a discrete order parameter ($n = 1$) undergoes a phase transition, accompanied by a spontaneous magnetization m_0, at a *finite* temperature T_c; naturally, one would expect the same if $d > 2$. On the other hand, the spherical model, which is characterized by a continuous order parameter (with $n = \infty$), undergoes such a transition only if $d > 2$. The question then arises whether intermediate models, with $n = 2, 3, \ldots$, would undergo a phase transition at a finite T_c if d were equal to 2. The answer to this question was provided by Mermin and Wagner (1966) who, making use of a well-known inequality due to Bogoliubov (1962), established the following theorem:[7]

> "Systems composed of spins with continuous symmetry ($n \geq 2$) and short-range interactions do not acquire spontaneous magnetization at any finite temperature T if the space dimensionality $d \leq 2$."

In this sense, systems with $n \geq 2$ behave in the same manner as the spherical model–and quite unlike the Ising model.

The marginal case ($n = 2, d = 2$), however, deserves a special mention. Clearly, this refers to an X–Y model in two dimensions, which has a direct relevance to superfluid He^4 adsorbed on a substrate. As shown by Kosterlitz and Thouless (1972, 1973), this model exhibits a curious phase transition in that, while no long-range order develops at any finite temperature T, various physical quantities such as the susceptibility, the correlation length, the specific heat and the superfluid density become singular at a finite temperature T_c whose precise value is determined by point defects, such as vortices or dislocations, in the system. The correlation length ξ, as $T \to T_c+$, displays an essential singularity, viz.

$$\xi \sim e^{b'/(T-T_c)^{1/2}}, \tag{1}$$

where b' is a non-universal constant; of course, the critical temperature itself is non-universal and for a square lattice is given by

$$K_c = J/kT_c = 2/\pi. \tag{2}$$

The singular part of the specific heat shows a somewhat similar behavior, viz.

$$c^{(s)} \sim \xi^{-2} \tag{3}$$

which, measured against a regular background, is rather indetectable. The superfluid density behaves rather strangely; it approaches a finite value, as $T \to T_c-$, preceded by a square-root cusp. The correlation function is no different; at $T = T_c$, Kosterlitz (1974) found a logarithmic factor along with a power law, viz.

$$g(r) \sim \frac{[\ln(r/a)]^{1/8}}{r^{1/4}}, \tag{4}$$

while for $T < T_c$ we encounter a temperature-dependent exponent η such that

$$g(r) \sim \frac{1}{r^{\eta(T)}} \quad [\eta(T) = T/4T_c]. \tag{5}$$

For further details of this transition, see Kosterlitz and Thouless (1978) and Nelson (1983); for a pedagogical account, see Plischke and Bergersen (1989), Sec. 5.E.

For other exactly soluble models in two dimensions, see Baxter (1982), Wu (1982), Nienhuis (1987) and Cardy (1987), where references to other relevant literature on the subject can also be found.

We now proceed to consider the situation in three dimensions. Here no exact solutions exist except for the extreme case $n = \infty$, which was discussed in Sec. 12.4. However, an enormous amount of effort has been spent in obtaining approximate solutions which, over the years, have become accurate enough to be regarded as "almost exact". Irrespective of the model under study, the problem has generally been attacked along three different lines which, after some refinements, have led to almost identical results. In summary, these approaches may be described as follows:

(i) *The method of series expansions.* In this approach, the partition function or other relevant properties of the system are expanded as high-temperature series such as (12.3.5), with expansion parameter $v = \tanh(\beta J)$, or as low-temperature series such as (12.3.9), with expansion parameter $w = \exp(-2\beta J)$; in the presence of an external field, one would have series with two expansion parameters instead of one. In either case, the first major task involves the numerical computation of coefficients, such as $n(r)$ and $m(r)$, on the basis of graph theory and other allied techniques, while the second major task involves analysis of these coefficients with a view to determining the location and the nature of the singularity governing the property in question. The latter task is generally accomplished by employing the *ratio method*, which examines the trend of the ratio of two consecutive coefficients, such as $n(r)$ and $n(r-1)$, as $r \to \infty$, or by constructing *Padé approximants* which, in a sense, provide a continuation of the known (finite) series beyond its normal region of validity up to its radius of convergence–thus locating and examining the nature of the relevant singularity. Since their inception (in the mid-1950s for the ratio method and the early 1960s for the use of Padé approximants), these techniques have been expanded, refined and enriched in so many ways that it would be hopeless to attempt a proper review here. Suffice it to say that the reader may refer to vol. 3 of the Domb–Green series, which is devoted solely to the method of series expansions–in particular, to the articles by Gaunt and Guttmann (1974) on the asymptotic analysis of the various coefficients, by Domb (1974) on the Ising model, by Rushbrooke, Baker and Wood (1974) on the Heisenberg model, by Stanley (1974) on the n-vector models, and by Betts (1974) on the X–Y model. For more recent reviews. see Guttmann (1989) and Baker (1990), where references to other relevant work on the subject are also available.

(ii) *The renormalization group method.* This method is based on the crucial observation (Wilson, 1971) that, as the critical point is approached, the correlation length of the system becomes exceedingly large–with the result that the sensitivity of the system to a *length transformation* (or renormalization) gets exceedingly diminished. At the critical point itself, the correlation length becomes infinite and, with it, the system becomes totally insensitive to such a transformation. The *fixed point* of the transformation is then identified with the critical point of the system, and the behavior of the svstem parameters such as K and h, see eqn. (12.4.2), in the neighborhood of the fixed point determines the critical exponents, etc. Since very few systems could be solved exactly, approximation procedures had to be developed to handle most of the cases under study. One such procedure starts with known results for the upper critical dimension $d = 4$ and carries out expansions in terms of the (small) parameter $\varepsilon = 4 - d$, while the other starts with known results for the spherical model ($n = \infty$) and carries out expansions in terms of the (small) parameter $1/n$; in the former case, the coefficients of the expansion would be n-dependent. while in the latter case they would be d-dependent. Highly sophisticated manipulations enable one to obtain reliable results for ε as large as 1 (which corresponds to the most practical dimension $d = 3$) and for n as small as 1 (which corresponds to a large variety of systems that can be described through an order parameter that is scalar). We propose to discuss this method at length in Chapter 13; for the present, suffice it to say that the reader interested in details may refer to vol. 6 of the Domb–Green series, which is devoted entirely to this topic.

(iii) *The Monte Carlo methods.* As the name implies, these methods employ "random numbers" to simulate statistical fluctuations in order to carry out numerical integrations and computer simulations of systems with large numbers of degrees of freedom. Such methods have been in vogue for quite some time and, fortunately, have kept pace with developments along other lines of approach–so much so that they have adapted themselves to the ideas of the renormalization group as well (see Ma, 1976b). We do not propose to discuss these methods here; the interested reader may refer to Binder (1986, 1987, 1992).

As mentioned earlier, the aforementioned methods lead to results that are mutually compatible and, within the stated margins of error, essentially identical. Table 12.1 lists the generally accepted values of the critical exponents of a *three-dimensional* system with $n = 0, 1, 2, 3$ and ∞. It includes all the major exponents except α, which can be determined by using the scaling relation $\alpha + 2\beta + \gamma = 2$ (or the hyperscaling relation $d\nu = 2 - \alpha$); we thus obtain, for $n = 0$, 1, 2, 3 and $\infty, \alpha = 0.235$, 0.111, −0.008, −0.114 and −1, respectively–of course, with appropriate margins of error. The theoretical results assembled here may be compared with the corresponding experimental ones listed in Table 11.1, remembering that, while all other entries there are Ising-like ($n = 1$), the case of superfluid He^4 pertains to $n = 2$.

TABLE 12.1. THEORETICAL VALUES OF THE CRITICAL EXPONENTS IN THREE DIMENSIONS (AFTER BAKER, 1990)

	$n=0$	$n=1$	$n=2$	$n=3$	$n=\infty$
β	0.302 ± 0.004	0.324 ± 0.006	0.346 ± 0.009	0.362 ± 0.012	0.5
γ	1.161 ± 0.003	1.241 ± 0.004	1.316 ± 0.009	1.39 ± 0.01	2.0
δ	4.85 ± 0.08	4.82 ± 0.06	4.81 ± 0.08	4.82 ± 0.12	5.0
ν	0.588 ± 0.001	0.630 ± 0.002	0.669 ± 0.003	0.705 ± 0.005	1.0
η	0.026 ± 0.014	0.031 ± 0.011	0.032 ± 0.015	0.031 ± 0.022	0.0

It will be noticed that Table 12.1 includes exponents for the case $n = 0$ as well. This relates to the fact that if the spin dimensionality n is treated as a continuously varying parameter then the limit $n \to 0$ corresponds to the statistical behavior of self-avoiding random walks and hence of polymers (de Gennes, 1972; des Cloizeaux, 1974). The role of t in that case is played by the parameter $1/N$ where N is the number of steps constituting the walk or the number of monomers constituting the polymer chain; thus, the limit $N \to \infty$ corresponds to $t \to 0$, whereby one approaches the critical point of the system. Concepts such as the correlation function, the correlation length, the susceptibility, the free energy, etc. can be introduced systematically into the problem and one obtains a well-defined model that fits neatly with the rest of the family. For details, see de Gennes (1979) and des Cloizeaux (1982).[8]

Finally, we look at the situation with $d \geq 4$. As mentioned earlier, especially towards the end of Sec. 12.4, critical exponents for $d > 4$ are independent of n and have values as predicted by the mean field theory. To recapitulate, they are:

$$\left.\begin{array}{c} \alpha = 0, \quad \beta = \tfrac{1}{2}, \quad \gamma = 1, \quad \delta = 3 \\ \nu = \tfrac{1}{2}, \quad \eta = 0. \end{array}\right\} \tag{6}$$

At the borderline dimensionality 4, two nonclassical features appear. First, the nature of the singularity is such that it cannot be represented by a power law alone; logarithmic factors are also present. Second, the dependence on n shows up in a striking fashion. We simply quote the results [for details, see Brézin *et al.* (1976)]:

$$c^{(s)} \sim |\ln t|^{(4-n)/(n+8)} \qquad (h = 0, t \gtrsim 0) \tag{7}$$

$$m_0 \sim |t|^{1/2} |\ln |t||^{3/(n+8)} \qquad (h = 0, t \lesssim 0) \tag{8}$$

$$\chi \sim t^{-1} |\ln t|^{(n+2)/(n+8)} \qquad (h = 0, t \gtrsim 0) \tag{9}$$

$$h \sim m^3 |\ln m|^{-1} \qquad (t = 0, h \gtrsim 0), \tag{10}$$

along with the fact that $\eta = 0$ and hence $\xi \sim \chi^{1/2}$. In the limit $n \to \infty$, these results go over to the ones pertaining to the spherical model; see Sec. 12.4.

Problems

12.1. Making use of expressions (11.3.17–19) and (12.1.12–13), show that the expectation values of the numbers N_+, N_-, N_{++}, N_{--} and N_{+-} in the case of an Ising chain are

$$\overline{N}_\pm = N \frac{P(\beta, B) \pm \sinh(\beta\mu B)}{2P(\beta, B)},$$

$$\overline{N}_{++} = \frac{N}{2D(\beta, B)} e^{\beta\mu B}[P(\beta, B) + \sinh(\beta\mu B)],$$

$$\overline{N}_{--} = \frac{N}{2D(\beta, B)} e^{-\beta\mu B}[P(\beta, B) - \sinh(\beta\mu B)]$$

and

$$\overline{N}_{+-} = \frac{N}{D(\beta, B)} e^{-4\beta J},$$

where

$$P(\beta, B) = \{e^{-4\beta J} + \sinh^2(\beta\mu B)\}^{1/2}$$

and

$$D(\beta, B) = P(\beta, B)[P(\beta, B) + \cosh(\beta\mu B)].$$

Check that (i) the above expressions satisfy the requirement that $\overline{N}_{+-} + \overline{N}_{--} + \overline{N}_{+-} = N$ and (ii) they agree with the quasi-chemical approximation (11.6.22), regardless of the value of B.

12.2. (a) Show that the partition function of an Ising lattice can be written as

$$Q_N(B, T) = \sum_{N_+, N_{+-}}{}' g_N(N_+, N_{+-}) \exp\{-\beta H_N(N_+, N_{+}\ )\},$$

where

$$H_N(N_+, N_{+-}) = -J\left(\tfrac{1}{2}qN - 2N_{+-}\right) - \mu B(2N_+ - N),$$

while the other symbols have their usual meanings; cf. eqns (11.3.19, 20).

(b) Next, determine the combinatorial factor $g_N(N_+, N_{+-})$ for an Ising chain ($q = 2$) and show that, asymptotically,

$$\begin{aligned}\ln g_N(N_-, N_{+-}) \approx\ & N_+ \ln N_+ + (N - N_+)\ln(N - N_+) \\ & - (N_+ - \tfrac{1}{2}N_{+-})\ln(N_+ - \tfrac{1}{2}N_{+-}) \\ & - (N - N_+ - \tfrac{1}{2}N_{+-})\ln(N - N_+ - \tfrac{1}{2}N_{+-}) \\ & - 2(\tfrac{1}{2}N_{+-})\ln(\tfrac{1}{2}N_{+-}).\end{aligned}$$

Now, assuming that $\ln Q_V \approx$ (the logarithm of the *largest* term in the sum $\sum'$), evaluate the Helmholtz free energy $A(B, T)$ of the system and show that this leads to precisely the same results as the ones quoted in the preceding problem as well as the ones obtained in Sec. 12.1.

12.3. Using the approximate expression, see Fowler and Guggenheim (1940),

$$g_N(N_1, N_{12}) \simeq \frac{\left(\tfrac{1}{2}qN\right)!}{N_{11}!N_{22}!\left[\left(\tfrac{1}{2}N_{12}\right)!\right]^2}\left(\frac{N_1!N_2!}{N!}\right)^{q-1},$$

for evaluating the partition function of an Ising lattice, show that one is led to the same results as the ones following from the Bethe approximation.

[Note that, for $q = 2$, the quantity $\ln g$ here is *asymptotically* exact; cf. Problem 12.2(b). No wonder that the Bethe approximation gives exact results in the case of an Ising chain.]

12.4. Making use of relation (12.1.36), along with expressions (12.1.8) for the eigenvalues λ_1 and λ_2 of the transfer matrix P, determine the correlation length $\xi(B, T)$ of the Ising chain in the presence of a magnetic field. Evaluate the critical exponent ν^c, as defined in Problem 11.23, and check that $\nu^c = \nu/\Delta$, where ν and Δ are standard exponents defined in Secs 11.10 and 11.12.

12.5. Consider a one-dimensional Ising system in a *fluctuating* magnetic field B, so that

$$Q_N(B, T) \sim \int_{-\infty}^{\infty} dB \sum_{\{\sigma_i\}} \exp\left\{-\frac{\beta N B^2}{2s} + \sum_{i=1}^{N}[\beta\mu B\sigma_i + \beta J\sigma_i\sigma_{i+1}]\right\},$$

with $\sigma_{N+1} = \sigma_1$. Note that when the system is very large (i.e. $N \gg 1$) the typical value of B is very small; nevertheless, the presence of this small fluctuating field leads to an order–disorder transition in this one-dimensional system! Determine the critical temperature of this transition.

12.6. Solve exactly the problem of a field-free Ising chain with nearest-neighbor and next-nearest-neighbor interactions, so that

$$H\{\sigma_i\} = -J_1 \sum_i \sigma_i \sigma_{i+1} - J_2 \sum_i \sigma_i \sigma_{i+2},$$

and examine the various properties of interest of this model.

[*Hint*: Introduce a new variable $\tau_i = \sigma_i \sigma_{i+1} = \pm 1$, with the result that

$$H\{\tau_i\} = -J_1 \sum_i \tau_i - J_2 \sum_i \tau_i \tau_{i+1},$$

which is formally similar to expression (12.1.1)].

12.7. Consider a double Ising chain such that the nearest-neighbor coupling constant along either chain is J_1 while the one linking adjacent spins in the two chains is J_2. Then, in the absence of the field,

$$H\{\sigma_i, \sigma_i'\} = -J_1 \sum_i (\sigma_i \sigma_{i+1} + \sigma_i' \sigma_{i+1}') - J_2 \sum_i \sigma_i \sigma_i'.$$

Show that the partition function of this system is given by

$$\frac{1}{2N} \ln Q \approx \frac{1}{2} \ln [2 \cosh K_2 \{\cosh 2K_1 + \sqrt{(1 + \sinh^2 2K_1 \tanh^2 K_2)}\}],$$

where $K_1 = \beta J_1$ and $K_2 = \beta J_2$. Examine the various thermodynamic properties of this system.

[*Hint*: Express the Hamiltonian H in a *symmetric* form by writing the last term as $-\frac{1}{2} J_2 \Sigma_i (\sigma_i \sigma_i' + \sigma_{i+1} \sigma_{i+1}')$ and use the transfer matrix method.]

12.8. Write down the transfer matrix P for a one-dimensional spin-1 Ising model in zero field, described by the Hamiltonian

$$H_N\{\sigma_i\} = -J \sum_i \sigma_i \sigma_{i+1} \qquad \sigma_i = -1, 0, +1.$$

Show that the free energy of this model is given by

$$\frac{1}{N} A(T) = -kT \ln \left\{ \frac{1}{2} \left[(1 + 2\cosh \beta J) + \sqrt{\{8 + (2 \cosh \beta J - 1)^2\}} \right] \right\}.$$

Examine the limiting behavior of this quantity in the limits $T \to 0$ and $T \to \infty$.

12.9. (a) Apply the theory of Sec. 12.1 to a one-dimensional lattice gas and show that the pressure P and the volume per particle v are given by

$$\frac{P}{kT} = \ln \left[\frac{1}{2} \left\{ (y + 1) + \sqrt{[(y - 1)^2 + 4y\eta^2]} \right\} \right]$$

and

$$\frac{1}{v} = \frac{1}{2} \left[1 + \frac{y - 1}{\sqrt{[(y - 1)^2 + 4y\eta^2]}} \right]$$

where $y = z \exp(4\beta J)$ and $\eta = \exp(-2\beta J)$, z being the fugacity of the gas. Examine the high and the low temperature limits of these expressions.

(b) A *hard-core* lattice gas pertains to the limit $J \to -\infty$; this makes $y \to 0$ and $\eta \to \infty$ such that the quantity $y\eta^2$, which is equal to z, stays finite. Show that this leads to the equation of state

$$\frac{P}{kT} = \ln \left(\frac{1 - \rho}{1 - 2\rho} \right) \quad \left(\rho = \frac{1}{v} \right).$$

12.10. For a one-dimensional system, such as the ones discussed in Secs 12.1 and 12.2, the correlation function $g(r)$ at all temperatures is of the form $\exp(-ra/\xi)$, where a is the lattice constant of the system. Using the fluctuation-dissipation relation (11.11.11), show that the low-field susceptibility of such a system is given by

$$\chi_0 = N\beta\mu^2 \coth(a/2\xi).$$

Note that as $T \to 0$ and, along with it, $\xi \to \infty$, χ_0 becomes directly proportional to ξ–consistent with the fact that the critical exponent $\eta = 1$.

For an n-vector model (including the scalar case $n = 1$), ξ is given by eqn. (12.2.17), which leads to the result

$$\chi_0 = N\beta\mu^2 \frac{I_{(n-2)/2}(\beta J) + I_{n/2}(\beta J)}{I_{(n-2)/2}(\beta J) - I_{n/2}(\beta J)}.$$

Check that for the special case $n = 1$ this result reduces to eqn. (12.1.14).

12.11. Show that for a one-dimensional, field-free Ising model

$$\overline{\sigma_k\sigma_l\sigma_m\sigma_n} = \{\tanh \beta J\}^{n-m+l-k},$$

where $k \le l \le m \le n$.

12.12. Recall the symbol $n(r)$, of eqn. (12.3.5), which denotes the "number of *closed graphs* that can be drawn on a given lattice using exactly r bonds". Show that for a square lattice wrapped on a torus (which is equivalent to imposing periodic boundary conditions)

$$n(4) = N, \quad n(6) = 2N, \quad n(8) = \tfrac{1}{2}N^2 + \tfrac{9}{2}N, \cdots$$

Substituting these numbers into eqn. (12.3.5) and taking logs, one gets

$$\ln Q(N, T) = N\left\{\ln(2\cosh^2 K) + v^4 + 2v^6 + \tfrac{9}{2}v^8 + \cdots\right\}, v = \tanh K.$$

Note that the term in N^2 has disappeared–in fact, all higher powers of N do the same. Why?

12.13. According to Onsager, the field-free partition function of a rectangular lattice (with interaction parameters J and J' in the two perpendicular directions) is given by

$$\frac{1}{N}\ln Q(T) = \ln 2 + \frac{1}{2\pi^2}\int_0^{\pi}\int_0^{\pi} \ln\{\cosh(2\gamma)\cosh(2\gamma') - \sinh(2\gamma)\cos\omega - \sinh(2\gamma')\cos\omega'\}\,d\omega\,d\omega',$$

where $\gamma = J/kT$ and $\gamma' = J'/kT$. Show that if $J' = 0$, this leads to expression (12.1.9) for the linear chain with $B = 0$ while if $J' = J$, one is led to expression (12.3.22) for the square net. Locate the critical point of the rectangular lattice and study its thermodynamic behavior in the neighborhood of that point.

12.14. Write the elliptic integral $K_1(\kappa)$ in the form

$$K_1(\kappa) = \int_0^{\pi/2} \frac{1 - \kappa\sin\phi}{\sqrt{(1 - \kappa^2\sin^2\phi)}}\,d\phi + \int_0^{\pi/2} \frac{\kappa\sin\phi}{\sqrt{(1 - \kappa^2\sin^2\phi)}}\,d\phi.$$

Show that, as $\kappa \to 1-$, the first integral $\to \ln 2$ while the second $\approx \ln[2/(1-\kappa^2)^{1/2}]$. Hence $K_1(\kappa) \approx \ln(4/|\kappa'|)$, as in eqn. (12.3.34).

[*Hint*: In the second integral, substitute $\cos\phi = x$.]

12.15. Using eqns (12.3.22) and (12.3.28) at $T = T_c$, show that the entropy of the two-dimensional Ising model on a square lattice at its critical point is given by

$$\frac{S_c}{Nk} = \frac{2G}{\pi} + \frac{1}{2}\ln 2 - \sqrt{2}K_c \simeq 0.3065;$$

here, G is Catalan's constant which is approximately equal to 0.915966. Compare this result with the ones following from the Bragg–Williams approximation and from the Bethe approximation; see Problem 11.15.

12.16. The spontaneous magnetization of a two-dimensional Ising model on a square lattice at $T < T_c$ is given by eqn. (12.3.38a). Express this result in the form $B|t|^\beta\{1 + b|t| + \cdots\}$, where B and β are stated in eqn. (12.3.40) while $b = (1 - 9K_c/\sqrt{2})/8$. As usual, $t = (T - T_c)/T_c < 0$ and $|t| \ll 1$.

12.17. Apply the theory of Sec. 12.3 to a two-dimensional lattice gas and show that, at $T = T_c$, the quantity $kT_c/P_cv_c \simeq 10.35$.

12.18. Show that for the spherical model in one dimension the free energy *at constant* λ is given by

$$\frac{\beta A_\lambda}{N} = \frac{1}{2}\ln\left[\frac{\beta\{\lambda + \sqrt{(\lambda^2 - J^2)}\}}{2\pi}\right] - \frac{\beta\mu^2B^2}{4(\lambda - J)},$$

while λ is determined by the constraint equation

$$\frac{1}{2\beta\sqrt{(\lambda^2-J^2)}}+\frac{\mu^2B^2}{4(\lambda-J)^2}=1.$$

In the absence of the field ($B=0$), $\lambda=\sqrt{(1+4\beta^2J^2)}+/2\beta$; the free energy *at constant* $\mathcal{J}$is then given by

$$\frac{\beta A_{\mathcal{J}}}{N}=\frac{1}{2}\ln\left[\frac{\sqrt{(1+4\beta^2J^2)}+1}{4\pi}\right]-\frac{1}{2}\sqrt{(1+4\beta^2J^2)}.$$

12.19. Starting with expression (12.2.8) for the partition function of a one-dimensional n-vector model, with $J_i=nJ'$, show that

$$\operatorname*{Lim}_{n,N\to\infty}\frac{1}{nN}\ln Q_N(nK)=\frac{1}{2}\left[\sqrt{(4K^2+1)}-1-\ln\left\{\frac{\sqrt{(4K^2+1)}+1}{2}\right\}\right],$$

where $K=\beta J'$. Note that, apart from a constant term, this result is exactly the same as for the spherical model; the difference arises from the fact that the present result is normalized to give $Q_N=1$ when $K=0$.

Hint: Use the asymptotic formulae (for $\nu\gg1$)

$$\Gamma(\nu)\approx(2\pi/\nu)^{1/2}(\nu/e)^{\nu}$$

and

$$I_\nu(\nu z)\approx(2\pi\nu)^{-1/2}(z^2+1)^{-1/4}e^{\nu\eta},$$

where

$$\eta=\sqrt{(z^2+1)}-\ln[\{\sqrt{(z^2+1)}+1\}/z].$$

12.20. Show that the low-field susceptibility, χ_0, of the spherical model at $T<T_c$ is given by the asymptotic expression

$$\chi_0\approx(N\mu^2/k_BT)\cdot Nm_0^2(T),$$

where $m_0(T)$ is the spontaneous magnetization of the system; note that in the thermodynamic limit the reduced susceptibility, $k_BT\chi_0/N\mu^2$, is infinite at all $T<T_c$. Compare Problem 12.26.

12.21. In view of the fact that only those fluctuations whose length scale is large play a dominant role in determining the nature of a phase transition, the quantity $(\lambda-\mu_{\mathbf{k}})$ in the expression for the correlation function of the spherical model, see eqn. (12.4.57), may be replaced by

$$\lambda-\mu_{\mathbf{k}}=\lambda-J\sum_{j=1}^{d}\cos(k_ja)\simeq J\left[\phi+\frac{k^2a^2}{2}\right],$$

where $\phi=(\lambda/J)-d$. Show that this approximation leads to the same result for $G(\boldsymbol{R})$ as we have in eqn. (12.4.61), with the same ξ as in eqn. (12.4.62).

For a similar reason, the quantity $[\exp(\alpha+\beta\varepsilon)-1]$ in the correlation function (12.5.33) of the ideal Bose gas may be replaced by

$$e^{\alpha+\beta\varepsilon}-1\simeq\alpha+\beta\varepsilon=\alpha+(\beta\hbar^2/2m)k^2,$$

leading to the same $G(\boldsymbol{R})$ as in eqn. (12.5.35), with the same ξ as in eqn. (12.5.36).[9]

12.22. Consider a spherical model whose spins interact through a *long-range* potential varying as $(a/r)^{d+\sigma}(\sigma>0)$, r being the distance between two spins. This replaces the quantity $(\lambda-\mu_{\mathbf{k}})$ of eqns (12.4.16) and (12.4.57) by an expression approximating $J[\phi+\frac{1}{2}(ka)^\sigma]$ for $\sigma<2$ and $J[\phi+\frac{1}{2}(ka)^2]$ for $\sigma>2$; note that the nearest-neighbor interaction corresponds to the limit $\sigma\to\infty$ and hence to the latter case.

Assuming σ to be less than 2, show that the above system undergoes a phase transition at a *finite* temperature T_c for all $d>\sigma$. Further show that the critical exponents for this model are

$$\alpha=\frac{d-2\sigma}{d-\sigma},\quad\beta=\frac{1}{2},\quad\gamma=\frac{\sigma}{d-\sigma},\quad\delta=\frac{d+\sigma}{d-\sigma},$$

$$\nu=\frac{1}{d-\sigma},\quad\eta=2-\sigma$$

for $\sigma < d < 2\sigma$, and

$$\alpha = 0, \quad \beta = \tfrac{1}{2}, \quad \gamma = 1, \quad \delta = 3, \quad \nu = \frac{1}{\sigma}, \quad \eta = 2 - \sigma$$

for $d > 2\sigma$.

12.23. Refer to Sec. 12.5 on the ideal Bose gas in d dimensions, and complete the steps leading to eqns (12.5.9)–(12.5.15) and (12.5.23).

12.24. Show that for an ideal Bose gas in d dimensions and at $T > T_c$

$$V\left(\frac{\partial^2 P}{\partial T^2}\right)_v = \frac{Nk_B}{T}\left[\frac{d(d+2)}{4}\frac{g_{(d+2)/2}(z)}{g_{d/2}(z)} - \frac{d}{2}\frac{g_{d/2}(z)}{g_{(d-2)/2}(z)} - \frac{d^2}{4}\frac{\{g_{d/2}(z)\}^2 g_{(d-4)/2}(z)}{\{g_{(d-2)/2}(z)\}^3}\right]$$

and

$$\left(\frac{\partial^2 \mu}{\partial T^2}\right)_v = \frac{k_B}{T}\left[\frac{d(d-2)}{4}\frac{g_{d/2}(z)}{g_{(d-2)/2}(z)} - \frac{d^2}{4}\frac{\{g_{d/2}(z)\}^2 g_{(d-4)/2}(z)}{\{g_{(d-2)/2}(z)\}^3}\right]$$

where $\mu(= kT \ln z)$ is the chemical potential of the gas while other symbols have the same meanings as in Sec. 12.5. Note that these quantities satisfy the thermodynamic relationship

$$VT\left(\frac{\partial^2 P}{\partial T^2}\right)_v - NT\left(\frac{\partial^2 \mu}{\partial T^2}\right)_v = C_V.$$

Also note that, since P here equals $(2U/dV)$, the quantities $(\partial P/\partial T)_v$ and $(\partial^2 P/\partial T^2)_v$ are directly proportional to C_V and $(\partial C_V/\partial T)$, respectively. Finally, examine the singular behavior of these quantities as $T \to T_c$ from above.

12.25. Show that for any given fluid

$$C_P = VT(\partial P/\partial T)_S(\partial P/\partial T)_V \kappa_T$$

and

$$C_V = VT(\partial P/\partial T)_S(\partial P/\partial T)_V \kappa_S,$$

where the various symbols have their usual meanings. In the two-phase region, these formulae take the form

$$C_P = VT(dP/dT)^2\kappa_T \quad \text{and} \quad C_V = VT(dP/dT)^2\kappa_S,$$

respectively. Using the last of these results, rederive eqn. (12.5.30) for C_V at $T < T_c$.

12.26. Show that for any given fluid

$$\kappa_T = \rho^{-2}(\partial\rho/\partial\mu)_T,$$

where $\rho(= N/V)$ is the particle density and μ the chemical potential of the fluid. For the ideal Bose gas at $T < T_c$,

$$\rho = \rho_0 + \rho_e \approx -\frac{k_B T}{V\mu} + \frac{\zeta(d/2)}{\lambda^d}.$$

Using these results, show that[10]

$$\kappa_T \approx (V/k_B T)(\rho_0/\rho)^2 \quad (T < T_c);$$

note that in the thermodynamic limit the reduced compressibility, $k_B T\kappa_T/v$, is infinite at all $T < T_c$. Compare Problem 12.20.

12.27. Consider an ideal relativistic Bose gas composed of N_1 particles and N_2 antiparticles, each of rest mass m_0, with occupation numbers

$$\frac{1}{\exp[\beta(\varepsilon - \mu_1)] - 1} \quad \text{and} \quad \frac{1}{\exp[\beta(\varepsilon - \mu_2)] - 1},$$

respectively, and the energy spectrum $\varepsilon = c\sqrt{(p^2 + m_0^2 c^2)}$. Since particles and antiparticles are supposed to be created in pairs, the system is constrained by the conservation of the number $Q(= N_1 - N_2)$, rather than of N_1 and N_2 separately; accordingly, $\mu_1 = -\mu_2 = \mu$, say.

Set up the thermodynamic functions of this system in three dimensions and examine the onset of Bose–Einstein condensation as T approaches a critical value, T_c, determined by the "number density" Q/V. Show that the nature of the singularity at $T = T_c$ is such that, regardless of the severity of the relativistic effects, this system falls in the *same* universality class as the nonrelativistic one.

Notes

[1] Note that, since $T_c = 0$ here, the assertion that the free energy function $\psi \equiv (A/NkT) \sim (T - T_c)^{2-\alpha}$ implies that, near $T = T_c$ in the present case, $A \sim T^{3-\alpha}$; accordingly, the specific heat $C_0 \sim T^{2-\alpha}$. Comparison with expression (19a) would be inappropriate because C_0, in the limit $T \to 0$, cannot be nonzero; the result quoted in (19a) is an artifice of the model considered, which must somehow be "subtracted away". The next approximation yields a result proportional to the *first* power in T, giving $\alpha = 1$. For a parallel situation, see eqn. (12.4.35) for the spherical model (which pertains to the case $n = \infty$).

[2] Note, however, that in order to calculate the field-dependent susceptibility, $(\partial\overline{M}/\partial B)_{T,r}$, subject to the spherical constraint (19), one must keep in mind the field dependence of λ while differentiating (20a).

[3] We denote Boltzmann's constant by the symbol k_B so as to avoid confusion with the wave number k.

[4] Note that our definition of the function $W_d(\phi)$ differs slightly from the one adopted by Barber and Fisher (1973); this difference arises from the fact that our J is *twice*, and our ϕ is *one-half*, of theirs.

[5] This is reminiscent of a similar problem, and a similar *ad hoc* solution, encountered in the study of Bose–Einstein condensation in Sec. 7.1; see also Sec. 12.5.

[6] We equate this exponent with $(1 + \alpha)$ because if $C_V \sim t^{-\alpha}$, then $\partial C_V/\partial t \sim t^{-\alpha-1}$. We hasten to add that the critical exponent α should not be confused with the physical quantity denoted by the same symbol, viz. $\alpha(= -\ln z)$, which was introduced just before eqns (16) and has been used throughout this section.

[7] For a review of this theorem and other allied questions, see Griffiths (1972), pp. 84–89.

[8] Values of n other than the ones appearing in Table 12.1 are sometimes encountered. For instance, certain antiferromagnetic order–disorder transitions require for their description an order parameter with $n = 4$, 6 8 or 12; see Mukamel (1975), and Bak, Krinsky and Mukamel (1976). Another example of this is provided by the superfluidity of liquid He^3 which seems to require an order parameter with $n = 18$; see, for instance, Anderson (1984), and Vollhardt and Wölfle (1990).

[9] A comparison with the mean field results (11.11.25, 26) brings out a close similarity that exists between these models and the mean field picture of a phase transition; for instance, they all share a common critical exponent η, which is zero. There are, however, significant differences; for one, the correlation length ξ for these models is characterized by a critical exponent ν which is *nonclassical* – in the sense that it is d-dependent whereas in the mean field case it is independent of d.

[10] This remarkable relationship between the isothermal compressibility of a *finite-sized* Bose gas and the condensate density in the corresponding *bulk* system was first noticed by Singh and Pathria (1987b).

CHAPTER 13

PHASE TRANSITIONS: THE RENORMALIZATION GROUP APPROACH

IN THIS chapter we propose to discuss what has turned out to be the most successful approach to the problem of phase transitions. This approach is based on ideas first propounded by Kadanoff (1966b) and subsequently developed by Wilson (1971) and others into a powerful calculational tool. The main point of Kadanoff's argument is that, as the critical point of a system is approached, its correlation length becomes exceedingly large–with the result that the sensitivity of the system to a *length transformation* (or a *change of scale*, as one may call it) gets exceedingly diminished. At the critical point itself, the correlation length becomes infinitely large and with it the system becomes totally insensitive to such a transformation. It is then conceivable that, if one is not too far removed from the critical point (i.e. $|t|, h \ll 1$), the given system (with lattice constant a) may bear a close resemblance to a transformed system (with lattice constant $a' = la$, where $l > 1$, and presumably modified parameters t' and h'), renormalized so that all distances in it are measured in terms of the new lattice constant a'; clearly, the rescaled correlation length ξ' (in units of a') would be one-lth of the original correlation length ξ (in units of a). This resemblance in respect of critical behavior is expected only if ξ' is also much larger than a', just as ξ was in comparison with a, which in turn requires that $l \ll (\xi/a)$; by implication, $|t'|$, h' would also be $\ll 1$. These considerations lead to a formulation similar to the one presented in Sec. 11.10, with the difference that, while there we had to rely on a (scaling) *hypothesis*, here we have a convincing argument based on the role played by correlations among the microscopic constituents of the system which, in the vicinity of the critical point, are so large-scale that they make all other length scales, including the one that determines the structure of the lattice, essentially irrelevant. Unfortunately, Kadanoff's approach did not provide a systematic means of deriving the critical exponents or of constructing the scaling functions that appear in the formulation. Those deficiencies were remedied by Wilson by introducing the concept of a renormalization group (RG) into the theory.

We propose to discuss Wilson's approach in Secs 13.3 and 13.4, but first we present a formulation of scaling along the lines indicated above and follow it with an exploration of simple examples of renormalization which pave the way for establishing Wilson's theory. Finally, in Sec. 13.5 we outline the theory of finite-sizing scaling which too has benefitted greatly from the RG approach.

13.1. The conceptual basis of scaling

The scale change in a given system can be effected in several different ways, the earliest one being due to Kadanoff who suggested that, when large-scale correlations prevail, the individual spins in the system may be grouped into "blocks of spins", each block consisting of l^d original spins, and let each block play the role of a "single spin" in the transformed system; see Fig. 13.1, where a block–spin transformation with $l = 2$ and $d = 2$ is shown. The spin variable of a block may be denoted by the symbol σ', which arises from the values of the individual spins in the block, but a rule has to be established so that σ' too is either $+1$ or -1, just as the original σ_i were.[1] The transformed system then consists of N' "spins", where

$$N' = l^{-d}N, \tag{1}$$

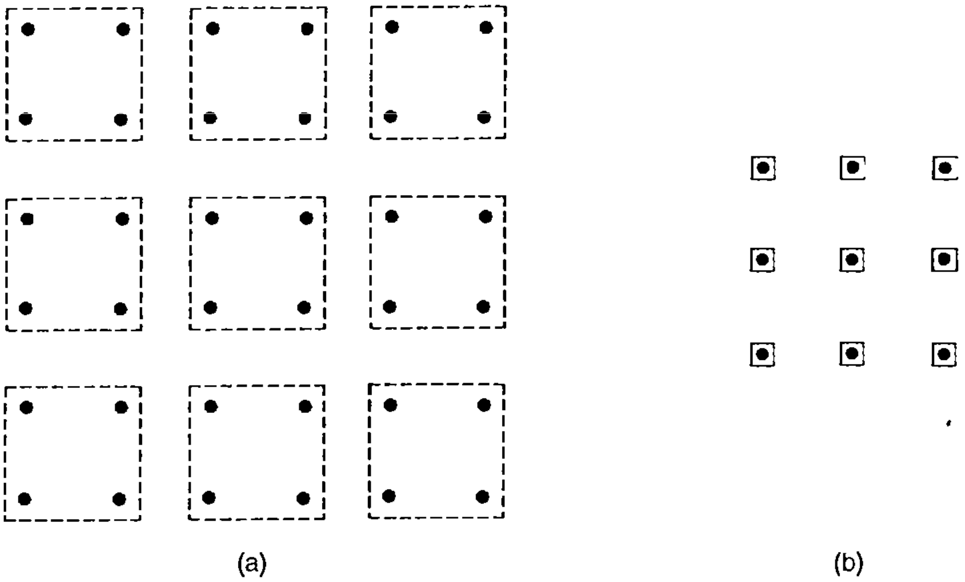


FIG. 13.1. A block–spin transformation, with $l = 2$ and $d = 2$. The original lattice has $N(= 36)$ spins, the transformed one has $N'(= 9)$; after rescaling, the latter looks very much the same as the former, especially in the limit $N, N' \to \infty$.

occupying a lattice structure with lattice constant $a' = la$. In order to preserve the spatial density of the degrees of freedom in the system, all spatial distances must be rescaled by the factor l, so that for any two "spins" in the transformed system

$$\boldsymbol{r}' = l^{-1}\boldsymbol{r}. \tag{2}$$

A second way of effecting a scale change is to write down the partition function of the system,

$$Q = \sum_{\{\sigma_i\}} \exp\left[-\beta H_N\{\sigma_i\}\right], \tag{3}$$

and carry [illegible] ummation over a subset of $(N - N')$ spins, such that one is left with a summation, over the remaining N' spins, which can (hopefully) be expressed in a form similar to (3), namely

$$Q = \sum_{\{\sigma_i'\}} \exp[-\beta H_{N'}\{\sigma_i'\}]. \tag{4}$$

If the desired passage, from expression (3) to (4), can be accomplished with some degree of accuracy, we should expect a close resemblance between the critical behavior of the original system represented by eqn. (3) and the transformed one represented by (4); see Fig. 13.2, where an example of this procedure with $l = \sqrt{2}$ and $d = 2$ is shown. We note that this procedure forms the very backbone of the Wilson approach and is generally referred to as "decimation", though the fraction of the spins removed, $(N - N')/N$, is rarely equal to 1/10. Other ways of effecting a scale change will be mentioned later.

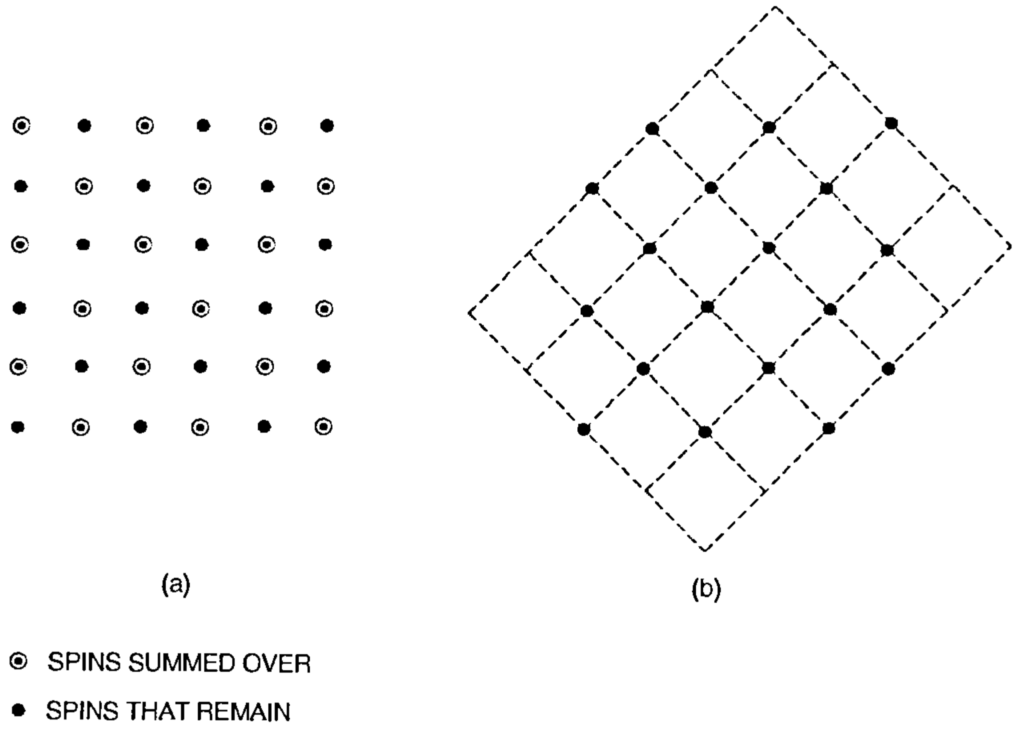

(a) (b)

FIG. 13.2. A scale transformation by "decimation", with $l = \sqrt{2}$ and $d = 2$. The original lattice has $N(= 36)$ spins, the transformed one has $N'(= 18)$; the latter is yet to be rescaled (and rotated through an angle $\pi/4$) so that it looks very much the same as the former, especially in the limit $N, N' \to \infty$.

It is quite natural to expect that the free energy of the transformed system (or, at least, that part of it which determines the critical behavior) is the same as that of the original system. The singular parts of the free energy *per spin* of the two systems should, therefore, be related as

$$N'\psi^{(s)}(t', h') = N\psi^{(s)}(t, h), \tag{5}$$

so that

$$\psi^{(s)}(t, h) = l^{-d}\psi^{(s)}(t', h'). \tag{6}$$

Now, since both t and t' are small in magnitude, one may assume that they are linearly related, i.e.

$$t' = l^{y_t} t, \tag{7}$$

where y_t, as yet, is an unknown number. Similarly, one may assume that

$$h' = l^{y_h} h, \tag{8}$$

with the result that

$$\psi^{(s)}(t, h) = l^{-d} \psi^{(s)}(l^{y_t} t, l^{y_h} h); \tag{9}$$

like y_t, the number y_h is also unknown at this stage of the game.

We now assert that the function $\psi^{(s)}$, which governs much of the critical behavior of the system, is essentially insensitive to a change of scale; we should, therefore, be able to eliminate the scale factor l from expression (9). This essentially forces us to replace the variables t' and h' by a single, l-independent variable

$$\frac{h'}{|t'|^{y_h/y_t}} = \frac{h}{|t|^{y_h/y_t}} = \frac{h}{|t|^{\Delta}}, \text{ say } \left(\Delta = \frac{y_h}{y_t}\right); \tag{10}$$

at the same time, it requires us to write

$$\psi^{(s)}(t', h') = |t'|^{d/y_t} \tilde{\psi}(h'/|t'|^{\Delta}), \tag{11}$$

leading to the identical result

$$\psi^{(s)}(t, h) = |t|^{d/y_t} \tilde{\psi}(h/|t|^{\Delta}); \tag{12}$$

note that, as of now, the function $\tilde{\psi}$ is also unknown.[2] Comparing (12) with eqn. (11.10.7), we readily identify the critical exponent α as

$$\alpha = 2 - (d/y_t); \tag{13}$$

more importantly, the present considerations have led to the same scaled form for the free energy density of the system as was hypothesized in Sec. 11.10. We have thus raised the status of expression (11.10.7) from being a mere hypothesis to being a well-founded result whose conceptual basis lies in the prevalence of large-scale correlations in the system. As in Sec. 11.10, we infer that the exponents β, γ and δ are now given by

$$\beta = 2 - \alpha - \Delta = (d - y_h)/y_t \tag{14}$$

$$\gamma = -(2 - \alpha - 2\Delta) = (2y_h - d)/y_t \tag{15}$$

and

$$\delta = \Delta/\beta = y_h/(d - y_h). \tag{16}$$

As remarked earlier, the rescaled correlation length ξ' of the transformed system and the original correlation length ξ of the given system are related as

$$\xi' = l^{-1}\xi. \tag{17}$$

At the same time, we expect ξ' to be $\sim |t'|^{-\nu}$, just as $\xi \sim |t|^{-\nu}$. It follows that

$$\left(\frac{\xi'}{\xi}\right) = \left(\frac{t'}{t}\right)^{-\nu} = l^{-\nu y_t}. \tag{18}$$

Comparing (17) and (18), we conclude that

$$\nu = 1/y_t \tag{19}$$

and hence, by (13),

$$d\nu = 2 - \alpha. \tag{20}$$

We thus obtain not only a useful expression for the critical exponent ν but also the *hyperscaling relation* (11.12.15) on a basis far sounder than the one employed in Sec. 11.12.

Finally we look at the correlation functions of the two systems. At the critical point we expect that for the transformed system

$$g(\boldsymbol{r}'_1, \boldsymbol{r}'_2) = \langle \sigma'(\boldsymbol{r}'_1)\sigma'(\boldsymbol{r}'_2)\rangle \sim (r')^{-(d-2+\eta)}, \tag{21}$$

just as for the original system

$$g(\boldsymbol{r}_1, \boldsymbol{r}_2) = \langle \sigma(\boldsymbol{r}_1)\sigma(\boldsymbol{r}_2)\rangle \sim r^{-(d-2+\eta)}. \tag{22}$$

In order that eqns (21) and (22) be mutually compatible, we must rescale the spin variables such that

$$\sigma'(\boldsymbol{r}') = l^{(d-2+\eta)/2}\sigma(\boldsymbol{r}). \tag{23}$$

As for η, we may use the scaling relation $\gamma = (2-\eta)\nu$, to get

$$\eta = d + 2 - 2y_h. \tag{24}$$

13.2. Some simple examples of renormalization

A. The Ising model in one dimension

We start with the partition function (12.1.3a) of a closed Ising chain consisting of N spins, viz.

$$Q_N(B, T) = \sum_{\{\sigma_i\}} \exp\left[\sum_{i=1}^{N}\left\{K_0 + K_1\sigma_i\sigma_{i+1} + \tfrac{1}{2}K_2(\sigma_i + \sigma_{i+1})\right\}\right]$$

$$(K_0 = 0, K_1 = \beta J, K_2 = \beta\mu B); \tag{1}$$

the parameter K_0 has been introduced here for reasons that will become clear in the sequel. For simplicity, we assume N to be even and carry out summation in (1) over all σ_i for which i is even, i.e. over $\sigma_2, \sigma_4, \ldots$; see Fig. 13.3. Writing the summand in (1) as

$$\prod_{i=1}^{N} \exp\left\{K_0 + K_1\sigma_i\sigma_{i+1} + \tfrac{1}{2}K_2(\sigma_i + \sigma_{i+1})\right\}$$

$$= \prod_{j=1}^{\frac{1}{2}N} \exp\left\{2K_0 + K_1(\sigma_{2j-1}\sigma_{2j} + \sigma_{2j}\sigma_{2j+1}) + \tfrac{1}{2}K_2(\sigma_{2j-1} + 2\sigma_{2j} + \sigma_{2j+1})\right\}, \tag{2}$$

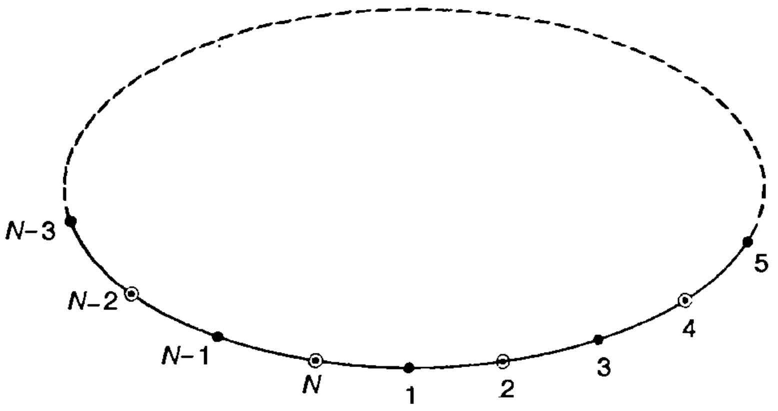


FIG. 13.3. A closed Ising chain to be "decimated" by carrying out summations over $\sigma_2, \sigma_4, \ldots$.

the summations over $\sigma_{2j} (= +1$ or $-1)$ can be carried out straightforwardly, with the result

$$\prod_{j=1}^{\frac{1}{2}N} \exp(2K_0) \cdot 2\cosh\{K_1(\sigma_{2j-1} + \sigma_{2j+1}) + K_2\} \cdot \exp\left\{\tfrac{1}{2}K_2(\sigma_{2j-1} + \sigma_{2j+1})\right\}. \tag{3}$$

Denoting σ_{2j-1} by σ'_j, the partition function Q_N assumes the form

$$Q_N(B, T) = \sum_{\{\sigma'_j\}} \prod_{j=1}^{\frac{1}{2}N} \exp(2K_0) \cdot 2\cosh\{K_1(\sigma'_j + \sigma'_{j+1}) + K_2\} \cdot \exp\left\{\tfrac{1}{2}K_2(\sigma'_j + \sigma'_{j+1})\right\}. \tag{4}$$

The crucial step now consists in expressing (4) in a form similar to (1), viz.

$$Q_N(B, T) = \sum_{\{\sigma'_j\}} \exp\left[\sum_{j=1}^{N'} \left\{K'_0 + K'_1\sigma'_j\sigma'_{j+1} + \tfrac{1}{2}K'_2(\sigma'_j + \sigma'_{j+1})\right\}\right]. \tag{5}$$

This requires that, for *all* choices of the variables σ'_j and σ'_{j+1},

$$\begin{aligned}&\exp\left\{K'_0 + K'_1\sigma'_j\sigma'_{j+1} + \tfrac{1}{2}K'_2(\sigma'_j + \sigma'_{j+1})\right\}\\ &\quad = \exp(2K_0) \cdot 2\cosh\{K_1(\sigma'_j + \sigma'_{j+1}) + K_2\} \cdot \exp\left\{\tfrac{1}{2}K_2(\sigma'_j + \sigma'_{j+1})\right\}.\end{aligned} \tag{6}$$

The various choices being $\sigma'_j = \sigma'_{j+1} = +1$, $\sigma'_j = \sigma'_{j+1} = -1$ and $\sigma'_j = -\sigma'_{j+1} = \pm 1$, we obtain from (6)

$$\exp(K'_0 + K'_1 + K'_2) = \exp(2K_0 + K_2) \cdot 2\cosh(2K_1 + K_2), \tag{7a}$$

$$\exp(K'_0 + K'_1 - K'_2) = \exp(2K_0 - K_2) \cdot 2\cosh(2K_1 - K_2) \tag{7b}$$

and

$$\exp(K'_0 - K'_1) = \exp(2K_0) \cdot 2\cosh K_2. \tag{7c}$$

Solving for K'_0, K'_1 and K'_2, we get

$$e^{K'_0} = 2e^{2K_0}\{\cosh(2K_1 + K_2)\cosh(2K_1 - K_2)\cosh^2 K_2\}^{1/4}, \tag{8a}$$

$$e^{K'_1} = \{\cosh(2K_1 + K_2)\cosh(2K_1 - K_2)/\cosh^2 K_2\}^{1/4} \tag{8b}$$

and

$$e^{K'_2} = e^{K_2}\{\cosh(2K_1 + K_2)/\cosh(2K_1 - K_2)\}^{1/2}. \tag{8c}$$

We may now remark on the need to have the parameter K_0 included in expression (1) and, accordingly, K'_0 in (5). Since we end up having three eqns (7a), (7b) and (7c), to determine the parameters appropriate to the transformed system, the problem could not be solved with K_1 and K_2 only; thus, even if K_0 were set equal to zero, a $K'_0 \neq 0$ is essential for a proper representation of the transformed system. To highlight the role played by this parameter in determining the free energy of the given system, we observe on the basis of eqns (1) and (5) that, with $K_0 = 0$,

$$Q_N(K_1, K_2) = e^{N'K'_0}Q_{N'}(K'_1, K'_2) \tag{9}$$

and hence for the free energy (in units of kT)

$$A_N(K_1, K_2) = -N'K'_0 + A_{N'}(K'_1, K'_2). \tag{10}$$

Since $N' = \frac{1}{2}N$, we obtain for the free energy *per spin* the recurrence relation

$$f(K_1, K_2) = -\tfrac{1}{2}K'_0 + \tfrac{1}{2}f(K'_1, K'_2), \tag{11}$$

which relates the free energy per spin of the given system with that of the transformed system; the role played by K'_0 is clearly significant. For example, in the limit $T \to \infty$, when (K_1, K_2) and along with them (K'_1, K'_2) tend to zero, eqns (8) and (11) give

$$f(0, 0) = -K'_0 = -\ln 2, \tag{12}$$

which is indeed the correct result (arising from the limiting value of the entropy, $Nk \ln 2$, of the system).

We note that the parameter K_0 does not appear in eqns (8b) and (8c) which determine K'_1 and K'_2 in terms of K_1 and K_2. As will be seen in Secs 13.3 and 13.4, it is these two relations that determine the singular part of the free energy of the system and hence its critical behavior; the parameters K_0 and K'_0 affect only the regular part of the free energy and hence play no direct role in determining the critical behavior of the system. We might hasten to add that, while renormalization is generally used as a technique for studying the properties of a given system in the vicinity of its critical point, it can be useful over a much broader range of the variables K_1 and K_2. For instance, in the absence of the field ($K_2 = 0$), eqns (8) and (11) give

$$K'_0 = \ln\{2\surd[\cosh(2K_1)]\}, \quad K'_1 = \ln\surd[\cosh(2K_1)], \quad K'_2 = 0 \tag{13}$$

and hence

$$f(K_1, 0) = -\tfrac{1}{2}\ln\{2\surd[\cosh(2K_1)]\} + \tfrac{1}{2}f(\ln\surd[\cosh(2K_1)], 0); \tag{14}$$

the functional n. (14) has the solution

$$f(K_1, 0) = -\ln(2\cosh K_1), \tag{15}$$

valid at *all* K_1. On the other hand, in the paramagnetic case ($K_1 = 0$), we get

$$K_0' = \ln(2\cosh K_2), \quad K_1' = 0, \quad K_2' = K_2 \tag{16}$$

and hence

$$f(0, K_2) = -\tfrac{1}{2}\ln(2\cosh K_2) + \tfrac{1}{2}f(0, K_2), \tag{17}$$

with the solution

$$f(0, K_2) = -\ln(2\cosh K_2), \tag{18}$$

valid at *all* K_2. The case when both K_1 and K_2 are present is left as an exercise for the reader; see Problem 13.2. The critical behavior of this system will be studied in Sec. 13.4.

B. The spherical model in one dimension

We adopt the same topology as in Fig. 13.3 and write down the partition function of the one-dimensional spherical model consisting of N spins, see eqn. (12.4.12a),

$$Q_N = \int_{-\infty}^{\infty}\!\!\cdots\!\int \exp\left[\sum_{i=1}^{N}\left\{K_0 + K_1\sigma_i\sigma_{i+1} + K_2\sigma_i - \Lambda\sigma_i^2\right\}\right] d\sigma_1 \ldots d\sigma_N$$
$$(K_0 = 0, K_1 = \beta J, K_2 = \beta\mu B, \Lambda = \beta\lambda), \tag{19}$$

where Λ is chosen so that

$$\left\langle \sum_{i=1}^{N} \sigma_i^2 \right\rangle = -\frac{\partial}{\partial\Lambda}\ln Q_N = N; \tag{20}$$

see eqns (12.4.9) and (12.4.13). Assuming N to be even, we carry out integrations over $\sigma_2, \sigma_4, \ldots$. For this, we write our integrand as

$$\begin{aligned}
&\prod_{i=1}^{N} \exp\{K_0 + K_1\sigma_i\sigma_{i+1} + K_2\sigma_i - \Lambda\sigma_i^2\} \\
&= \prod_{j=1}^{\frac{1}{2}N} \exp\{2K_0 + K_1(\sigma_{2j-1}\sigma_{2j} + \sigma_{2j}\sigma_{2j+1}) \\
&\qquad + K_2(\sigma_{2j-1} + \sigma_{2j}) - \Lambda(\sigma_{2j-1}^2 + \sigma_{2j}^2)\}
\end{aligned} \tag{21}$$

and integrate over σ_{2j}, using the formula

$$\int_{-\infty}^{\infty} e^{-px^2+qx}\,dx = \left(\frac{\pi}{p}\right)^{1/2} e^{q^2/4p} \quad (p > 0). \tag{22}$$

After simplification, we get

$$\prod_{j=1}^{\frac{1}{2}N}\left(\frac{\pi}{\Lambda}\right)^{1/2}\exp\left\{\left(2K_0+\frac{K_2^2}{4\Lambda}\right)+\frac{K_1^2}{2\Lambda}\sigma_{2j-1}\sigma_{2j+1}\right.$$
$$\left.+\left(K_2+\frac{K_1K_2}{\Lambda}\right)\sigma_{2j-1}-\left(\Lambda-\frac{K_1^2}{2\Lambda}\right)\sigma_{2j-1}^2\right\}. \qquad (23)$$

Denoting σ_{2j-1} by σ'_j, expression (19) may be written in the renormalized form

$$Q_N=\int_{-\infty}^{\infty}\!\!\cdots\int\exp\left[\sum_{j=1}^{N'}\left\{K'_0+K'_1\sigma'_j\sigma'_{j+1}+K'_2\sigma'_j-\Lambda'\sigma'^2_j\right\}\right]d\sigma'_1\ldots d\sigma'_{N'}, \qquad (24)$$

where

$$K'_0=\frac{1}{2}\ln\left(\frac{\pi}{\Lambda}\right)+2K_0+\frac{K_2^2}{4\Lambda},\quad K'_1=\frac{K_1^2}{2\Lambda}, \qquad (25\text{a, b})$$

$$K'_2=K_2\left(1+\frac{K_1}{\Lambda}\right),\quad \Lambda'=\Lambda-\frac{K_1^2}{2\Lambda} \qquad (25\text{c, d})$$

and, of course, $N'=\frac{1}{2}N$. It follows that, with $K_0=0$,

$$Q_N(K_1,K_2,\Lambda)=e^{N'K'_0}Q_{N'}(K'_1,K'_2,\Lambda') \qquad (26)$$

and hence for the free energy *per spin* (in units of kT) we have the recurrence relation

$$f(K_1,K_2,\Lambda)=-\tfrac{1}{2}K'_0+\tfrac{1}{2}f(K'_1,K'_2,\Lambda'). \qquad (27)$$

The critical behavior of this system will be studied in Sec. 13.4. Presently we would like to demonstrate how the free energy of the system, over a broad range of the variables K_1 and K_2, can be determined by using the recurrence relation (27) along with the transformation eqns (25).

First of all we identify two *invariants* of the transformation, viz.

$$\Lambda'^2-K'^2_1=\Lambda^2-K_1^2 \qquad (28\text{a})$$

and

$$(\Lambda'-K'_1)/K'_2=(\Lambda-K_1)/K_2. \qquad (28\text{b})$$

It turns out that precisely these combinations appear in the constraint equation of the system as well; see Problem 12.18. It follows that the constraint eqn. (20) is *RG-invariant*, i.e. once it is satisfied in the original system, its counterpart

$$\left\langle\sum_{j=1}^{N'}\sigma'^2_j\right\rangle=N' \qquad (29)$$

is automatically satisfied in the transformed system—without any need to rescale the spin variables.[3] Now, in the absence of the field ($K_2=0$), eqns (25) and (27) give

$$K'_0=\frac{1}{2}\ln\left(\frac{\pi}{\Lambda}\right),\quad K'_1=\frac{K_1^2}{2\Lambda},\quad K'_2=0,\quad \Lambda'=\Lambda-\frac{K_1^2}{2\Lambda} \qquad (30)$$

and hence

$$f(K_1, \Lambda) = -\frac{1}{4}\ln\left(\frac{\pi}{\Lambda}\right) + \frac{1}{2}f\left(\frac{K_1^2}{2\Lambda}, \Lambda - \frac{K_1^2}{2\Lambda}\right). \tag{31}$$

The functional eqn. (31) has the solution

$$f(K_1, \Lambda) = \frac{1}{2}\ln\left[\frac{\Lambda + \sqrt{(\Lambda^2 - K_1^2)}}{2\pi}\right], \tag{32}$$

valid at *all* K_1; the appropriate value of Λ is given by the constraint equation

$$\frac{\partial f}{\partial \Lambda} = \frac{1}{2\sqrt{(\Lambda^2 - K_1^2)}} = 1, \quad \text{that is} \quad \Lambda = \frac{1}{2}\sqrt{(1 + 4K_1^2)}. \tag{33}$$

Note that the invariant (28a) in this case is equal to $\frac{1}{4}$. On the other hand, in the paramagnetic case ($K_1 = 0$), we get

$$K_0' = \frac{1}{2}\ln\left(\frac{\pi}{\Lambda}\right) + \frac{K_2^2}{4\Lambda}, \quad K_1' = 0, \quad K_2' = K_2, \quad \Lambda' = \Lambda \tag{34}$$

and hence

$$f(K_2, \Lambda) = -\frac{1}{4}\ln\left(\frac{\pi}{\Lambda}\right) - \frac{K_2^2}{8\Lambda} + \frac{1}{2}f(K_2, \Lambda), \tag{35}$$

with the solution

$$f(K_2, \Lambda) = -\frac{1}{2}\ln\left(\frac{\pi}{\Lambda}\right) - \frac{K_2^2}{4\Lambda}, \tag{36}$$

valid at *all* K_2; the appropriate value of Λ is now given by

$$\frac{\partial f}{\partial \Lambda} = \frac{1}{2\Lambda} + \frac{K_2^2}{4\Lambda^2} = 1, \quad \text{that is} \quad \Lambda = \frac{\sqrt{(1 + 4K_2^2)} + 1}{4}. \tag{37}$$

The case when both K_1 and K_2 are present is left as an exercise for the reader; see Problem 13.3.

C. The Ising model in two dimensions

As our third example of renormalization, we consider an Ising model on a square lattice in two dimensions. In the absence of the field, the partition function of this system is given by

$$Q_N(T) = \sum_{\{\sigma_i\}} \exp\left\{\sum_{\text{n.n.}} K\sigma_i\sigma_j\right\} \quad (K = \beta J), \tag{38}$$

where the summation in the exponent goes over all nearest-neighbor pairs of spins in the lattice. Writing the summand in (38) as a product of factors pertaining to different pairs of spins, we may highlight those factors that contain a particular spin, say σ_5, and carry out summation over this spin (see Fig. 13.4):

$$\sum_{\sigma_5=\pm 1}\prod_{\text{n.n.}} \cdots e^{K\sigma_2\sigma_5}\cdot e^{K\sigma_4\sigma_5}\cdot e^{K\sigma_5\sigma_6}\cdot e^{K\sigma_5\sigma_8}\cdots \tag{39}$$

$$=\prod_{\text{n.n.}}{}' \cdots [2\cosh K(\sigma_2+\sigma_4+\sigma_6+\sigma_8)]\ldots, \tag{40}$$

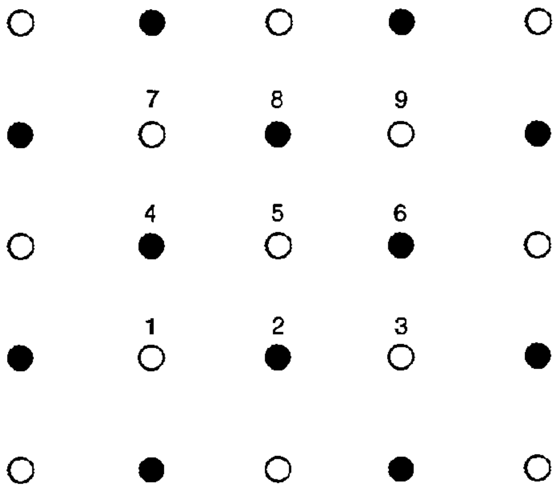


FIG. 13.4. A section of the two-dimensional Ising lattice. The summed-over spins are denoted by open circles, the remaining ones by solid dots. We concentrate on summation over σ_5.

where the primed product goes over the remaining nearest-neighbor pairs in the lattice. This procedure of summation is supposed to be continued till one-half of the spins, shown as open circles in Fig. 13.4, are all summed over. Clearly, this will generate a host of factors of the type shown in expression (40) but the real task now is to express these factors in a form similar, or at least as close as possible, to the factors appearing in the original expression (38); moreover, this mode of expression should be valid for all possible values of the remaining spins, namely $\sigma_2, \sigma_4, \ldots = \pm 1$.

For the factor explicitly displayed in (40), there are 16 possible values for the spins involved, of which only 4 turn out to be non-equivalent; they are

(i) $\sigma_2 = \sigma_4 = \sigma_6 = \sigma_8$, (41a)

(ii) $\sigma_2 = \sigma_4 = \sigma_6 = -\sigma_8$, (41b)

(iii) $\sigma_2 = \sigma_4 = -\sigma_6 = -\sigma_8$, (41c)

(iv) $\sigma_2 = -\sigma_4 = -\sigma_6 = \sigma_8$. (41d)

However, even 4 are too many to accommodate by a factor of the form

$$\exp\{A + B(\sigma_2\sigma_4 + \sigma_2\sigma_6 + \sigma_4\sigma_8 + \sigma_6\sigma_8)\},$$

which contains only nearest-neighbor interactions in the transformed lattice and hence only two parameters to choose. Clearly, we need two more degrees of freedom, and it turns out that these are provided by the next-nearest-neighbor interactions and by interactions among a quartet of spins sitting on the corners of an elementary square in the new lattice. Thus, we are obliged to set

$$2\cosh K(\sigma_2+\sigma_4+\sigma_6+\sigma_8)=\exp\left\{K_0'+\tfrac{1}{2}K'(\sigma_2\sigma_4+\sigma_2\sigma_6+\sigma_4\sigma_8+\sigma_6\sigma_8)\right.$$
$$\left.+L'(\sigma_2\sigma_8+\sigma_4\sigma_6)+M'\sigma_2\sigma_4\sigma_6\sigma_8\right\}; \qquad (42)$$

the reason why we have written $\frac{1}{2}K'$, rather than K', will become clear shortly. Now, the four possibilities listed above require that

$$2\cosh 4K=\exp(K_0'+2K'+2L'+M'), \qquad (43a)$$

$$2\cosh 2K=\exp(K_0'-M'), \qquad (43b)$$

$$2=\exp(K_0'-2L'+M'), \qquad (43c)$$

$$2=\exp(K_0'-2K'+2L'+M'), \qquad (43d)$$

with the result that

$$K_0'=\ln 2+\tfrac{1}{2}\ln\cosh 2K+\tfrac{1}{8}\ln\cosh 4K. \qquad (44)$$

$$K'=\tfrac{1}{4}\ln\cosh 4K, \qquad (45)$$

$$L'=\tfrac{1}{8}\ln\cosh 4K, \qquad (46)$$

$$M'=\tfrac{1}{8}\ln\cosh 4K-\tfrac{1}{2}\ln\cosh 2K. \qquad (47)$$

Continuing the process, we find that the factor $\exp\left(\frac{1}{2}K'\sigma_2\sigma_4\right)$ appears once again when summation over σ_1 is carried out, the factor $\exp\left(\frac{1}{2}K'\sigma_2\sigma_6\right)$ appears once again when summation over σ_3 is carried out, and so on; no further factors involving the products $\sigma_2\sigma_8$, $\sigma_4\sigma_6$ and $\sigma_2\sigma_4\sigma_6\sigma_8$ appear. The net result is that the partition function (38) assumes the form

$$Q_N=e^{N'K_0'}\sum_{\{\sigma_j'\}}\exp\left\{K'\sum_{\text{n.n.}}\sigma_j'\sigma_k'+L'\sum_{\text{n.n.n.}}\sigma_j'\sigma_k'+M'\sum_{\text{sq.}}\sigma_j'\sigma_k'\sigma_l'\sigma_m'\right\}, \qquad (48)$$

where $N'=\frac{1}{2}N$.

Clearly, we have not been able to establish an *exact* correspondence between the transformed system and the original one (in which no interactions other than the nearest-neighbor ones were present). It seems more reasonable now that we redefine the original system as one having all the interactions appearing in expression (48), but with $L=M=0$. We may then write

$$Q_N(K,0,0)=e^{N'K_0'}Q_{N'}(K',L',M') \qquad (49)$$

and hence for the free energy *per spin* (in units of kT)

$$f(K,0,0)=-\tfrac{1}{2}K_0'+\tfrac{1}{2}f(K',L',M'), \qquad (50)$$

where K_0', K', L' and M' are given by eqns (44)–(47). In view of the appearance of new parameters, L' and M', in the recurrence relation (50), no further progress can be made without introducing some sort of an approximation, for which see Sec. 13.4. But one thing is clear: when renormalization is carried out in dimensions two or more, the connectivity of the lattice requires that the Hamiltonian of the decimated system contain some higher-order interactions not present in the original system, in order that the latter be represented correctly on transformation. It is obvious that further renormalizations would require more and more such interactions, and hence the need for more and more parameters would grow without limit. It may then be advisable that the Hamiltonian of the given system be regarded as a function of an "infinitely large number of parameters" (all but a few of which are zero to begin with), such that the number of parameters with a nonzero value grows indefinitely as renormalization transformations are carried out in succession and the number of degrees of freedom of the system steadily reduced.

We now present a formulation of the renormalization group approach to the study of critical phenomena.

13.3. The renormalization group: general formulation

We start with a system whose Hamiltonian depends on a large number of parameters $K_1, K_2, \ldots$ (all but a few of which are zero to begin with) and on the spin configuration $\{\sigma_i\}$ of the lattice. The free energy of the system is then given by

$$\exp(-\beta A) = \sum_{\{\sigma_i\}} \exp[-\beta H_{\{\sigma_i\}}(\{K_\alpha\})] \quad \alpha = 1, 2, \ldots. \tag{1}$$

We now effect a "decimation" of the system which reduces the number of degrees of freedom from N to N' and the correlation length from ξ to ξ', such that

$$N' = l^{-d}N, \quad \xi' = l^{-1}\xi \quad (l > 1). \tag{2a, b}$$

Expressing the Hamiltonian of the transformed system in a form similar to the one for the original system, except that we now have new parameters K'_α, along with the additional K'_0, and new spins σ'_j, eqn. (1) takes the form

$$\exp(-\beta A) = \exp(N'K'_0) \sum_{\{\sigma'_j\}} \exp\left[-\beta H_{\{\sigma'_j\}}(\{K'_\alpha\})\right], \tag{3}$$

so that the free energy *per spin* (in units of β^{-1}) is given by

$$f(\{K_\alpha\}) = l^{-d}\left[-K'_0 + f(\{K'_\alpha\})\right]. \tag{4}$$

We now look closely at the transformation $\{K_\alpha\} \to \{K'_\alpha\}$ by introducing a *vector space* $\mathcal{K}$ in which the set of parameters K_α is represented by the tip of a position vector K; on transformation, K changes to K', which may be looked upon as a kind of "flow from one position in the vector space to another". This flow may be represented by the transformation equation

$$K' = \mathcal{R}_l(K), \tag{5}$$

where $\mathcal{R}_l$ is the renormalization-group operator appropriate to the case under consideration. A repeated application of this process leads to a sequence of vectors

$K', K'', \ldots$, such that

$$K^{(n)} = \mathscr{R}_l(K^{(n-1)}) = \ldots = \mathscr{R}_l^n(K^{(0)}) \quad n = 0, 1, 2, \ldots, \tag{6}$$

where $K^{(0)}$ denotes the original K. At the end of this sequence, the correlation length ξ and the singular part of the free energy f_s are given by

$$\xi^{(n)} = l^{-n}\xi^{(0)}, \quad f_s^{(n)} = l^{nd} f_s^{(0)}; \tag{7a, b}$$

see eqns (2b) and (4).

Now, the transformation (5) may have a *fixed point*, K^*, so that

$$\mathscr{R}_l(K^*) = K^*. \tag{8}$$

Equation (2b) then tells us that $\xi(K^*) = l^{-1}\xi(K^*)$, which means that $\xi(K^*)$ is either zero or infinite! The former possibility is of little interest to us, so let us dwell only on the latter (which makes the system with parameters $K = K^*$ critical); in simple situations, the fixed point, K^*, will correspond to the critical point, K_c, of the given system. Conceivably, an arbitrary point K, on successive transformations such as (6), may end up at the fixed point K^*. Since the correlation length ξ can only decrease on transformation, see eqn. (7a), and is infinite at the end of this sequence of transformations, it must be infinite at K as well (the same being true for *all* points intermediate between K and K^*). The collection of all those points which, on successive transformations, flow into the fixed point, constitutes a surface all over which ξ is infinite; this surface is generally referred to as the *critical surface*. All flow lines in this surface are directed towards, and terminate at, the fixed point, while points off this surface may initially move towards the fixed point but eventually their flow lines veer away from it; see Fig. 13.5. Reasons behind this pattern of flow will become clear soon.

For the analysis of the critical behavior we examine the pattern of flow in the neighborhood of the fixed point K^*.[4] Setting

$$K = K^* + k, \tag{9}$$

we have by eqns (5) and (8)

$$K' = K^* + k' = \mathscr{R}_l(K^* + k) = K^* + \mathscr{R}_l(k), \tag{10}$$

so that

$$k' = \mathscr{R}_l(k). \tag{11}$$

Assuming $\{k_\alpha\}$, and hence $\{k'_\alpha\}$, to be small, we may linearize eqn. (11) to write

$$k' = \mathscr{A}_l^* k, \tag{12}$$

where $\mathscr{A}_l^*$ is a matrix arising from the linearization of the operator $\mathscr{R}_l$ around the fixed point K^*. The eigenvalues λ_i and the eigenvectors ϕ_i of the matrix $\mathscr{A}_l^*$ play a vital role in determining the critical behavior of the system.

If the vectors ϕ_i form a complete set, we may expand k and k' in terms of ϕ_i,

$$k = \sum_i u_i \phi_i, \quad k' = \sum_i u'_i \phi_i, \tag{13a, b}$$

with the result that

$$u'_i = \lambda_i u_i \quad i = 1, 2, \ldots; \tag{14}$$

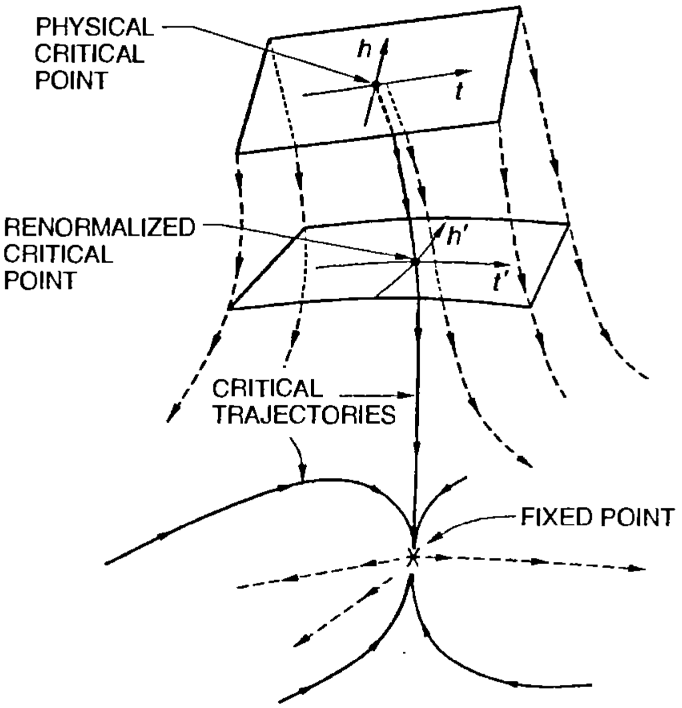


FIG. 13.5. The parameter space of a physical system, showing critical trajectories (solid lines) flowing into the fixed point. The sub-space of the relevant variables, t and h, is everywhere "orthogonal" to these trajectories (on all of which $t = h = 0$); the critical trajectories differ from one another only in respect of irrelevant variables which vanish as the fixed point is approached. The dashed lines depict that part of the flow in which the relevant variables play a decisive role.

the coefficients u_i appearing here are generally referred to as the *scaling fields* of the problem. Under successive transformations (all in the neighborhood of the fixed point), these fields are given by

$$u_i^{(n)} = \lambda_i^n u_i^{(0)}. \tag{15}$$

It is obvious that the fields u_i are certain linear combinations of the original parameters k_α; they may, therefore, be looked upon as the "generalized coordinates" of the problem. The relative influence of these coordinates in determining the critical behavior of the system depends crucially on the respective eigenvalues λ_i. With $u_i^{(n)}$ given by (15), we have three possible courses for a given coordinate u_i.

(a) If $\lambda_i > 1$, the coordinate u_i grows with n and, with successive transformations, becomes more and more significant. Clearly, u_i in this case is a *relevant variable* which, by itself, drives the system away from the fixed point–thus making the fixed point unstable. By experience, we know that the temperature parameter $t[= (T - T_c)/T_c]$ and the magnetic field parameter $h[= \mu B/kT_c]$ are two basic quantities that vanish at the critical point and are clearly relevant to the problem of phase transitions. We therefore expect that our analysis will produce at least two relevant variables, u_1 and u_2 say, which could be identified with t and h, respectively, so that

$$u_1 = at + O(t^2), \quad u_2 = bh + O(h^2), \tag{16, 17}$$

with both λ_1 and λ_2 greater than unity.

(b) If $\lambda_i < 1$, the coordinate u_i decays with n and, with successive transformations, becomes less and less significant. Clearly, u_i in this case is an *irrelevant variable* which, by itself, drives the system towards the fixed point. Now, if all the relevant variables are set at zero, then successive transformations (by virtue of the irrelevant variables) *will* drive the system to the fixed point. We must then be cruising on the critical surface itself (where all trajectories are known to flow into the fixed point). It follows that on the critical surface all relevant variables of the problem are zero; furthermore, the divergence of the correlation length is also tied to the same fact.

(c) If $\lambda_i = 1$, the coordinate u_i, in the linear approximation, stays constant; it neither grows nor decays very rapidly unless one enters the nonlinear, beyond-scaling, regime of the variable u_i. Quite appropriately, u_i in this case is termed a *marginal variable*; it does not affect the critical behavior of the system as significantly as a relevant variable does, but it may throw in logarithmic factors along with the conventional power laws. The ability to identify such variables and to predict the consequent departures from simple power-law scaling is one of the added virtues of the RG approach.

The above considerations enable us to understand the pattern of flow shown in Fig. 13.5. While the points *on* the critical surface flow into the fixed point, those *off* this surface flow towards the fixed point by virtue of the irrelevant variables and, at the same time. away from it by virtue of the relevant variables; the net result is an initial approach towards but a final recession away from the fixed point. Points close to the fixed point and in directions "orthogonal" to the critical surface have only relevant variables to contend with, so right away they move away from the fixed point. It is this part of the flow that determines the critical behavior of the given system.

Disregarding the irrelevant variables, we now examine the manner in which the correlation length. ξ, and the (singular part of the) free energy, f_s, of the system are affected by the transformation (15). In view of eqns (7), we have

$$\xi(u_1, u_2, \ldots) = l^n \xi(\lambda_1^n u_1, \lambda_2^n u_2, \ldots) \tag{18}$$

and

$$f_s(u_1, u_2, \ldots) = l^{-nd} f_s(\lambda_1^n u_1, \lambda_2^n u_2, \ldots). \tag{19}$$

Identifying u_1 with t, as in (16), and remembering the definition of the critical exponent ν, we obtain from (18)

$$u_1^{-\nu} = l^n (\lambda_1^n u_1)^{-\nu}, \tag{20}$$

with the result that

$$\nu = \ln l / \ln \lambda_1. \tag{21}$$

At first sight one might wonder why ν should depend on l. In fact, it doesn't because of the following argument. On physical grounds we expect that two successive transformations with scale factors l_1 and l_2 would be equivalent to a single transformation with scale factor $l_1 l_2$, i.e.[5]

$$\mathcal{R}_{l_1}^* \mathcal{R}_{l_2}^* = \mathcal{R}_{l_1 l_2}^*. \tag{22}$$

This forces the eigenvalues λ_i to be of the form l^{y_i}, for

$$l_1^{y_i} l_2^{y_i} = (l_1 l_2)^{y_i}. \tag{23}$$

Relation (21) then becomes

$$\nu = 1/y_1, \tag{24}$$

manifestly independent of l.

Equation (19) may now be written as

$$f_s(t, h, \ldots) = l^{-nd} f_s(l^{ny_1} t, l^{ny_2} h, \ldots). \tag{25}$$

To ensure that the above relationship is independent of the choice of l, we follow the same line of argument as in Sec. 13.1, after eqn. (13.1.9), with the result

$$f_s(t, h, \ldots) = |t|^{d\nu} \tilde{f}_s(h/|t|^{\Delta}, \ldots), \tag{26}$$

where

$$\Delta = y_2/y_1. \tag{27}$$

Equation (26) is formally the same as the scaling relationship postulated in Sec. 11.10 and, one might say, argued out in Sec. 13.1. The big difference here is that not only has this relationship been derived on a firmer basis but now we also have a recipe for evaluating the critical exponents ν, Δ, etc. from first principles. What one has to do in this approach is to determine the RG operator $\mathscr{R}_l$ for the given problem, linearize it around the appropriate fixed point $\boldsymbol{K}^*$, determine the eigenvalues $\lambda_i (= l^{y_i})$ and use eqns (24) and (27) to evaluate ν and Δ. At the same time, recalling the definition of the critical exponent α, we infer from (26) that

$$2 - \alpha = d\nu; \tag{28}$$

the remaining exponents follow with the help of the scaling relations

$$\beta = (2-\alpha) - \Delta, \quad \gamma = 2\Delta - (2-\alpha), \quad \delta = \Delta/\beta, \quad \eta = 2 - (\gamma/\nu). \tag{29}$$

We find that the hyperscaling relation (28) is an integral part of the RG formulation; it comes out naturally–with no external imposition whatsoever. It is, however, disconcerting that, according to the above argument, this relation should hold for *all* d–notwithstanding the fact that, for $d > 4$, all critical exponents are "stuck" at the mean field values and the relation (28) gets replaced by

$$2 - \alpha = 4\nu \quad (d > 4), \tag{30}$$

with $\alpha = 0$ and $\nu = \frac{1}{2}$. The reason for this peculiar behavior is somewhat subtle; it arises from the fact that in certain situations an "irrelevant variable" may well raise its "dangerous" head and affect the outcome of a calculation in a rather significant manner. To see how this happens, we may consider a *continuous spin model*, very much like the one considered in Sec. 12.4, with the probability distribution law

$$p(\sigma_i)\, d\sigma_i = \text{const.}\, e^{-\frac{1}{2}\sigma_i^2 - \tilde{u}\sigma_i^4} \quad (\tilde{u} > 0); \tag{31}$$

cf. eqn. (12.4.1). The free energy of the system then depends on the parameter $\tilde{u}$ as well as on t and h, and we obtain instead of (25)

$$f_s(t, h, \tilde{u}) = l^{-nd} f_s(l^{ny_1} t, l^{ny_2} h, l^{ny_3} \tilde{u}, \ldots). \tag{32}$$

Now, using the RG approach, one finds (see Fisher, 1983; Appendix D) that, for $d > 4$,

$$y_1 = 2, \quad y_2 = \tfrac{1}{2}d + 1, \quad y_3 = 4 - d, \tag{33}$$

showing very clearly that, for $d > 4$, the parameter $\tilde{u}$ is an irrelevant variable. One is, therefore, tempted to ignore $\tilde{u}$ and arrive at eqn. (26), with

$$\nu = \tfrac{1}{2} \quad \text{and} \quad \Delta = (d+2)/4. \tag{34}$$

The very fact that Δ turns out to be d-dependent shows that there is something wrong here. It turns out that though, on successive transformations, the variable $\tilde{u}$ does tend to zero, its influence on the function f_s does not. It may, therefore, be prudent to write

$$f_s(t, h, \tilde{u}) = |t|^{d\nu} \tilde{f}_s(h/|t|^{\Delta}, \tilde{u}/|t|^{\phi}), \tag{35}$$

where

$$\phi = \frac{y_3}{y_1} = \frac{4-d}{2}. \tag{36}$$

Now, by analysis, one finds that

$$\operatorname*{Lim}_{w \to 0} \tilde{f}_s(v, w) \approx \frac{1}{w} F(vw^{1/2}), \tag{37}$$

where $F(vw^{1/2})$ is a perfectly well-behaved function. The singularity of $\tilde{f}_s$ in w changes the picture altogether, and we get in the desired limit

$$f_s(t, h, \tilde{u}) \approx |t|^{d\nu + \phi} F\left(h/|t|^{\Delta + \frac{1}{2}\phi}\right). \tag{38}$$

The "revised" value of Δ now is

$$\Delta_{\text{rev}} = \frac{d+2}{4} + \frac{4-d}{4} = \frac{3}{2},$$

which is indeed independent of d and agrees with the corresponding mean field value. At the same time, the "revised" form of the hyperscaling relation now is

$$2 - \alpha = \frac{d}{2} + \frac{4-d}{2} = 2. \tag{39}$$

as stated in (30).

The lesson to be learnt here is that the standard derivations of the scaling relations rest on assumptions, often left unstated, about the nonsingular or nonvanishing behavior of various scaling functions and their arguments. In many cases these assumptions are valid and may even be confirmed by explicit calculations or otherwise; in certain circumstances, however, they fail–in which case a scaling relation may change its form. Fortunately, such circumstances are not too common.

13.4. Applications of the renormalization group

We start ou onsiderations with the models examined in Sec. 13.2.

A. The Ising model in one dimension

The renormalization group transformation in this case is given by eqns (13.2.8b and c), viz.

$$K_1' = \tfrac{1}{4}\ln[\cosh(2K_1 + K_2)\cosh(2K_1 - K_2)] - \tfrac{1}{2}\ln\cosh K_2 \tag{1a}$$

and

$$K_2' = K_2 + \tfrac{1}{2}\ln\left[\cosh(2K_1 + K_2)/\cosh(2K_1 - K_2)\right], \tag{1b}$$

where $K_1 = J/kT$ and $K_2 = \mu B/kT$. It is straightforward to see that this transformation has a "line of *trivial* fixed points", with $K_1 = 0$ and K_2 arbitrary. These fixed points pertain to either $J = 0$ or $T = \infty$; in either case, one has a correlation length that vanishes. There is also a *nontrivial* fixed point at $K_1 = \infty$ and $K_2 = 0$, which may be realized by first setting $B = 0$ and then letting $T \to 0$; the correlation length at this fixed point will be infinite. In the vicinity of this point, we have

$$K_1' \simeq K_1 - \tfrac{1}{2}\ln 2, \quad K_2' \simeq 2K_2. \tag{2ab}$$

Now, since $K_1^* = \infty$, K_1 is not an appropriate variable to carry out an expansion around the fixed point. We may instead adopt a new variable, see eqn. (12.1.17), namely

$$t = \exp(-pK_1) \quad (p > 0), \tag{3}$$

so that $t^* = 0$; now, in the vicinity of the fixed point, we have

$$t' \simeq 2^{p/2}t. \tag{4}$$

Identifying K_2 as the variable h, and remembering that the scale factor l here is 2, we readily obtain from eqns (2b) and (4)

$$y_1 = p/2, \quad y_2 = 1. \tag{5}$$

The critical exponents of the model follow straightforwardly from (5); we get

$$\nu = 2/p, \quad \Delta = 2/p, \tag{6}$$

whence, see eqns (13.3.28 and 29),

$$\alpha = 2 - 2/p, \quad \beta = 0, \quad \gamma = 2/p, \quad \delta = \infty, \quad \eta = 1, \tag{7}$$

in complete agreement with the results found in Sec. 12.1. As for the choice of p, see remarks following eqn. (12.2.21).

B. The spherical model in one dimension

The RG transformation in this case is given by eqns (13.2.25b, c and d), viz.

$$K_1' = \frac{K_1^2}{2\Lambda}, \quad K_2' = K_2\left(1 + \frac{K_1}{\Lambda}\right), \quad \Lambda' = \Lambda - \frac{K_1^2}{2\Lambda}, \tag{8a, b, c}$$

where $K_1 = J/kT$, $K_2 = \mu B/kT$ and $\Lambda = \lambda/kT$, λ being the "spherical field" that enabled us to take care of the constraint on the model. The nontrivial fixed point is again at $T = 0$ where $\lambda \simeq J$ [see eqn. (12.4.24), with $d = 1$] and hence $\Lambda = K_1$. Equations (8) [illegible] reduce to the linearized form (valid for small T)

$$K_1' \simeq \tfrac{1}{2}K_1, \quad K_2' \simeq 2K_2, \quad \Lambda' \simeq \tfrac{1}{2}\Lambda. \tag{9a, b, c}$$

Equations (9a) and (9c) contain essentially the same information, viz. $T' \simeq 2T$. Clearly, T itself is a good variable for expansion in this case–giving $y_1 = 1$. Equation (9b), just like (2b), gives $y_2 = 1$, and we obtain

$$\nu = 1, \quad \Delta = 1, \tag{10}$$

whence

$$\alpha = 1, \quad \beta = 0, \quad \gamma = 1, \quad \delta = \infty, \quad \eta = 1, \tag{11}$$

in complete agreement with the results for one-dimensional models with $n \geq 2$, as quoted in eqn. (12.2.20).

C. *The Ising model in two dimensions*

The RG transformation in this case is given by eqns (13.2.45–47), viz.

$$K' = \tfrac{1}{4}\ln\cosh 4K, \tag{12}$$

$$L' = \tfrac{1}{8}\ln\cosh 4K \tag{13}$$

and

$$M' = \tfrac{1}{8}\ln\cosh 4K - \tfrac{1}{2}\ln\cosh 2K. \tag{14}$$

It will be recalled that, in effecting this transformation, we started only with nearest-neighbor interactions (characterized by the parameter $K = \beta J$) but, due to the connectivity of the lattice, ended up with more, viz. the next-nearest-neighbor interactions (characterized by L') and the four-spin interactions (characterized by M'), in addition to the nearest-neighbor interactions (characterized by K'). On subsequent transformations, still higher-order interactions come into the picture and the problem becomes formidable unless some approximations are introduced. In one such approximation, due originally to Wilson (1975), we discard all interactions other than the ones represented by the parameters K and L, and at the same time assume K and L to be small enough so that eqns (12) and (13) reduce to

$$K' = 2K^2, \quad L' = K^2. \tag{15a, b}$$

Now, if the parameter L had been introduced right in the beginning, the transformation equations in the same approximation would have been

$$K' = 2K^2 + L, \quad L' = K^2. \tag{16a, b}$$

We shall treat eqns (16) as if they were the exact transformation equations of the problem and see what they lead to.

It is straightforward to see that the transformation (16) has a nontrivial fixed point at

$$K^* = \tfrac{1}{3}, \quad L^* = \tfrac{1}{9}. \tag{17}$$

Linearizing around the fixed point, we get

$$k_1' = \tfrac{4}{3}k_1 + k_2, \quad k_2' = \tfrac{2}{3}k_1, \tag{18}$$

where k_1 and k_2 represent deviations of the parameters K and L from the fixed-point values K^* and L^*, respectively. The transformation matrix $\mathcal{T}_l^*$ of eqn. (13.3.12) is thus given by

$$\mathcal{T}_{\surd 2}^* = \begin{pmatrix} \frac{4}{3} & 1 \\ \frac{2}{3} & 0 \end{pmatrix} \tag{19}$$

whose eigenvalues are

$$\lambda_1 = \tfrac{1}{3}(2 + \surd 10), \quad \lambda_2 = \tfrac{1}{3}(2 - \surd 10) \tag{20a, b}$$

and whose eigenvectors are

$$\phi_1 \sim \begin{pmatrix} 2 + \surd 10 \\ 2 \end{pmatrix}, \quad \phi_2 \sim \begin{pmatrix} 2 - \surd 10 \\ 2 \end{pmatrix}. \tag{21a, b}$$

The scaling fields u_i are then determined by eqn. (13.3.13a) which, on inversion, gives

$$u_1 \sim \{2k_1 + (\surd 10 - 2)k_2\}, \quad u_2 \sim \{2k_1 - (\surd 10 + 2)k_2\}. \tag{22a, b}$$

Clearly, the field u_1, with $\lambda_1 > 1$, is the relevant variable of the problem, while the field u_2 is irrelevant. The "critical curve" in the $(K,\ L)$-plane is determined by the condition $u_1 = 0$, while the linear part of this curve, in the vicinity of the fixed point ($\boldsymbol{u} = 0$), is mapped by the relation (22a). In terms of the variables k_1 and k_2, this segment of the critical curve is given by the equation

$$k_2 \approx -\frac{\surd 10 + 2}{3} k_1, \tag{23}$$

which represents a straight line of slope -1.7208; see Fig. 13.6.

To determine the *physical* critical point of this model we have to locate on the critical curve a point with $L = 0$, for the original problem had no interactions other than the one represented by the parameter K; the corresponding value of K would be our K_c.[6] This requires a mapping of the critical curve right up to the K-axis. While this has been done numerically (Wilson, 1975), a crude estimate of K_c can be made by simply extending the straight-line segment (23) down to the desired limit. One thus obtains[7]

$$K_c = \frac{1}{3} + \frac{1}{9}\frac{3}{\surd 10 + 2} = \frac{4 + \surd 10}{18} = 0.3979, \tag{24}$$

which may be compared with the exact result found in Sec. 12.3, viz. 0.4407. Another quantity that can be estimated here is the critical exponent ν. From eqns (13.3.21) and (20a), one obtains

$$\nu = \frac{\ln \surd 2}{\ln[(2 + \surd 10)/3]} = 0.6385, \tag{25}$$

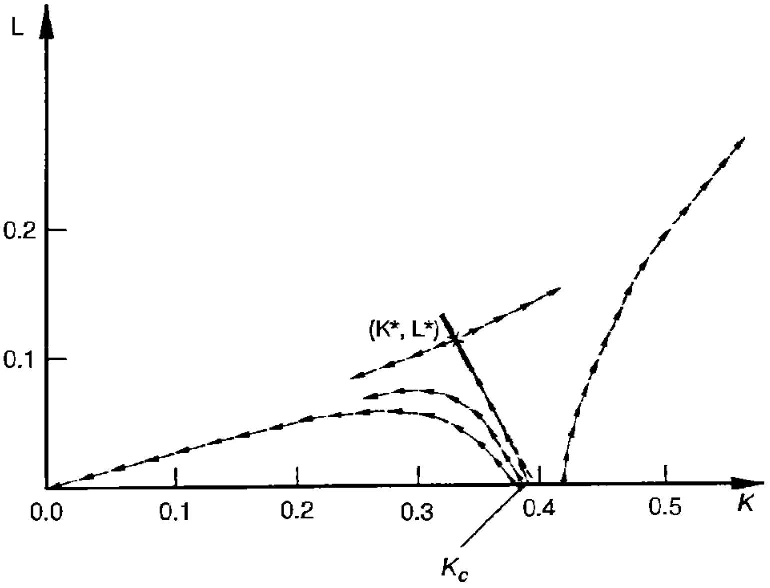


FIG. 13.6. A section of the critical curve for the two-dimensional Ising model near the nontrivial fixed point ($K^* = \frac{1}{3}, L^* = \frac{1}{9}$). Points on the critical curve flow into the fixed point, while those off it flow away towards the trivial fixed point ($K^* = L^* = 0$) or ($K^* = L^* = \infty$).

which may be compared with the exact value 1. Although a comparison of the results obtained here with the ones following from exact analysis is not very flattering, the basic merits of the RG approach are quite obvious.

One important aspect of critical phenomena, viz. their *universality* over a large class of systems, is manifest even in this simple example. Imagine, for instance, that in the case of the given system a next-nearest-neighbor interaction L_0 were indeed present. Our approximate treatment would then lead to the same fixed point and the same critical curve as above, but our *physical* critical point would now be given by that "point on the critical curve whose L-value is L_0"; we may denote this critical point by $K_c(L_0)$. As for the critical behavior, it will still be determined by an expansion around the fixed point, for that is where the "relevant part" of the flow is; see again Fig. 13.6. Clearly, the critical behavior of the given system, insofar as exponents are concerned, will be the same, regardless of the actual value of L_0. And, by extension, the same will be true of any two systems which have the same basic topology but differ only in the details of the spin–spin interactions.

As for the accuracy of the results obtained, improvements are needed in several important respects. First of all, the exclusion of all interaction parameters other than K and L constitutes a rather inadequate approximation; one should at least include the four-spin interaction, represented by the parameter M, and may possibly ignore the ones that appear on successive transformations. Next, the assumption that the parameters K and L are small is also unjustified–especially for K; this makes a numerical approach to the problem rather essential. Thirdly, we disregarded the renormalization of the spins, from the original $\sigma(\boldsymbol{r})$ to $\sigma'(\boldsymbol{r}')$, as required by eqn. (13.1.23); in the present problem, this would amount to introducing a factor of $(\sqrt{2})^{\eta/2}$, i.e. $2^{1/16}$, for η here is $\frac{1}{4}$. In the actual treatment, one may have to introduce an unknown parameter, ρ, and determine its "true" value by analysis (see Wilson, 1975). Highly sophisticated procedures have been developed over the years to accommodate (or circumvent) these problems, leading to very accurate–in fact, almost exact–results for the model under considerations. For details, see the review article by Niemeijer and van Leeuwen (1976), where references to other pertinent literature on the subject can also be found.[8]

D. The ε-expansion

Application of the RG approach to systems in higher dimensions, viz. with $d > 2$, presents serious mathematical difficulties. One is then forced to resort to approximation procedures such as the ε-expansion, first introduced by Wilson (1972); see also Wilson and Fisher (1972). This procedure was inspired by the observation that the field-theoretic calculations of the RG formulation become especially simple as the upper critical dimension, $d = 4$, is approached; it therefore seemed desirable to introduce a parameter ε $(= 4 - d)$ and carry out expansions of the various quantities of interest around $\varepsilon = 0$. The model adopted for these calculations was the same as the one referred to in Sec. 13.3, namely the *continuous, n-vector spin model*, with the probability distribution given by eqn. (13.3.31).[9] An important advantage of using continuous spins $\sigma(\boldsymbol{r})\{= \sigma^{(\mu)}(\boldsymbol{r}),\ \mu = 1, \ldots, n\}$, with $-\infty < \sigma^{(\mu)} < \infty$, is that one can introduce Fourier transforms $\sigma(\boldsymbol{q})$ and make use of the "momentum shell integration" technique of Wilson (1971). The parameters of interest now are (see Fisher, 1983)

$$r = \frac{T - T_0}{T_0 R_0^2} = \frac{t_0}{R_0^2}, \qquad u = \tilde{u}\frac{T^2 a^d}{T_0^2 R_0^4} \tag{26a, b}$$

and, of course, the magnetic field parameter h; here, T_0 denotes the mean-field critical temperature qJ/k (q being the coordination number), R_0 is a measure of the range of interactions, a the lattice constant whereas $\tilde{u}$ is the real-space parameter appearing in eqn. (13.3.31). The transformation equations, with a scale factor l, turn out to be

$$r' = l^2 r + 4(l^2 - 1)c(n + 2)u - l^2 \ln l (n + 2)(2\pi^2)^{-1} ru, \tag{27a}$$

$$u' = (1 + \varepsilon \ln l)u - (n + 8) \ln l (2\pi^2)^{-1} u^2, \tag{27b}$$

and

$$h' = l^3(1 - \tfrac{1}{2}\varepsilon \ln l)h, \tag{27c}$$

correct to the orders displayed; the parameter c in eqn. (27a) is related to a cut-off in the momentum space which, in turn, is a reflection of the underlying lattice structure.

Transformation (27) has two fixed points–the so-called *Gaussian* fixed point, with

$$r^* = u^* = h^* = 0, \tag{28}$$

and a *non-Gaussian* fixed point, with

$$r^* = -\frac{8\pi^2 c(n + 2)}{(n + 8)}\varepsilon, \qquad u^* = \frac{2\pi^2}{(n + 8)}\varepsilon, \qquad h^* = 0. \tag{29}$$

We now examine the two distinct situations.

(i) $d \gtrsim 4$, so that ε is a small negative number: One readily sees from eqn. (27b) that the parameter u in this case decreases on transformation, so on successive transformations it will tend to zero. Clearly, only the Gaussian fixed point is the one appropriate to this case. Linearizing around this point, we obtain for the transformation matrix $\mathcal{A}_l^*$

$$\mathcal{A}_l^* = \begin{pmatrix} l^2 & 4(l^2-1)c(n+2) \\ 0 & 1+\varepsilon \ln l \end{pmatrix}, \tag{30}$$

with eigenvalues

$$\lambda_r = l^2, \quad \lambda_u = (1+\varepsilon \ln l) < 1 \tag{31}$$

and, of course,

$$\lambda_h = l^3(1 - \tfrac{1}{2}\varepsilon \ln l). \tag{32}$$

It follows that

$$y_1 = 2, \quad y_2 \approx 3 - \tfrac{1}{2}\varepsilon, \quad y_3 \approx \varepsilon, \tag{33}$$

as in eqn. (13.3.33). Note that the parameter u in this case is an irrelevant variable but, as discussed at the end of Sec. 13.3, it is a *dangerously irrelevant variable* which does eventually affect the results of the calculation in hand.

(ii) $d \lesssim 4$, so that ε is a small positive number: The parameter u now behaves very differently. If it is already zero, it stays so; otherwise, it moves away from that value, carrying the system to some other fixed point–possibly the non-Gaussian one, with coordinates given in (29). The resulting pattern of flow in the $(r,\ u)$-plane is shown in Fig. 13.7; clearly, the Gaussian fixed point is no longer appropriate and the problem now revolves around the non-Gaussian fixed point

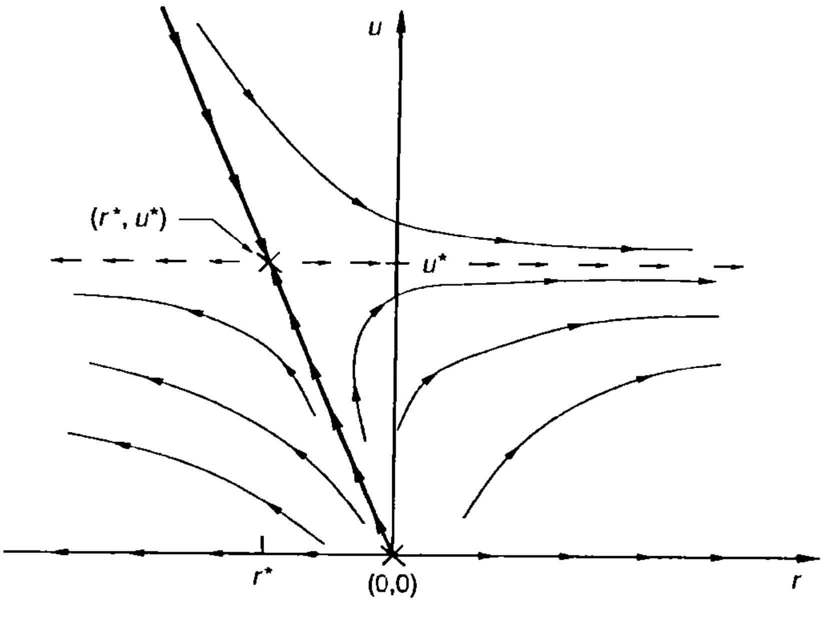


FIG. 13.7. A section of the critical curve and a sketch of the RG flows in the (r, u)-plane for $0 < \varepsilon \ll 1$. Note that the critical curve is straight only to order ε.

instead. Linearizing around the latter, we obtain

$$\mathscr{A}_l^* = \begin{pmatrix} l^2\left\{1 - \dfrac{n+2}{n+8}\varepsilon \ln l\right\} & 4(l^2-1)c(n+2) \\ 0 & 1 - \varepsilon \ln l \end{pmatrix}, \tag{34}$$

with eigenvalues

$$l^2\left\{1 - \frac{n+2}{n+8}\varepsilon \ln l\right\} \quad \text{and} \quad (1 - \varepsilon \ln l) < 1. \tag{35}$$

We note that of the "generalized coordinates" u_1 and u_2, which are certain linear combinations of the parameters $\Delta r(= r - r^*)$ and $\Delta u(= u - u^*)$, only u_1 is a relevant variable of the problem.[10] Identifying u_1 with the temperature parameter t, we obtain

$$y_t \approx 2 - \frac{n+2}{n+8}\varepsilon. \tag{36}$$

Combining this with expression (33) for y_h, viz.

$$y_h \approx 3 - \tfrac{1}{2}\varepsilon, \tag{37}$$

we obtain, see eqns (13.3.24, 27, 28 and 29),

$$\nu \approx \tfrac{1}{2} + \frac{n+2}{4(n+8)}\varepsilon, \quad \Delta \approx \tfrac{3}{2} + \frac{n-1}{2(n+8)}\varepsilon, \tag{38}$$

whence

$$\alpha \approx \frac{4-n}{2(n+8)}\varepsilon, \quad \beta \approx \frac{1}{2} - \frac{3}{2(n+8)}\varepsilon, \quad \gamma \approx 1 + \frac{n+2}{2(n+8)}\varepsilon, \tag{39}$$

$$\delta \approx 3 + \varepsilon, \quad \eta \approx 0, \tag{40}$$

correct to the first power in ε.

For obvious reasons the value of ε of greatest interest to us is $\varepsilon = 1$, for which the above results are totally inadequate; they do show the correct trends, though. For better numerical accuracy it is essential to extend these calculations to higher orders in ε. Considerable progress has been made in this direction, so that we now have expressions available that include terms up to ε^3 and, in some cases, even ε^4; for details, see Wallace (1976). One wonders if that degree of extension would be enough for obtaining reliable results for ε as large as 1. The answer is yes, and we will illustrate it with an example. For the spherical model we know exact values of the various critical exponents which may, for the purpose of illustration, be expressed as power series in ε. Thus, for instance,

$$\gamma = \frac{2}{d-2} = (1 - \frac{1}{2}\varepsilon)^{-1} = 1 + \frac{1}{2}\varepsilon + \frac{1}{4}\varepsilon^2 + \frac{1}{8}\varepsilon^3 + \frac{1}{16}\varepsilon^4 + \cdots \tag{41}$$

Since the radius of convergence of this series is 2, the value 1 of ε is not really as large as it seems. In fact, the terms displayed in (41) already give $\gamma = 1.9375$, as opposed to the correct value 2. The situation is clearly encouraging and, with better methods of summing up diagrams, the convergence of the ε-expansions can be improved greatly. In fact, some of the entries in Table 12.1 were originally obtained (or at least rechecked) with the help of this method.

Before closing this sub-section we would like to point out a somewhat unusual piece of information contained in the first-order results obtained above. This refers to the exponent α, for which we note the prediction that for large n it is negative and hence the (singular part of the) specific heat vanishes at $T = T_c$ (which we know to be the case with the spherical model) whereas for small n it is positive and hence the specific heat diverges (which we know to be the case with Ising-like systems). The inversion, from one case to the other, takes place at $n = 4$ where α, according to the first-order expression (39), vanishes. The inclusion of the second-order term in ε upholds this prediction qualitatively but changes it quantitatively. We now have

$$\alpha \approx \frac{4-n}{2(n+8)}\varepsilon - \frac{(n+2)^2(n+28)}{4(n+8)^3}\varepsilon^2, \tag{42}$$

so that, with $\varepsilon = 1$, the inversion takes place between $n = 1$ and $n = 2$–in agreement with the more accurate results quoted in Sec. 12.6.

E. The $1/n$ expansion

Another approach to the problem of determining critical exponents, as functions of d and n, is to adopt the limiting case $n = \infty$ as the starting point and carry out expansions in powers of the small quantity $1/n$. Clearly, the leading terms in these expansions would pertain to the spherical model, which has been studied in Sec. 12.4, and the correction terms would enable us to get some useful information on models with finite n. We quote some first-order results here:[11]

$$\eta = \frac{4(4-d)S_d}{d}\frac{1}{n} + O\left(\frac{1}{n^2}\right), \tag{43}$$

$$\gamma = \frac{2}{d-2}\left\{1 - \frac{6S_d}{n} + O\left(\frac{1}{n^2}\right)\right\}, \tag{44}$$

and

$$\alpha = -\frac{4-d}{d-2}\left\{1 - \frac{8(d-1)S_d}{4-d}\frac{1}{n} + O\left(\frac{1}{n^2}\right)\right\}, \tag{45}$$

where

$$S_d = \frac{\sin\{\pi(d-2)/2\}\Gamma(d-1)}{2\pi\{\Gamma(d/2)\}^2} \quad (2 < d < 4). \tag{46}$$

We note that the coefficients of expansion in this approach are functions of d just as the coefficients of expansion in the preceding approach were functions of n; in this sense, the two expansions are complementary to one another. Unfortunately, there has not been much progress in the evaluation of further terms of these expansions (except for the one mentioned in the note); accordingly, the usefulness of this approach has been rather limited.

F. Other topics

As mentioned earlier, the renormalization group approach has provided a very clear explanation of the concept of *universality*, in that it arises when several physical systems, despite their microscopic structural differences, are governed by a common fixed point and hence display a common critical behavior. Typically, this behavior is linked to the dimensionality d of the physical space, the dimensionality n of the spin vector σ, the range of the spin–spin interaction, etc. Now, depending on the precise nature of the Hamiltonian and the relative importance of the various parameters therein, it is quite possible that under certain circumstances the critical behavior of the system may "cross over" from being characteristic of one fixed point to being characteristic of another fixed point. For instance, we may write for the spin–spin interaction in the lattice

$$H_{\text{int}} = -\tfrac{1}{2}\sum_{\boldsymbol{r},\boldsymbol{r}'}\sum_{\alpha,\beta} J^{\alpha\beta}(\boldsymbol{r}-\boldsymbol{r}')\sigma^{\alpha}(\boldsymbol{r})\sigma^{\beta}(\boldsymbol{r}') \qquad \alpha, \beta = 1, \ldots, n. \tag{47}$$

If the given interaction is isotropic in the physical space but anisotropic in the spin components (assumed three in number), so that $J^{\alpha\beta} = J^{\alpha}\delta_{\alpha\beta}$, then the system is ordinarily supposed to be a Heisenberg ferromagnet; however, the anisotropy of the interaction may finally drive the system towards an Ising fixed point (if one of the J^{α} dominates over the other two) or towards an X–Y fixed point (if two of the J^{α} are equally strong and dominate over the third one). In either case, we encounter what is generally referred to as a *crossover phenomenon*. Similarly, anisotropy in the physical space, $J(\boldsymbol{R}) = J(R_i)\delta_{ij}$ or $K(R)R_iR_j$, may result in a crossover from a d-dimensional behavior to a d'-dimensional behavior (where $d' < d$). In the same vein, one may consider a long-range interaction, $J(\boldsymbol{R}) \sim R^{-d-\sigma}\delta_{ij}$, leading to a critical behavior which, for $\sigma < 2$, is quite seriously σ-dependent; see, for instance, Problem 12.22. However, as σ goes over from the value 2– to 2+, the system crosses over to the universality class characterized by a short-range interaction and remains in that class for all $\sigma > 2$. Crossover phenomena constitute a very fascinating topic in the subject of phase transitions but we cannot pursue them here any more; the interested reader may refer to an excellent review by Aharony (1976).

Another topic of considerable interest deals with the so-called *interfacial phase transitions* in both magnets and fluids. In his seminal paper of 1944, Onsager included in his model a row of "mismatched spins", calculated the boundary tension (or what is more commonly referred to as the interfacial free energy) of this row and examined how this quantity vanished as T approached T_c. In the case of a fluid system, this corresponds to the disappearance of the meniscus between the liquid and the vapor and hence to the vanishing of the conventional surface tension as $T \to T_c-$; see, in this connection, Problem 11.27. A theoretical study of such an interfacial layer involves consideration of the free energy of an *inhomogeneous* system, which has been a subject of considerable research for quite some time. We refer the interested reader to two review articles–by Abraham (1986) and by Jasnow (1986)–for further reading on this topic.

A major ingredient employed by the RG approach is the fact that the critical behavior of a system is invariant under a *scale transformation*. It did not take very long to realize that an important connection exists between this transformation and the well-known *conformal transformation* (in a complex plane), for the latter too is, roughly speaking, a scale transformation in which the scale factor l varies continuously with position. Though, in principle, this connection could be relevant in all dimensions, the most fruitful applications have been in the realm of two dimensions (where the conformal group consists of analytic functions of a complex variable). Among the important results emerging from the conformal transformation approach, one may mention the form of the many-point correlation functions, the critical behavior of finite-sized strips of different sizes and shapes, and the nature of the surface critical effects. For details, see the review article by Cardy (1987).

Another area of interest pertains to the so-called *multicritical points*, for which reference may be made to Lawrie and Sarbach (1984) for theoretical studies and to Knobler and Scott (1984) for experimental results.

13.5. Finite-size scaling

In our study of phase transitions so far, we generally worked in the *thermodynamic limit*, i.e. we started with a lattice of size $L_1 \times \ldots \times L_d$, containing $N_1 \times \ldots \times N_d$ spins (where $N_j = L_j/a$, a being the lattice constant), but at some appropriate stage of the calculation resorted to the limit $L_j \to \infty$. This limiting process is crucial in some important respects; while it simplifies subsequent calculations, it also generates singularities which, as we know, are a hallmark of systems undergoing phase transitions. It is of considerable interest, both theoretically and experimentally, to find out what happens (or does not happen) if some of the L_j are allowed to stay finite. The resulting analysis is quite complicated, but considerable progress has been made in this direction during the last twenty-five years or so. Accordingly, a whole new subject entitled "finite-size scaling" has emerged, of which only a brief summary will be presented here. The reader interested in further details may refer to Barber (1983), Cardy (1988) and Privman (1990).

To fix ideas, we start with a d-dimensional *bulk* system ("bulk" in the sense that it is infinite in all its dimensions) that undergoes a phase transition at a *finite* critical temperature $T_c(\infty)$; clearly, the dimensionality d must be greater than the "lower critical dimension" $d_<$. We also assume that d is less than the "upper critical dimension" $d_>$, so that the critical exponents of the system are d-dependent and obey the hyperscaling relation

$$d\nu = 2 - \alpha = 2\beta + \gamma. \tag{1}$$

We now consider a similar system which is infinite in only d' dimensions, where $d' < d$, and finite in the remaining dimensions; the geometry of this system may be expressed as $L^{d-d'} \times \infty^{d'}$, where $L \gg a$ and, for simplicity, is taken to be the same in all finite dimensions. We may expect this system to be critical at a finite temperature $T_c(L)$, not very different from $T_c(\infty)$. In reality, this is so only if d' too is greater than $d_<$; otherwise, the system continues to be regular at all finite temperatures and the criticality sets in only at $T = 0$.[12] The cases $d' > d_<$ and $d' \leq d_<$, therefore, merit separate treatments.

Our primary goal here is to determine the L-dependence of the various physical quantities pertaining to the system when the system is undergoing a phase transition. We attain this goal by setting up a *finite-size scaling law* that generalizes eqn. (11.10.7) or eqn. (13.3.26) to systems with a finite L. Now, since the only relevant length in the region of a phase transition is the correlation length ξ of the system, it is natural that we scale L with ξ–leading to the combination

$$(L/\xi) \sim Lt^{\nu} = (L^{1/\nu}t)^{\nu}. \tag{2}$$

At the same time, the combination (h/t^{Δ}) appearing in the bulk scaling law may be written as

$$(h/t^{\Delta}) = (hL^{\Delta/\nu})/(L^{1/\nu}t)^{\Delta}, \tag{3}$$

The appropriate combinations of L with t and h, therefore, are $L^{1/\nu}t$ and $L^{\Delta/\nu}h$, respectively. The "singular" part of the free energy density of the system may then be written in the form (see Privman and Fisher, 1984)

$$f^{(s)}(t,h;L) \equiv \frac{A^{(s)}}{VkT} \approx L^{-d}Y(x_1,x_2), \tag{4}$$

where x_1 and x_2 are the scaled variables of the system, viz.

$$x_1 = C_1L^{1/\nu}t, \quad x_2 = C_2L^{\Delta/\nu}h, \tag{5a, b}$$

with

$$t = \frac{T - T_c(\infty)}{T_c(\infty)}, \quad h = \frac{\mu B}{kT} \quad |t|, h \ll 1, \tag{6a, b}$$

while C_1 and C_2 are certain *non-universal* scale factors peculiar to the system under study. Expressed in terms of the variables x_1 and x_2, the function Y is expected to be a *universal* function–common to all systems in the same universality class as the system under study. Of course, the definition of the universality class will now include (apart from the conventional parameters d, n and the range of the spin–spin interaction) the parameter d' as well as the nature of the boundary conditions imposed on the system (which, unless stated otherwise, will be assumed to be *periodic*).

We note that, in the limit $L \to \infty$, expression (4) indeed reduces to eqn. (11.10.7), provided that the function Y has the asymptotic form

$$Y(x_1,x_2) \approx |x_1|^{d\nu}f_{\pm}(x_2/|x_1|^{\Delta}) \quad |x_1|, x_2 \gg 1, \tag{7}$$

thus identifying the non-universal parameters F and G with $C_1^{d\nu}$ and C_2/C_1^{Δ}, respectively. This enables us to write C_1 and C_2 in terms of F and G, viz.

$$C_1 \sim F^{1/(2-\alpha)}, \quad C_2 \sim F^{(\beta+\gamma)/(2-\alpha)}G, \tag{8a, b}$$

which provides a means of determining the non-universal parameters C_1 and C_2 from a knowledge of the bulk parameters F and G; any other factors appearing in (8) would be universal. Once C_1 and C_2 are known, no more non-universal amplitudes are needed to describe the behavior of the system–regardless of whether it is finite-sized or infinite in extent. We are now in a position to examine the consequences of the scaling law (4).

With appropriate differentiations, we obtain from eqn. (4) the following expressions for the zero-field susceptibility per unit volume and the "singular" part of the zero-field specific heat per unit volume:

$$\chi_0(t;L) = -\frac{1}{V}\left(\frac{\partial^2 A^{(s)}}{\partial B^2}\right)_{B=0} \approx -\frac{kT\mu^2C_2^2L^{2\Delta/\nu-d}}{(kT)^2}\left(\frac{\partial^2 Y(x_1,x_2)}{\partial x_2^2}\right)_{x_2=0}$$

$$= \frac{\mu^2C_2^2L^{\gamma/\nu}}{kT}Y_{\chi}(x_1), \tag{9}$$

and

$$c_0^{(s)}(t;L) = -\frac{T}{V}\left(\frac{\partial^2 A^{(s)}}{\partial T^2}\right)_{B=0} \approx -\frac{kT^2C_1^2L^{2/\nu-d}}{T_c^2(\infty)}\left(\frac{\partial^2 Y(x_1,x_2)}{\partial x_1^2}\right)_{x_2=0}$$
$$= \frac{kT^2C_1^2L^{\alpha/\nu}}{T_c^2(\infty)}Y_c(x_1), \tag{10}$$

where $Y_\chi(x_1)$ and $Y_c(x_1)$ are appropriate derivatives of the universal function $Y(x_1,x_2)$ and, hence, are themselves universal. We may, for further analysis, supplement the above results with the corresponding ones for the correlation length of the finite-sized system, viz.

$$\xi(t,h;L) = LS(x_1,x_2) \tag{11}$$

and

$$\xi_0(t;L) = LS(x_1), \tag{12}$$

where $S(x_1) = S(x_1, 0)$; note that the functions $S(x_1, x_2)$ and $S(x_1)$ are also universal. We shall now focus our attention on eqns (9), (10) and (12), and see what messages they deliver in different regimes of the variables T and L.

Case A: $T \gtrsim T_c(\infty)$

With $t > 0$ and $L \gg a$, the variable x_1 in this regime would be positive and much greater than unity. The functions Y_χ, Y_c and S are then expected to assume the form

$$Y_\chi(x_1) \approx \Gamma x_1^{-\gamma}, \quad Y_c(x_1) \approx Ax_1^{-\alpha}, \quad S(x_1) \approx Nx_1^{-\nu}, \tag{13a, b, c}$$

so that we recover the standard bulk results

$$\chi_0 \approx \frac{\mu^2\Gamma C_1^{-\gamma}C_2^2}{kT_c(\infty)}t^{-\gamma}, \quad c_0^{(s)} \approx kAC_1^{2-\alpha}t^{-\alpha}, \quad \xi_0 \approx NC_1^{-\nu}t^{-\nu}, \tag{14a, b, c}$$

complete with non-universal amplitudes and universal factors. The effect of L in this regime appears only as a *correction* to the bulk results; under periodic boundary conditions, such correction terms turn out to be exponentially small, i.e. $O(e^{-L/\xi_0})$ where $\xi_0 \sim a$.[13]

Case B: $T \simeq T_c(\infty)$

This case refers to the "core region" where $|x_1|$ is of order unity and hence $|t|$ is of order $L^{-1/\nu}$; the bulk critical point, $t = 0$, is at the heart of this region. Equations (9), (10) and (12) now yield the first significant results of finite-size scaling, viz.

$$\chi_0 \sim \frac{\mu^2C_2^2}{kT_c(\infty)}L^{\gamma/\nu}, \quad c_0^{(s)} \sim kC_1^2L^{\alpha/\nu}, \quad \xi_0 \sim L. \tag{15a, b, c}$$

Case C: $T < T_c(\infty)$

Here we must distinguish between the cases $d' > d_<$ and $d' \leq d_<$. In the first case, the system becomes critical at a temperature $T_c(L)$ which is not too far removed from $T_c(\infty)$; in the second, the system remains regular at all finite temperatures and becomes critical only at $T = 0$.

(i) $d' > d_<$. In view of the fact that the system is now singular at $T = T_c(L)$ rather than at $T_c(\infty)$, it seems natural to define a *shifted temperature variable* $\dot{t}$ such that

$$\dot{t} = \frac{T - T_c(L)}{T_c(\infty)}; \tag{16}$$

cf. (6a). Thus, for any temperature T,

$$\dot{t} = t - \tau \quad \tau = [T_c(L) - T_c(\infty)]/T_c(\infty), \tag{17}$$

which prompts us to define a *shifted scaled variable*

$$\dot{x}_1 = C_1 L^{1/\nu} \dot{t} = x_1 - X \quad X = C_1 L^{1/\nu} \tau; \tag{18}$$

cf. (5a). Clearly, the scaling functions governing the system would be singular at $\dot{x}_1 = 0$, i.e. at $x_1 = X$. With no other arguments present, we presume that $|X|$ will be of order unity; the shift in T_c is this given by

$$|\tau| = |X| C_1^{-1} L^{-1/\nu} = O(L^{-1/\nu}). \tag{19}$$

Now, as $T \to T_c(L)$, the correlation length of the system approaches infinity–with the result that, insofar as the qualitative nature of the critical behavior is concerned, the variable L, however large, becomes essentially unimportant. The behavior of the system, in the immediate neighborhood of $T_c(L)$, would, therefore, be characteristic of a d'-dimensional bulk system rather than of a d-dimensional one; accordingly, it would be governed by the critical exponents $\dot{\alpha}, \dot{\beta}, \ldots$ pertaining to d' dimensions rather than by the exponents $\alpha, \beta, \ldots$ pertaining to d dimensions. We therefore expect that, as $\dot{x}_1 \to 0$, the functions Y_χ, Y_c and S of eqns (9), (10) and (12) assume the form

$$Y_\chi(x_1) \approx \dot{\Gamma} \dot{x}_1^{-\dot{\gamma}}, \quad Y_c(x_1) \approx \dot{A} \dot{x}_1^{-\dot{\alpha}}, \quad S(x_1) \approx \dot{N} \dot{x}_1^{-\dot{\nu}}, \tag{20a, b, c}$$

with the result that

$$\chi_0 \approx \left[\mu^2/kT_c(\infty)\right] \dot{\Gamma} C_1^{-\dot{\gamma}} C_2^2 L^{(\gamma-\dot{\gamma})/\nu} \dot{t}^{-\dot{\gamma}}, \tag{21a}$$

$$c_0^{(s)} \approx k \dot{A} C_1^{2-\dot{\alpha}} L^{(\alpha-\dot{\alpha})/\nu} \dot{t}^{-\dot{\alpha}}, \tag{21b}$$

and

$$\xi_0 \approx \dot{N} C_1^{-\dot{\nu}} L^{(\nu-\dot{\nu})/\nu} \dot{t}^{-\dot{\nu}}. \tag{21c}$$

It is obvious that, for $\dot{t} < 0$ but such that $|\dot{t}| \ll 1$, the same results would hold–except that $\dot{t}$ would be replaced by $|\dot{t}|$.

Expressions (21), insofar as the dependence on L and $\dot{t}$ is concerned, have been verified by direct calculation on a variety of systems over the years; for details, see the review articles by Barber (1983) and Privman (1990) cited earlier. More recently, Allen and Pathria (1989) have verified the non-universal amplitudes as well by carrying out an explicit calculation for the spherical model ($n = \infty$) in general geometry $L^{d-d'} \times \infty^{d'}$, with both d and d' greater than $d_<$. Remarkably enough, they found that, just as the critical exponents $\dot{\alpha}, \dot{\beta}, \dots$ are the same functions of d' as the exponents $\alpha, \beta, \dots$ are of d, the universal coefficients $\dot{\Gamma}, \dot{A}, \dots$ too are the same functions of d' as the coefficients $\Gamma, A, \dots$ are of d; the same is true of the coefficients appearing in the presence of a magnetic field (see Allen and Pathria, 1991). One wonders if this would be the case for general n too!

(ii) $d' \le d_<$. In this case the singularity of the problem lies at $T = 0$, so that at *all* finite temperatures the system is regular and hence expressible by smooth, analytic functions. We may, therefore, generalize the scaling law (4) to apply at all temperatures down to $T = 0$ by simply allowing the scale factors C_1 and C_2 to become T-dependent and writing (after Singh and Pathria, 1985b, 1986a)

$$x_1 = \tilde{C}_1(T)L^{1/\nu}t, \quad x_2 = \tilde{C}_2(T)L^{\Delta/\nu}h, \tag{22a, b}$$

leaving t and h unchanged; the quantities $\tilde{C}_1$ and $\tilde{C}_2$ must be such that, as $T \to T_c(\infty)$ from below, they approach the quantities C_1 and C_2 of (5). Expressions (9) and (10) now take the form

$$\chi_0(t; L) = \frac{\mu^2\tilde{C}_2^2L^{\gamma/\nu}}{kT}Y_\chi(x_1) \tag{23}$$

and

$$c_0^{(s)}(t; L) = kT^2\left[\frac{\partial}{\partial T}(\tilde{C}_1 t)\right]^2 L^{\alpha/\nu}Y_c(x_1), \tag{24}$$

respectively, while expression (12) remains formally the same. Now, as we approach the critical temperature T_c (which is zero here), we again expect the quantities χ_0, $c_0^{(s)}$ and ξ_0 to behave in a manner characteristic of a d'-dimensional bulk system. Let us assume that, in that limit, our scale factors behave as

$$\tilde{C}_1 \sim T^r, \quad \tilde{C}_2 \sim T^s \quad (T \to 0), \tag{25a, b}$$

and our universal functions as

$$Y_\chi(x_1) \sim |x_1|^\theta, \quad Y_c(x_1) \sim |x_1|^\phi, \quad S(x_1) \sim |x_1|^\sigma \quad (x_1 \to -\infty). \tag{26a, b, c}$$

The resulting T-dependence of χ_0, $c_0^{(s)}$ and ξ_0 then is

$$\chi_0 \sim T^{2s-1+\theta r}, \quad c_0^{(s)} \sim T^{(2+\phi)r}, \quad \xi_0 \sim T^{\sigma r}. \tag{27a, b, c}$$

The corresponding results for n-vector, d'-dimensional models (with $n \ge 2$ and $d' < d_<$, where $d_< = 2$) are[14]

$$\chi_0 \sim T^{-2/(2-d')}, \quad c_0^{(s)} \sim T^{d'/(2-d')}, \quad \xi_0 \sim T^{-1/(2-d')}. \tag{28a, b, c}$$

Comparing (27) with (28), we infer that

$$\theta = \frac{1}{r}\left[1 - 2s - \frac{2}{2-d'}\right], \quad \phi = \frac{d'}{r(2-d')} - 2, \quad \sigma = \frac{-1}{r(2-d')}. \tag{29a, b, c}$$

Very shortly we shall find, see eqns (44), that

$$r = -1/\nu(d-2), \quad s = \beta/\nu(d-2), \tag{30a, b}$$

with the results

$$\theta = 2\beta + \frac{\nu d'(d-2)}{(2-d')}, \quad \phi = -\frac{\nu d'(d-2)}{(2-d')} - 2, \quad \sigma = \frac{\nu(d-2)}{(2-d')}. \tag{31a, b, c}$$

The accompanying L-dependence of the various quantities turns out to be

$$\chi_0 \sim L^{(\gamma+\theta)/\nu} \sim L^{2(d-d')/(2-d')} \tag{32a}$$

$$c_0^{(s)} \sim L^{(\alpha+\phi)/\nu} \sim L^{-2(d-d')/(2-d')} \tag{32b}$$

and

$$\xi_0 \sim L^{1+\sigma/\nu} \sim L^{(d-d')/(2-d')}. \tag{32c}$$

It is remarkable that in these last expressions the critical exponents pertaining to d dimensions have disappeared altogether and the resulting powers of L depend *entirely* on the geometry of the system. Expression (32a) agrees with the earlier results for χ_0 pertaining to a "block" geometry ($d' = 0$) and to a "cylindrical" geometry ($d' = 1$), viz.

$$\chi_0 \sim \begin{cases} L^d & \text{for} \quad d' = 0 \qquad (33a) \\ L^{2(d-1)} & \text{for} \quad d' = 1; \qquad (33b) \end{cases}$$

see Fisher and Privman (1985). For $d' = 2$, the L-dependence of the various quantities studied here becomes exponential instead of a power law.

To obtain results valid for *all* T in the range $0 < T < T_c(\infty)$, we need to know the full T-dependence of the scale factors $\tilde{C}_1$ and $\tilde{C}_2$. It turns out that this too can be determined from the properties of the corresponding bulk system–in particular, from the field-free bulk correlation function $G(\boldsymbol{R}, T)$, which is known to possess the following forms:

$$G(\boldsymbol{R}, T) \sim R^{-(d-2+\eta)} \quad T = T_c(\infty) \tag{34}$$

and

$$G(\boldsymbol{R}, T) = m_0^2(T) + A(T)R^{-(d-2)} \quad T < T_c(\infty), \tag{35}$$

where $m_0(T)$ is the order parameter of the bulk system and $A(T)$ another system-dependent parameter.[15] Now, the correlation function of the finite-sized system may be written in the scaled form

$$G(\boldsymbol{R}, t, h; L) \approx \tilde{D}(T)R^{-(d-2+\eta)}Z(x_1, x_2, x_3). \tag{36}$$

where the scaled variables x_1 and x_2 are given by eqns (22) while $x_3 = \boldsymbol{R}/L$; as usual, the scale factor $\tilde{D}(T)$ is non-universal while the function $Z(x_1, x_2, x_3)$ is universal. Expression (36) already conforms to (34), with $x_1 = x_2 = x_3 = 0$. For conformity with (35), the function Z must possess the following asymptotic form:

$$Z(x_1, x_2, x_3) \approx Z_1(|x_1|^\nu x_3)^{d-2+\eta} + Z_2(|x_1|^\nu x_3)^\eta \quad x_1 \to -\infty, x_2 = 0, x_3 \to 0, \tag{37}$$

with

$$Z_1 = \frac{m_0^2(T)}{\tilde{D}(T)[\tilde{C}_1(T)|t|]^{\nu(d-2+\eta)}}, \quad Z_2 = \frac{A(T)}{\tilde{D}(T)[\tilde{C}_1(T)|t|]^{\nu\eta}}. \tag{38a, b}$$

It follows that

$$\tilde{C}_1(T)|t| \sim \left[\frac{m_0^2(T)}{A(T)}\right]^{1/\nu(d-2)}, \quad \tilde{D}(T) \sim \left[\frac{A^\beta(T)}{m_0^{\nu\eta}(T)}\right]^{2/\nu(d-2)}; \tag{39a, b}$$

here, use has been made of the fact that

$$\nu(d-2+\eta) = (2-\alpha) - \gamma = 2\beta. \tag{40}$$

We shall now establish a relationship between the scale factors $\tilde{C}_2$ and $\tilde{D}$. For this, we utilize the fluctuation–dissipation relation (11.11.12) which, with the help of expression (36), gives for the zero-field susceptibility per unit volume

$$\begin{aligned}\chi_0(t; L) &= \frac{\mu^2 \tilde{D}(T)}{a^{2d} kT} \int \frac{Z(x_1, 0, R/L)}{R^{d-2+\eta}} d^d R \\ &= \frac{\mu^2 \tilde{D}(T)}{a^{2d} kT} L^{2-\eta} Z_\chi(x_1),\end{aligned} \tag{41}$$

where Z_χ is another universal function. Comparing (41) with (23), we get

$$\tilde{D}(T) \sim a^{2d} \tilde{C}_2^2(T). \tag{42}$$

Equation (39b) now gives

$$\tilde{C}_2(T) \sim a^{-d} [A^\beta(T)/m_0^{\nu\eta}(T)]^{1/\nu(d-2)}. \tag{43}$$

Equations (39) and (43) give us the full T-dependence of the scale factors $\tilde{C}_1$, $\tilde{C}_2$ and $\tilde{D}$ for *all* T in the range $0 < T < T_c(\infty)$; these equations were first derived by Singh and Pathria (1987a).

Before utilizing these results we note that since, in the limit $T \to 0$, $m_0(T)$ approaches a constant value while $A(T) \sim T$, expressions (39a) and (43) yield

$$\tilde{C}_1|t| \sim T^{-1/\nu(d-2)}, \quad \tilde{C}_2 \sim T^{\beta/\nu(d-2)}, \tag{44a, b}$$

exactly as stipulated in eqns (25) and (30). As for the T-dependence of the quantities χ_0, $c_0^{(s)}$ and ξ_0, we observe that, regardless of whether we keep L fixed and let $T \to 0$ *or* keep T fixed and let $L \to \infty$, in either case $x_1 \to -\infty$; the asymptotic forms (26) of the universal functions Y_χ, Y_c and S, therefore, apply throughout the region under study. Now, with θ, ϕ and σ given by eqns (31), our final results for χ_0, $c_0^{(s)}$ and ξ_0 turn out to be

$$\chi_0 \sim \frac{\mu^2 A(T)}{a^{2d} kT} \left[\frac{m_0^2(T)}{A(T)}\right]^{2/(2-d')} L^{2(d-d')/(2-d')}, \tag{45}$$

$$c_0^{(s)} \sim k \left\{ T \frac{\partial}{\partial T} \left[\frac{m_0^2(T)}{A(T)}\right] \right\}^2 \left[\frac{m_0^2(T)}{A(T)}\right]^{-(4-d')/(2-d')} L^{-2(d-d')/(2-d')}, \tag{46}$$

and

$$\xi_0 \sim \left[\frac{m_0^2(T)}{A(T)}\right]^{1/(2-d')} L^{(d-d')/(2-d')}, \tag{47}$$

complete with non-universal amplitudes. Comparing (45) with (47), we find that, in the regime under study,

$$\chi_0/\xi_0^2 \sim \mu^2 A(T)/a^{2d} kT, \tag{48}$$

a function of T only. In the case of the spherical model, since $A(T)$ is proportional to T at *all* $T < T_c(\infty)$, see eqn. (12.4.71), the quantity χ_0/ξ_0^2 is a constant — independent of both T and L; see also eqn. (12.4.64).

It is important to note that the above formulation ties very neatly with the one provided by the scale factors C_1 and C_2 of eqns (5) that covered cases A and B pertaining to the regions $T \gtrsim T_c(\infty)$ and $T \simeq T_c(\infty)$, respectively. To see this, we observe that as $T \to T_c(\infty)$ from below, $m_0(T)$ becomes $\sim |t|^\beta$ and $A(T) \sim |t|^{\nu\eta}$; expressions (39a) and (43) then assume the form

$$\tilde{C}_1(T)|t| \sim |t|^{(2\beta-\nu\eta)/\nu(d-2)} \sim |t|^1 \tag{49a}$$

and

$$\tilde{C}_2(T) \sim |t|^0. \tag{49b}$$

Clearly, $\tilde{C}_1(T)$ and $\tilde{C}_2(T)$ now tend to some constant values which may be identified with C_1 and C_2–thus providing a *unified* formulation, through the *same* universal functions $Y(x_1, x_2)$, $S(x_1, x_2)$ and $Z(x_1, x_2, x_3)$, covering the regions of *both* first-order and second-order phase transition. Remarkable though it is, this finding is not really surprising because, with L finite and $d' \le d_<$, the system is critical only at $T = 0$ and analytic everywhere else; so, it should indeed be expressible by a single set of functions throughout. Of course, as $L \to \infty$, the criticality spreads all the way from $T = 0$ to $T = T_c(\infty)$.

As regards the spin dimensionality n, our results for cases A, B and C(i) are quite general; only in case C(ii) did we specialize to systems with *continuous* symmetry ($n \ge 2$). With a slight modification, the case of discrete symmetry ($n = 1$) can also be taken care of. The net result essentially is the replacement of the number 2 in eqns (28) and henceforth by the "lower critical dimension", $d_<$, of the system–leading to results such as[16]

$$\chi_0 \sim L^\zeta, \quad c_0^{(s)} \sim L^{-\zeta}, \quad \xi_0 \sim L^{\zeta/d_<} \quad T < T_c(\infty), \tag{50a, b, c}$$

with $\zeta = d_<(d - d')/(d_< - d')$; cf. eqns (32). Once again, the L-dependence of the various quantities of interest follows a power law, which changes to an exponential at $d' = d_<$; in the case of scalar models, this happens at $d' = 1$.

Throughout this discussion we have assumed that the total dimensionality d of the system is less than the "upper critical dimension" $d_>$. The case $d \ge d_>$ presents some special problems but the net result is that, while the situation in the region $T < T_c(\infty)$ is described by the same set of expressions as above, that in the region $T \simeq T_c(\infty)$ is considerably modified. For instance, one now gets in the region $T \simeq T_c(\infty)$, for $d > d_>$ and $d' < d_<$,

$$\chi_0 \sim L^{2(d-d')/(d_>-d')}, \quad c_0^{(s)} \sim L^0, \quad \xi_0 \sim L^{(d-d')/(d_>-d')}, \tag{51a, b, c}$$

which may be compared with the corresponding results, (15a, b, c), for $d < d_>$. Furthermore, if $d = d_>$ and/or $d' = d_<$, factors containing $\ln L$ appear along with the power l [illegible] displayed in (51). For details, see Singh and Pathria (1986b, 1988, 1992).

Finally we would like to emphasize the fact that finite-size effects in a given system are quite sensitive to the choice of the boundary conditions imposed on the system. For simplicity, we assumed the boundary conditions to be *periodic*. In real situations there may be reasons to adopt different boundary conditions such as *antiperiodic*, *free*, etc. This, in general, changes the mathematical character of the finite-size effects and the finite-size corrections in not only the singular part(s) of the various quantities studied but in their regular part(s) as well. For comparison between theory and experiment, this aspect of the problem is of vital importance and deserves a close scrutiny. For lack of space we cannot go into this matter any further here; the interested reader may refer to a review article by Privman (1990), where other references on this topic can also be found. An allied subject in this context is the "critical behavior of surfaces", for which reference may be made to Binder (1983) and Diehl (1986).

Problems

13.1. Show that the decimation transformation of a one-dimensional Ising model, with $l = 2$, can be written in terms of the transfer matrix P as

$$P\{K'\} = P^2\{K\}, \tag{1}$$

where K and K' are the coupling constants of the original and the decimated lattice, respectively. Next show that, with P given by

$$(P\{K\}) = e^{K_0}\begin{pmatrix} e^{K_1+K_2} & e^{-K_1} \\ e^{-K_1} & e^{K_1-K_2} \end{pmatrix} \tag{2}$$

see eqn. (12.1.4), relation (1) leads to the same transformation equations among K and K' as (13.2.8a, b, and c).

13.2. Verify that expression (15) of Sec. 13.2 indeed satisfies the functional eqn. (14) for the field-free Ising model in one dimension. Next show (or at least verify) that, with the field present, the functional eqn. (11), with K' given by (8), is satisfied by the more general expression

$$f(K_1, K_2) = -\ln\left[e^{K_1}\cosh K_2 + \left\{e^{-2K_1} + e^{2K_1}\sinh^2 K_2\right\}^{1/2}\right].$$

13.3. Verify that expression (32) of Sec. 13.2 indeed satisfies the functional eqn. (31) for the field-free spherical model in one dimension. Next show (or at least verify) that, with the field present, the functional eqn. (27), with K' given by (25), is satisfied by the more general expression

$$f(K_1, K_2, \Lambda) = \frac{1}{2}\ln\left[\frac{\Lambda + \sqrt{(\Lambda^2 - K_1^2)}}{2\pi}\right] - \frac{K_2^2}{4(\Lambda - K_1)},$$

where Λ is determined by the constraint equation

$$\frac{\partial f}{\partial \Lambda} = \frac{1}{2\sqrt{(\Lambda^2 - K_1^2)}} + \frac{K_2^2}{4(\Lambda - K_1)^2} = 1.$$

13.4. Consider the field-free spherical model in one dimension whose partition function is given by eqn. (13.2.24) as well as by (13.2.19), with $K_2' = K_2 = 0$. Substituting $\sigma_j' = (2\Lambda/K_1)^{1/2}s_j'$ in the former and comparing the resulting expression with the latter, show that

$$Q_N(K_1, \Lambda) = \left(\frac{2\pi}{K_1}\right)^{N'/2} Q_{N'}(K_1, \Lambda'),$$

where $N' = \frac{1}{2}N$ and $\Lambda' = (2\Lambda^2/K_1) - K_1$. This leads to the functional relation

$$f(K_1, \Lambda) = -\frac{1}{4}\ln\left(\frac{2\pi}{K_1}\right) + \frac{1}{2}f(K_1, \Lambda').$$

Check that expression (13.2.32) satisfies this relation.

13.5. An approximate way of implementing an RG transformation on a square lattice is provided by the so-called Migdal–Kadanoff transformation[17] shown in Fig. 13.8. It consists of two essential steps:

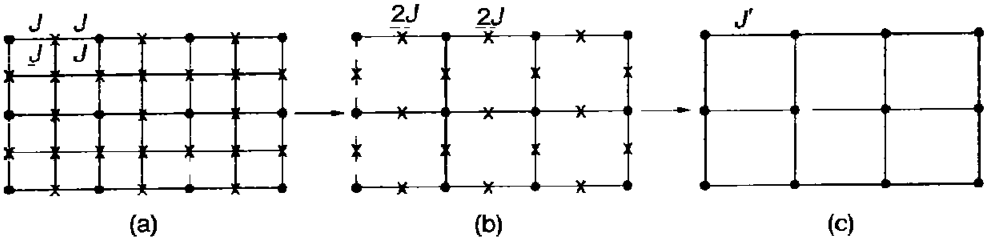


Fig. 13.8. Migdal–Kadanoff transformation on a square lattice.

(i) First, one-half of the bonds in the lattice are simply removed so as to change the length scale of the lattice by a factor of 2; to compensate for this, the coupling strength of the remaining bonds is changed from J to $2J$. This takes us from Fig. 13.8(a) to Fig. 13.8(b).

(ii) Next, the sites marked by crosses in Fig. 13.8(b) are eliminated by a set of one-dimensional decimation transformations, leading to Fig. 13.8(c) with coupling strength J'.

(a) Show that the recursion relation for a spin-$\frac{1}{2}$ Ising model on a square lattice, according to the above transformation, is

$$x' = 2x^2/(1+x^4),$$

where $x = \exp(-2K)$ and $x' = \exp(-2K')$. Disregarding the trivial fixed points $x^* = 0$ and $x^* = 1$, show that the nontrivial fixed point of this transformation is

$$x^* = \frac{1}{3}\left[-1 + 2\sqrt{2}\sinh\left\{\frac{1}{3}\sinh^{-1}\frac{17}{2\sqrt{2}}\right\}\right] \simeq 0.5437;$$

cf. the actual value of x_c, which is $(\sqrt{2}-1) \simeq 0.4142$.

(b) Linearizing around the nontrivial fixed point, show that the eigenvalue λ of this transformation is

$$\lambda = 2(1-x^*)/x^* \simeq 1.6785$$

and hence the exponent $\nu = \ln 2/\ln\lambda \simeq 1.338$; cf. the actual value of ν, which is 1.

13.6. Consider the linearized RG transformation (13.3.12), with

$$\mathscr{A}_l^* = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

such that $(a_{11}a_{22} - a_{12}a_{21}) \neq 0$. We now introduce the "generalized coordinates" u_1 and u_2 through eqns (13.3.13); clearly, u_1 and u_2 are certain linear combinations of the system parameters k_1 and k_2.

(a) Show that the slopes of the lines $u_1 = 0$ and $u_2 = 0$, in the (k_1, k_2)-place, are

$$m_1 = \frac{a_{21}}{\lambda_2 - a_{22}} = \frac{\lambda_2 - a_{11}}{a_{12}} \quad \text{and} \quad m_2 = \frac{a_{21}}{\lambda_1 - a_{22}} = \frac{\lambda_1 - a_{11}}{a_{12}},$$

respectively; here, λ_1 and λ_2 are the eigenvalues of the matrix $\mathscr{A}_l^*$. Verify that the product $m_1 m_2 = -a_{21}/a_{12}$ and hence the two lines are mutually perpendicular if and only if $a_{12} = a_{21}$.

(b) Check that in the special case when $a_{12} = 0$ but $a_{21} \neq 0$, the above slopes assume the simple form

$$m_1 = \infty \quad \text{and} \quad m_2 = a_{21}/(a_{11} - a_{22})$$

whereas in the special case when $a_{21} = 0$ but $a_{12} \neq 0$, they become

$$m_1 = (a_{22} - a_{11})/a_{12} \quad \text{and} \quad m_2 = 0;$$

note that Fig. 13.7 pertains to the latter case.

(c) Examine as well the cases where either a_{11} or a_{22} is zero; Fig. 13.6 pertains to the latter of these cases.

13.7. Check that the critical exponents (13.4.38)–(13.4.40) in the limit $n \to \infty$ agree with the corresponding exponents for the spherical model of Sec. 12.4 with $d \lesssim 4$.

13.8. Show, from eqns (13.4.43)–(13.4.46), that for $d \lesssim 4$

$$\eta \simeq \frac{1}{2n}\varepsilon^2, \quad \gamma \simeq 1 + \frac{1}{2}\left(1 - \frac{6}{n}\right)\varepsilon, \quad \alpha \simeq -\frac{1}{2}\left(1 - \frac{12}{n}\right)\varepsilon,$$

where $\varepsilon = (4 - d) \ll 1$. Check that these results agree with the ones following from eqns (13.4.38)–(13.4.40) for $n \gg 1$.

13.9. Using the various scaling relations, derive from eqns (13.4.43)–(13.4.45) comparable expressions for the remaining exponents β, δ and ν. Repeat for these exponents the exercise suggested in the preceding problem.

Notes

[1] For simplicity, we employ the language of the scalar model here.

[2] Some authors derive eqn. (12) from (9) by choosing l to be t^{-1/y_t}. As will be seen shortly, see eqn. (19), this amounts to letting l be $O(\xi/a)$, which violates the requirement, $l \ll \xi/a$, mentioned earlier.

[3] This is further related to the fact that the critical exponent η in this case is equal to 1; see eqn. (13.1.23) which, with $d = 1$, gives: $\sigma'(r') = \sigma(r)$.

[4] In general, the vector K^* will contain components not present in the original problem. In such a case, one has to locate, on the critical surface, a point K_c that is free of these "unnecessary" components; since ξ is infinite at $K = K_c$ as well, the latter may be identified as the *critical point* of the given system. As will be seen in the sequel, the critical behavior of the system is still determined by the flow pattern in the neighborhood of the fixed point.

[5] This requirement makes the set of operators $\mathscr{R}_l$ a *semi-group*–not a *group* because the inverse of $\mathscr{R}_l$ does not exist. The reason for the non-existence of $\mathscr{R}_l^{-1}$ is that once a number of degrees of freedom of the system are summed over there is no definitive way of recreating them.

[6] Remember that at each and every point on the critical surface–in this case, the critical curve–the correlation length is infinite; accordingly, each and every such point is qualified to be a critical point. The *physical* critical point is one that is free of unnecessary parameters.

[7] The result obtained through numerical analysis was 0.3921.

[8] In this reference one can also find a systematic method of constructing the scaling function $f_s(u_1, u_2, \ldots)$ from a knowledge of the regular function K_0' of eqn. (13.3.4).

[9] It can be shown that, by a suitable transformation, the Ising model ($n = 1$), which is a *discrete* (rather than a *continuous*) model, can also be rendered "continuous" with a probability distribution similar to (13.3.31). For details, see Fisher (1983), Appendix A.

[10] It can be seen quite easily that the generalized coordinate u_2 is directly proportional to Δu, making u an irrelevant variable of the problem; see Problem 13.6, with $a_{21} = 0$.

[11] In the special case $d = 3$, the expansion for η is known to a higher order, viz.

$$\eta = \frac{8}{3\pi^2 n} - \left(\frac{8}{3}\right)^3 \frac{1}{\pi^4 n^2} + O\left(\frac{1}{n^3}\right).$$

[12] For the special case $d' = 0$, when the system is fully finite, this point has already been emphasized in Sec. 11.1. Here we assert that, even when some of the system dimensions are infinite (and hence the total number of spins is infinite), a *finite-temperature singularity* does not arise unless the number of those infinite dimensions exceeds $d_<$.

[13] See, for instance, Luck (1985) and Singh and Pathria (1985b, 1987a).

[14] For the spherical model ($n = \infty$), these results appear in Sec. 12.4; see eqns (12.4.34, 35 and 64). Since the criticality in this case occurs at absolute zero, these results hold for all models with *continuous* symmetry, i.e. with $n \geq 2$. See, for instance, Sec. 12.2, where n is general but $d' = 1$.

[15] Note that the exponent η appears only in eqn. (34) and not in (35); for details, see Schultz *et al.* (1964) and Fisher *et al.* (1973).

[16] See, for instance, Singh and Pathria (1986b).

[17] See Kadan [illegible] '976a).

CHAPTER 14

FLUCTUATIONS

IN THIS course we have been mostly concerned with the evaluation of *statistical averages* of the various physical quantities; these averages represent, with a high degree of accuracy, the results expected from relevant measurements on the given system *in equilibrium*. Nevertheless, there do occur *deviations* from, or *fluctuations* about, these mean values. Though they are generally small, their study is of great physical interest for several reasons.

Firstly, it enables us to develop a mathematical scheme with the help of which the magnitude of the relevant fluctuations, under a variety of physical situations, can be estimated. Not surprisingly, we find that while in a single-phase system the fluctuations are thermodynamically negligible they can assume considerable importance in multi-phase systems, especially in the neighborhood of a critical point. In the latter case, we obtain a rather high degree of *spatial correlation* among the molecules of the system which, in turn, gives rise to phenomena such as *critical opalescence*.

Secondly, it provides a natural framework for understanding a class of phenomena that come under the heading "Brownian motion"; these phenomena relate properties such as the mobility of a fluid system, its coefficient of diffusion, etc. with temperature through the so-called *Einstein relations*. The mechanism of Brownian motion is vital in formulating, and in a certain sense answering, questions as to how "a given physical system, which is not in a state of equilibrium, finally approaches such a state" while "a physical system, which is already in a state of equilibrium, persists to stay in that state".

Thirdly, the study of fluctuations, as a function of time, leads to the concept of certain "correlation functions" that play a vital role in relating the dissipative properties of a system, such as the viscous resistance of a fluid or the electrical resistance of a conductor, with the microscopic properties of the system in a state of equilibrium; this relationship (between irreversible processes on one hand and equilibrium properties on the other) manifests itself in the so-called *fluctuation–dissipation theorem*. At the same time, a study of the "frequency spectrum" of fluctuations, which is related to the time-dependent correlation function through the fundamental theorem of Wiener and Khintchine, is of considerable value in assessing the "noise" met with in electrical circuits as well as in the transmission of electromagnetic signals.

14.1. Thermodynamic fluctuations

We begin by deriving a *probability distribution law* for the fluctuations of certain basic thermodynamic quantities pertaining to a given physical system; the *mean square fluctuations* can then be evaluated, in a straightforward manner, with the help of this law. We assume that the given system, which may be referred to as 1, is embedded in a reservoir, which may be referred to as 2, such that a mutual exchange of energy, and of volume, can take place between the two; of course, the overall energy E and the overall volume V are supposed to be fixed. For convenience, we do not envisage an exchange of particles here, so the numbers N_1 and N_2 remain individually constant. The equilibrium division of E into $\overline{E}_1$ and $\overline{E}_2$, and of V into $\overline{V}_1$ and $\overline{V}_2$, must be such that parts 1 and 2 of the composite system $(1+2)$ have a *common* temperature T^* and a *common* pressure P^*; see Secs 1.2 and 1.3, especially eqn. (1.3.6). Of course, the entropy of the composite system will have its largest value in the equilibrium state; in any other state, such as the one characterized by a fluctuation, it must have a lower value. Let ΔS denote the deviation, in the entropy of the composite system, from its equilibrium value S_0. Then

$$\Delta S \equiv S - S_0 = k \ln \Omega_f - k \ln \Omega_0, \tag{1}$$

where Ω_f (or Ω_0) denotes the number of distinct microstates of the system $(1+2)$ in the presence (or in the absence) of a fluctuation from the equilibrium state; see eqn. (1.2.6). The probability that the proposed fluctuation may occur is then given by

$$p \propto \Omega_f \propto \exp(\Delta S/k); \tag{2}$$

see Sec. 3.1, especially eqn. (3.1.3). In terms of other thermodynamic quantities, we may write

$$\Delta S = \Delta S_1 + \Delta S_2 = \Delta S_1 + \int_0^f \frac{dE_2 + P_2 dV_2}{T_2}; \tag{3}$$

note that the pressure P_2 and the temperature T_2 of the reservoir may, in principle, vary during the build-up of the fluctuation. Now, even if the fluctuation is sizable from the point of view of system 1 it will be small from the point of view of 2. The "variables" P_2 and T_2 may, therefore, be replaced by the constants P^* and T^*, respectively; at the same time, the increments dE_2 and dV_2 may be replaced by $-dE_1$ and $-dV_1$, respectively. Equation (3) then becomes

$$\Delta S = \Delta S_1 - (\Delta E_1 + P^* \Delta V_1)/T^*. \tag{4}$$

Accordingly, formula (2) takes the form

$$p \propto \exp\{-(\Delta E_1 - T^* \Delta S_1 + P^* \Delta V_1)/kT^*\}. \tag{5}$$

Clearly, the probability distribution law (5) does not depend, in any manner, on the peculiarities of the reservoir in which the given system was supposedly embedded. Formula (5), therefore, applies equally well to a system that attained equilibrium in a statistical ensemble (or, for that matter, to any macroscopic part of a given system itself). Consequently, we may drop the suffix 1 from the symbols ΔE_1, ΔS_1 and ΔV_1, and the star from the symbols P^* and T^*, and write

$$p \propto \exp\{-(\Delta E - T\,\Delta S + P\,\Delta V)/kT\}. \tag{6}$$

In most cases, the fluctuations are exceedingly small in magnitude; the quantity ΔE may, therefore, be expanded as a Taylor series about the equilibrium value $(\Delta E)_0 = 0$, with the result

$$\begin{aligned}\Delta E = {} & \left(\frac{\partial E}{\partial S}\right)_0 \Delta S + \left(\frac{\partial E}{\partial V}\right)_0 \Delta V \\ & + \frac{1}{2}\left[\left(\frac{\partial^2 E}{\partial S^2}\right)_0 (\Delta S)^2 + 2\left(\frac{\partial^2 E}{\partial S\,\partial V}\right)_0 \Delta S \Delta V + \left(\frac{\partial^2 E}{\partial V^2}\right)_0 (\Delta V)^2\right] + \cdots\end{aligned} \tag{7}$$

Substituting (7) into (6) and retaining terms up to second order only, we obtain

$$p \propto \exp\{-(\Delta T\,\Delta S - \Delta P\,\Delta V)/2kT\}; \tag{8}$$

here, use has been made of the relations

$$\left(\frac{\partial E}{\partial S}\right)_0 = T, \quad \left(\frac{\partial E}{\partial V}\right)_0 = -P, \tag{9}$$

and of the fact that the expression within the long brackets in (7) is equivalent to

$$\Delta\left(\frac{\partial E}{\partial S}\right)_0 \Delta S + \Delta\left(\frac{\partial E}{\partial V}\right)_0 \Delta V = \Delta T\,\Delta S - \Delta P\,\Delta V. \tag{10}$$

With the help of (8), the mean square fluctuations of various physical quantities, and the statistical correlations among different fluctuations, can be readily calculated. We note, however, that of the four Δ terms appearing in this formula only two can be chosen independently; the other two must assume the role of "derived quantities". For instance, if we choose ΔT and ΔV to be the *independent variables*, then ΔS and ΔP can be written as

$$\Delta S = \left(\frac{\partial S}{\partial T}\right)_V \Delta T + \left(\frac{\partial S}{\partial V}\right)_T \Delta V = \frac{C_V}{T}\Delta T + \left(\frac{\partial P}{\partial T}\right)_V \Delta V \tag{11}$$

and

$$\Delta P = \left(\frac{\partial P}{\partial T}\right)_V \Delta T + \left(\frac{\partial P}{\partial V}\right)_T \Delta V = \left(\frac{\partial P}{\partial T}\right)_V \Delta T - \frac{1}{\kappa_T V}\Delta V, \tag{12}$$

κ_T being the *isothermal compressibility* of the system. Substituting (11) and (12) into (8), we get

$$p \propto \exp\left\{-\frac{C_V}{2kT^2}(\Delta T)^2 - \frac{1}{2kT\kappa_T V}(\Delta V)^2\right\}, \tag{13}$$

which shows that the fluctuations in T and V are *statistically independent, Gaussian variables*. A quick glance at (13) yields the results

$$\overline{(\Delta T)^2} = \frac{kT^2}{C_V}, \quad \overline{(\Delta V)^2} = kT\kappa_T V, \tag{14a}$$

while

$$\overline{(\Delta T\, \Delta V)} = 0. \tag{14b}$$

Similarly, if we choose ΔS and ΔP as our *independent variables*, we are led to the distribution law

$$p \propto \exp\left\{-\frac{1}{2kC_P}(\Delta S)^2 - \frac{\kappa_S V}{2kT}(\Delta P)^2\right\}, \tag{15}$$

whence

$$\overline{(\Delta S)^2} = kC_P, \quad \overline{(\Delta P)^2} = \frac{kT}{\kappa_S V}, \tag{16a}$$

while

$$\overline{(\Delta S\, \Delta P)} = 0; \tag{16b}$$

here, κ_S denotes the *adiabatic compressibility* of the system.

We note that, in general, the mean square fluctuation of an extensive quantity is directly proportional to the size of the system while that of an intensive quantity is inversely proportional to the same; in either case, the *relative, root-mean-square fluctuation* of any quantity is inversely proportional to the square root of the size of the system. Thus, except for situations such as the ones met with in a critical region, normal fluctuations are thermodynamically negligible. This does not mean that fluctuations are altogether irrelevant to the physical phenomena taking place in the system; in fact, as will be seen in the sequel, the very presence of fluctuations at the *microscopic* level is of fundamental importance to several properties of the system displayed at the *macroscopic* level!

With the help of the foregoing results, we may evaluate the mean square fluctuation in the energy of the system. With T and V as independent variables, we have

$$\Delta E = \left(\frac{\partial E}{\partial T}\right)_V \Delta T + \left(\frac{\partial E}{\partial V}\right)_T \Delta V. \tag{17}$$

Squaring this expression and taking averages, keeping in mind eqns (14), we get

$$\begin{aligned}\overline{(\Delta E)^2} &= kT^2 C_V + kT\kappa_T V\left\{\left(\frac{\partial E}{\partial V}\right)_T\right\}^2 \\ &= kT^2 C_V + kT\kappa_T\left(\frac{N^2}{V}\right)\left\{\left(\frac{\partial E}{\partial N}\right)_T\right\}^2.\end{aligned} \tag{18}$$

Now, the results derived in the preceding paragraphs determine the fluctuations of the various physical quantities pertaining to *any macroscopic sub-system* of a given system, provided that the number of particles in the sub-system remains fixed. The expression (14a) for $\overline{(\Delta V)^2}$ may, therefore, be used to derive an expression for the mean square fluctuation of the variable v (the volume per particle) and the variable n (the particle density) of the sub-system. We readily obtain

$$\overline{(\Delta v)^2} = kT\kappa_T V/N^2, \quad \overline{(\Delta n)^2} = \frac{1}{v^4}\overline{(\Delta v)^2} = kT\kappa_T N^2/V^3; \tag{19}$$

note that the [illegible] result obtained here is in complete agreement with eqn. (4.5.7), which was derived on the basis of the grand canonical ensemble. A little reflection shows that this result applies equally well to a sub-system with a fixed volume V and a fluctuating number of particles N. The mean square fluctuation in N is then given by

$$\overline{(\Delta N)^2} = V^2\overline{(\Delta n)^2} = kT\kappa_T N^2/V. \tag{20}$$

Substituting (20) into (18), we obtain once again the grand canonical result for $\overline{(\Delta E)^2}$, namely

$$\overline{(\Delta E)^2} = kT^2 C_V + \overline{(\Delta N)^2}\{(\partial E/\partial N)_T\}^2; \tag{21}$$

cf. eqn. (4.5.14).

In passing, we note that the first part of expression (21) denotes the mean square fluctuation in the energy E of a sub-system for which both N and V are fixed, just as we have in the canonical ensemble (N, V, T). Conversely, if we assume the energy E to be fixed, then the temperature of the sub-system will fluctuate, and the mean square value of the quantity ΔT will be given by $(kT^2 C_V)$ divided by the square of the thermal capacity of the sub-system. The net result will, therefore, be (kT^2/C_V), which is the same as in (14a).

14.2. Spatial correlations in a fluid

It is well known that the particles constituting a homogeneous, isotropic system, such as a liquid or a gas, are equally likely to be at any point $\boldsymbol{r}$ in the space available to them. This statement applies to each individual particle, on the condition that the positions of all other particles in the system are completely arbitrary. If, however, we consider two particles at a time, then, for a *given* position of one particle, different positions of the other particle may no longer be equally likely to obtain. In fact, because of the interparticle interactions and the symmetry properties of the wave functions, different values of the relative position $(\boldsymbol{r}_2 - \boldsymbol{r}_1)$ of any two particles in the system do not occur with equal likelihood. In other words, there exists a *correlation* between the simultaneous positions $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ of the two particles. Interestingly enough, there also exists a simple and straightforward relationship between the space integral of the relevant correlation function on one hand and the mean square fluctuation in the particle density of the system on the other.

To study the precise nature of these correlations, we first introduce a few definitions. To begin with, we have the *configurational distribution function* $F_N(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$ which appears in the nature of a *probability density* and satisfies the normalization condition

$$\int_V F_N(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)\, d^3r_1 \ldots d^3r_N = 1. \tag{1}$$

Integrating $F_N(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$ over the coordinates $\boldsymbol{r}_2, \ldots, \boldsymbol{r}_N$ and multiplying the result by N, we obtain a *single-particle distribution function*

$$F_1(\boldsymbol{r}_1) = N\int_V F_N(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)\, d^3r_2 \ldots d^3r_N = n(\boldsymbol{r}_1). \tag{2}$$

The function $F_1(\boldsymbol{r}_1)$ represents the particle density at the point $\boldsymbol{r}_1$, for $\int_V F_1(\boldsymbol{r}_1)\, d^3r_1 = N$; for a homogeneous system, $F_1(\boldsymbol{r}_1)$ is a constant (which may be denoted by the symbol n). The *two-particle distribution function* is now defined as

$$F_2(\boldsymbol{r}_1, \boldsymbol{r}_2) = N(N-1)\int_V F_N(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)\, d^3r_3 \ldots d^3r_N = n^2 g(\boldsymbol{r}), \tag{3}$$

where $\boldsymbol{r} = (\boldsymbol{r}_2 - \boldsymbol{r}_1)$. Equation (3) defines the *pair distribution function* $g(\boldsymbol{r})$ of the system; clearly, the product $g(\boldsymbol{r})\, d^3r$ determines the probability of finding a particle in the volume element d^3r around the point $\boldsymbol{r}$ when we already have a particle at the point $\boldsymbol{r} = 0$. In the absence of spatial correlations, which only holds for a *classical* gas composed of *non-interacting* particles, the function $g(\boldsymbol{r})$ is identically equal to unity; for real systems, $g(\boldsymbol{r})$ is generally different from unity. It is then natural to introduce a function $\nu(\boldsymbol{r})$, defined by the formula

$$\nu(\boldsymbol{r}) = g(\boldsymbol{r}) - 1, \tag{4}$$

as a measure of the degree of spatial correlation in the system; the function $\nu(\boldsymbol{r})$ is generally referred to as the *pair correlation function*. In the absence of spatial correlations, $\nu(\boldsymbol{r})$ would be identically equal to zero.[1]

We now consider a *macroscopic* region V_A in the fluid and evaluate the mean square fluctuation in the number N_A of the particles occupying this region. To do this, we introduce a function $\mu(\boldsymbol{r})$ such that $\mu(\boldsymbol{r}) = 1$ (or 0) according as the point $\boldsymbol{r}$ lies inside (or outside) the region V_A. The number N_A, which is indeed fluctuating, is then given by the formula

$$N_A = \sum_{i=1}^{N} \mu(\boldsymbol{r}_i), \tag{5}$$

where the summation goes over all the N particles in the system. It follows that

$$\begin{aligned} \overline{N}_A &= \sum_{i=1}^{N} \int_V \mu(\boldsymbol{r}_i) F_N(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)\, d^{3N}r \\ &= N \int_V \mu(\boldsymbol{r}_1)\, d^3r_1 F_N(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)\, d^{3(N-1)}r \\ &= \int_V \mu(\boldsymbol{r}_1) F_1(\boldsymbol{r}_1)\, d^3r_1 = \int_{V_A} F_1(\boldsymbol{r})\, d^3r. \end{aligned} \tag{6}$$

Substituting from (2), we obtain for a homogeneous system

$$\overline{N}_A = nV_A. \tag{7}$$

Next, we write

$$\begin{aligned} N_A^2 &= \sum_{i=1}^{N} \mu(\boldsymbol{r}_i) \sum_{j=1}^{N} \mu(\boldsymbol{r}_j) \\ &= \sum_{i \neq j} \mu(\boldsymbol{r}_i)\mu(\boldsymbol{r}_j) + \sum_{i} \mu(\boldsymbol{r}_i); \end{aligned} \tag{8}$$

in writing the last term, we have made use of the fact that, for all i, $\mu(\boldsymbol{r}_i)\mu(\boldsymbol{r}_i) = \mu(\boldsymbol{r}_i)$. Accordingly,

$$\overline{N_A^2} = \overline{N}_A + \sum_{i \neq j} \int_V \mu(\boldsymbol{r}_i)\mu(\boldsymbol{r}_j) F_N(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)\, d^{3N}r$$

$$
\begin{aligned}
&= \overline{N}_A + N(N-1)\int_V \mu(\boldsymbol{r}_1)\mu(\boldsymbol{r}_2)\, d^3r_1\, d^3r_2 F_N(\boldsymbol{r}_1, \ldots, \mathbf{r}_N)\, d^{3(N-2)}r \\
&= \overline{N}_A + \int_V \mu(\boldsymbol{r}_1)\mu(\boldsymbol{r}_2) F_2(\boldsymbol{r}_1, \boldsymbol{r}_2)\, d^3r_1\, d^3r_2 \\
&= \overline{N}_A + n^2 V_A \int_{V_A} g(\boldsymbol{r})\, d^3r. \qquad (9)
\end{aligned}
$$

Combining (4), (7) and (9), we obtain the desired result

$$
\frac{\overline{N_A^2} - \overline{N}_A^2}{\overline{N}_A} = 1 + n \int_{V_A} \nu(\boldsymbol{r})\, d^3r. \qquad (10)
$$

Equation (10) brings out the intimate relationship that exists between the density fluctuations and the spatial correlations in any fluid system.[2]

Making use of eqn. (14.1.20), as applied to the region V_A of the given system, eqn. (10) yields the fundamental relationship

$$
\int \nu(\boldsymbol{r})\, d^3r = \frac{1}{n}(nkT\kappa_T - 1), \qquad (11)
$$

which may be compared with the corresponding result for a magnetic system, viz. eqn. (11.11.12). For a classical ideal gas, $PV = NkT$; therefore,

$$
\kappa_T \equiv -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T = \frac{1}{P} = \frac{1}{nkT}. \qquad (12)
$$

Accordingly, the space integral of the function $\nu(\boldsymbol{r})$ vanishes. This is in keeping with the fact that, in the absence of interparticle interactions and (quantum-mechanical) symmetry effects, there do not exist any spatial correlations among the particles of a given system.

At this stage, a mention may be made of an important connection that exists between the density fluctuations in a fluid and the scattering of electromagnetic waves by the fluid. It has been known for a long time that the scattering of electromagnetic waves is entirely due to the density fluctuations in a given fluid; in fact, if fluctuations were absent, the waves scattered by various molecules of the fluid would exactly cancel out by interference and no net scattering would result. For light waves, the scattered intensity is directly proportional to expression (10); for x-rays, on the other hand, it is proportional to the more general expression

$$
1 + n \int \nu(\boldsymbol{r}) \frac{\sin(sr)}{sr}\, d^3r. \qquad (13)
$$

Here, $\boldsymbol{s}(= \boldsymbol{k} - \boldsymbol{k}_0)$ is the vector difference between the scattered and the incident wave vectors; hence, $s = (4\pi \sin\theta)/\lambda$, where λ is the wavelength and θ the angle of scattering.[3] Denoting (13) by the symbol $i(s)$, we have

$$
s\{i(s) - 1\} = 4\pi n \int_0^\infty \{g(\boldsymbol{r}) - 1\} \sin(sr) r\, dr, \qquad (14)
$$

whence

$$\{g(r) - 1\} = \frac{1}{2\pi^2 nr} \int_0^\infty s\{i(s) - 1\} \sin(sr)\, ds. \tag{15}$$

Formula (15) enables us to derive the pair distribution function $g(r)$ from the experimental values of the quantity $i(s)$, i.e. from the intensity distribution of the scattered waves over various values of the variable s.

At this point we note that in the region of phase transitions, especially near a critical point, the isothermal compressibility of a fluid and, with it, the level of density fluctuations in the system become abnormally high. Consequently, the intensity of the scattered waves also becomes abnormally large, which gives rise to the spectacular phenomenon of *critical opalescence*.[4] Analysis shows that under these circumstances the correlation function $\nu(r)$, which is ordinarily *short-ranged* and is roughly given by the formula

$$\nu(r) \propto \frac{kT}{r} \exp(-r/\xi), \tag{16}$$

the correlation length ξ being of the order of the mean interparticle distance in the system, becomes *long-ranged* and is roughly given by the formula

$$\nu(r) \propto \frac{kT}{r}. \tag{17}$$

That spatial correlations among the molecules of a fluid should extend over *macroscopic* distances is a feature that typifies the onset of long-range order in the system.

14.3. The Einstein–Smoluchowski theory of the Brownian motion

The term "Brownian motion" derives its name from the botanist Robert Brown who, in 1828, made careful observations on the tiny pollen grains of a plant under a microscope. In his own words: "While examining the form of the particles immersed in water, I observed many of them very evidently in motion. These motions were such as to satisfy me ... that they arose neither from currents in the fluid nor from its gradual evaporation, but belonged to the particle itself." We now know that the real source of this motion lies in the incessant, and more or less random, bombardment of the *Brownian particles*, as these grains (or, for that matter, any colloidal suspensions) are usually referred to, by the *molecules* of the surrounding fluid. It was Einstein who, in a number of papers (beginning 1905), first provided a sound theoretical analysis of the Brownian motion on the basis of the so-called "random walk problem" and thereby established a far-reaching relationship between the irreversible nature of this phenomenon and the mechanism of molecular fluctuations.

To illustrate the essential theme of Einstein's approach, we first consider the problem in *one* dimension. Let $x(t)$ denote the position of the Brownian particle at time t, given that its position coincided with the point $x = 0$ at time $t = 0$. To simplify matter we assume that each molecular impact (which, on an average, takes place af a time τ^*) causes the particle to jump a (small) distance l–of *constant* magnitude–in either a positive or negative direction along the x-axis. It seems natural to regard the possibilities $\Delta x = +l$ and $\Delta x = -l$ as *equally likely*; though somewhat less natural, we may also regard the successive impacts on, and hence the successive jumps of, the Brownian particle to be *mutually uncorrelated.* The probability that the particle is found at the point x at time t is then equal to the probability that, in a series of $n(=t/\tau^*)$ successive jumps, the particle makes $m(=x/l)$ more jumps in the positive direction of the x-axis than in the negative, i.e. it makes $\frac{1}{2}(n+m)$ jumps in the positive direction and $\frac{1}{2}(n-m)$ in the negative.[5] The desired probability is given by the binomial expression

$$p_n(m) = \frac{n!}{\{\frac{1}{2}(n+m)\}!\,\{\frac{1}{2}(n-m)\}!}\left(\frac{1}{2}\right)^n \tag{1}$$

whence it follows that

$$\overline{m} = 0 \quad \text{and} \quad \overline{m^2} = n. \tag{2}$$

Thus, for $t \gg \tau^*$, we have for the net displacement of the particle

$$\overline{x(t)} = 0 \quad \text{and} \quad \overline{x^2(t)} = l^2\frac{t}{\tau^*} \propto t^1. \tag{3}$$

Accordingly, the root-mean-square displacement of the particle is proportional to the square root of the time elapsed:

$$x_{\text{r.m.s.}} = \sqrt{\left(\overline{x^2(t)}\right)} = l\sqrt{(t/\tau^*)} \propto t^{1/2}. \tag{4}$$

It should be noted that the proportionality of the *net* overall displacement of the Brownian particle to the *square root* of the total number of elementary steps is a typical consequence of the random nature of the steps and it manifests itself in a large variety of phenomena in nature. In contrast, if the successive steps were fully coherent (or else if the motion were completely predictable and reversible over the time interval t),[6] then the net displacement would have been proportional to t^1.

Smoluchowski's approach to the problem of Brownian motion, which appeared in 1906, was essentially the same as that of Einstein; the difference lay primarily in the mathematical procedure. Smoluchowski introduced the *probability function* $p_n(x_0|x)$ which denotes the "probability that, after a series of n steps, the Brownian particle, initially at the point x_0, reaches the point x"; the number x here denotes the distance in terms of the length of the elementary step. Clearly,

$$p_n(x_0|x) = \sum_{z=-\infty}^{\infty} p_{n-1}(x_0|z)p_1(z|x) \quad (n \geq 1); \tag{5}$$

moreover, since a single step is equally likely to take the particle to the right or to the left,

$$p_1(z|x) = \tfrac{1}{2}\delta_{z,x-1} + \tfrac{1}{2}\delta_{z,x+1}, \tag{6}$$

while

$$p_0(z|x) = \delta_{z,x}. \tag{7}$$

Equation (5) is known as the *Smoluchowski equation.* To solve it, we introduce a *generating function* $Q_n(\xi)$,

$$Q_n(\xi) = \sum_{x=-\infty}^{\infty} p_n(x_0|x)\xi^{x-x_0}, \tag{8}$$

whence it follows that

$$Q_0(\xi) = \sum_{x=-\infty}^{\infty} p_0(x_0|x)\xi^{x-x_0} = \sum_{x=-\infty}^{\infty} \delta_{x_0,x}\xi^{x-x_0} = 1. \tag{9}$$

Substituting (6) into (5), we obtain

$$p_n(x_0|x) = \tfrac{1}{2}p_{n-1}(x_0|x-1) + \tfrac{1}{2}p_{n-1}(x_0|x+1). \tag{10}$$

Multiplying (10) by ξ^{x-x_0} and adding over all x, we obtain the recurrence relation

$$Q_n(\xi) = \tfrac{1}{2}[\xi + (1/\xi)]Q_{n-1}(\xi), \tag{11}$$

so that, by iteration,

$$Q_n(\xi) = \{\tfrac{1}{2}[\xi + (1/\xi)]\}^n Q_0(\xi) = (1/2)^n[\xi + (1/\xi)]^n. \tag{12}$$

Expanding this expression binomially and comparing the result with (8), we get

$$p_n(x_0|x) = \begin{cases} \left(\dfrac{1}{2}\right)^n \dfrac{n!}{\left\{\frac{1}{2}(n+x-x_0)\right\}!\left\{\frac{1}{2}(n-x+x_0)\right\}!} & \text{for } |x-x_0| \le n \\[2ex] 0 & \text{for } |x-x_0| > n. \end{cases} \tag{13}$$

Identifying $(x - x_0)$ with m, we find this result to be in complete agreement with our previous result (1).[7] Accordingly, any conclusions drawn from the Smoluchowski approach will be the same as the ones drawn from the Einstein approach.

To obtain an asymptotic form of the function $p_n(m)$, we apply Stirling's formula, $n! \approx (2\pi n)^{1/2}(n/e)^n$, to the factorials appearing in (1), with the result

$$\begin{aligned} \ln p_n(m) \approx{}& (n+\tfrac{1}{2})\ln n - \tfrac{1}{2}(n+m+1)\ln\{\tfrac{1}{2}(n+m)\} \\ &- \tfrac{1}{2}(n-m+1)\ln\{\tfrac{1}{2}(n-m)\} - n\ln 2 - \tfrac{1}{2}\ln(2\pi). \end{aligned}$$

For $m \ll n$ (which is generally true because $\overline{m} = 0$ and $m_{\text{r.m.s.}} = n^{1/2}$ while $n \gg 1$), we obtain

$$p_n(m) \approx \frac{2}{\surd(2\pi n)}\exp(-m^2/2n). \tag{14}$$

Taking x to be a continuous variable (and remembering that $p_n(m) \equiv 0$ either for even values of m or for odd values of m, so that in the distribution (14), $\Delta m = 2$ and not 1), we may write this result in the *Gaussian* form:

$$p(x)\,dx = \frac{dx}{\surd(4\pi Dt)}\exp\left(-\frac{x^2}{4Dt}\right), \tag{15}$$

where

$$D = l^2/2\tau^*. \tag{16}$$

Later on we shall see that the quantity D introduced here is identical with the *diffusion coefficient* of the given system; eqn. (16) connects it with the (microscopic) quantities l and τ^*. To appreciate this connection, one has simply to note that the problem of Brownian motion can also be looked upon as a problem of "diffusion" of Brownian particles through the medium of the fluid; this point of view is also due to Einstein. However, before we embark upon these considerations, we would like to present here the results of an actual observation made on the Brownian motion of a spherical particle immersed in water; see Lee, Sears and Turcotte (1963). It was found that the 403 values of the net displacement Δx of the particle, observed after successive intervals of 2 seconds each, were distributed as follows:

Displacement Δx, *in units of* $\mu (= 10^{\ 4}$ *cm)*	*Frequency of occurrence n*
less than -5.5	0
between -5.5 and -4.5	1
between -4.5 and -3.5	2
between -3.5 and -2.5	15
between -2.5 and -1.5	32
between -1.5 and -0.5	95
between -0.5 and $+0.5$	111
between $+0.5$ and $+1.5$	87
between $+1.5$ and $+2.5$	47
between $+2.5$ and $+3.5$	8
between $+3.5$ and $+4.5$	5
greater than $+4.5$	0

The mean square value of the displacement turns out to be: $\overline{(\Delta x)^2} = 2.09 \times 10^{-8}$ cm^2. The observed frequency distribution has been plotted as a "block diagram" in Fig. 14.1. We have included, in the figure, a Gaussian curve based on the observed value of the mean square displacement; we find that the experimental data fit the theoretical curve fairly well. We can also derive here an experimental value for the diffusion coefficient of the medium; we obtain: $D = \overline{(\Delta x)^2}/2t = 5.22 \times 10^{-9}$ cm^2/sec.[8]

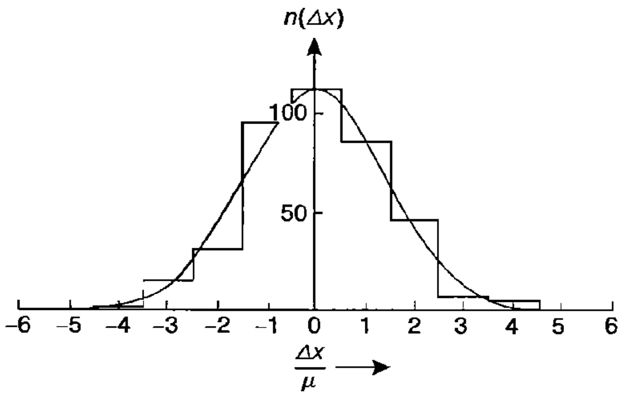


FIG. 14.1. The statistical distribution of the successive displacements, Δx, of a Brownian particle immersed in water: $(\Delta x)_{\text{r.m.s.}} \simeq 1.45\ \mu$.

We now turn to the study of the Brownian motion from the point of view of diffusion. We denote the number density of the Brownian particles in the fluid by the symbol $n(\boldsymbol{r}, t)$ and their current density by $\boldsymbol{j}(\boldsymbol{r}, t)\{= n(\boldsymbol{r}, t)\boldsymbol{v}(\boldsymbol{r}, t)\}$; then, according to Fick's law,

$$\boldsymbol{j}(\boldsymbol{r}, t) = -D\nabla n(\boldsymbol{r}, t), \tag{17}$$

where D stands for the *diffusion coefficient* of the medium. We also have the equation of continuity, viz.

$$\nabla \cdot \boldsymbol{j}(\boldsymbol{r}, t) + \frac{\partial n(\boldsymbol{r}, t)}{\partial t} = 0. \tag{18}$$

Substituting (17) into (18), we obtain the diffusion equation

$$\nabla^2 n(\boldsymbol{r}, t) - \frac{1}{D}\frac{\partial n(\boldsymbol{r}, t)}{\partial t} = 0. \tag{19}$$

Of the various possible solutions of this equation, the one relevant to the present problem is

$$n(r, t) = \frac{N}{(4\pi D t)^{3/2}} \exp\left(-\frac{r^2}{4Dt}\right), \tag{20}$$

which is a spherically symmetric solution and is already normalized:

$$\int_0^\infty n(r, t) 4\pi r^2 \, dr = N, \tag{21}$$

N being the total number of (Brownian) particles immersed in the fluid. A comparison of the (three-dimensional) result (20) with the (one-dimensional) result (15) brings out most vividly the relationship between the random walk problem on one hand and the phenomenon of diffusion on the other.

It is clear that in the last approach we have considered the motion of an "ensemble" of N Brownian particles placed under "equivalent" physical conditions, rather than considering the motion of a single particle over a length of time (as was done in the random walk approach). Accordingly, the averages of the various physical quantities obtained here will be in the nature of "ensemble averages"; they must, of course, agree with the long-time averages of the same quantities obtained earlier. Now, by virtue of the distribution (20), we obtain

$$\langle \boldsymbol{r}(t) \rangle = 0; \quad \langle r^2(t) \rangle = \frac{1}{N}\int_0^\infty n(r, t) 4\pi r^4 \, dr = 6Dt \propto t^1, \tag{22}$$

in complete agreement with our earlier results, namely

$$\overline{x(t)} = 0; \quad \overline{x^2(t)} = l^2 t/\tau^* = 2Dt \propto t^1. \tag{23}$$

Thus, the "ensemble" of the Brownian particles, initially concentrated at the origin, "diffuses out" as time increases, the nature and the extent of its spread at any time t being given by eqns (20) and (22), respectively. The diffusion process, which is clearly *irreversible,* gives us a fairly good picture of the statistical behavior of a single particle in the ensemble. However, the important thing to bear in mind is that, whether we focus our attention on a single particle in the ensemble or look at the ensemble as a whole, the ultimate source of the phenomenon lies in the incessant, and more or less random, impacts received by the Brownian particles from the molecules of the fluid. In other words, the irreversible character of the phenomenon ultimately arises from the random, fluctuating forces exerted by the fluid molecules on the Brownian particles. This leads to another systematic theory of the Brownian motion, viz. the theory of Langevin (1908). For a detailed analysis of the problem, see Uhlenbeck and Ornstein (1930), Chandrasekhar (1943, 1949), MacDonald (1948–49) and Wax (1954).

14.4. The Langevin theory of the Brownian motion

We consider the simplest case of a "free" Brownian particle, surrounded by a fluid environment; the particle is assumed to be free in the sense that it is not acted upon by any other force except the one arising from the molecular bombardment. The equation of motion of the particle will be

$$M\frac{dv}{dt} = \mathcal{F}(t), \tag{1}$$

where M is the particle mass, $v(t)$ the particle velocity and $\mathcal{F}(t)$ the force acting upon the particle by virtue of the impacts from the fluid molecules. Langevin suggested that the force $\mathcal{F}(t)$ may be written as a sum of two parts: (i) an "averaged-out" part, which represents the *viscous drag*, $-v/B$, experienced by the particle (accordingly, B is the *mobility* of the system, i.e. the drift velocity acquired by the particle by virtue of a unit "external" force)[9] and (ii) a "rapidly fluctuating" part $F(t)$ which, over long intervals of time (as compared to the characteristic time τ^*), averages out to zero; thus, we may write

$$M\frac{dv}{dt} = -\frac{v}{B} + F(t); \quad \overline{F(t)} = 0. \tag{2}$$

Taking the ensemble average of (2), we obtain[10]

$$M\frac{d}{dt}\langle v \rangle = -\frac{1}{B}\langle v \rangle, \tag{3}$$

whence

$$\langle v(t) \rangle = v(0)\exp(-t/\tau) \quad (\tau = MB). \tag{4}$$

Thus, the mean drift velocity of the particle decays, at a rate determined by the *relaxation time* τ, to the ultimate value zero. We note that this result is typical of the phenomena governed by *dissipative* properties such as the viscosity of the fluid; the *irreversible* nature of the result is also evident.

Dividing (2) by the mass of the particle, we obtain an equation for the *instantaneous* acceleration, viz.

$$\frac{dv}{dt} = -\frac{v}{\tau} + A(t); \quad \overline{A(t)} = 0. \tag{5}$$

We construct the scalar product of (5) with the *instantaneous* position $\boldsymbol{r}$ of the particle and take the ensemble average of the product. In doing so, we make use of the facts that (i) $\boldsymbol{r} \cdot \boldsymbol{v} = \frac{1}{2}(dr^2/dt)$, (ii) $\boldsymbol{r} \cdot (d\boldsymbol{v}/dt) = \frac{1}{2}(d^2r^2/dt^2) - v^2$, and (iii) $\langle \boldsymbol{r} \cdot \boldsymbol{A} \rangle = 0$. We obtain

$$\frac{d^2}{dt^2}\langle r^2 \rangle + \frac{1}{\tau}\frac{d}{dt}\langle r^2 \rangle = 2\langle v^2 \rangle. \tag{6}$$

If the Brownian particle has already attained thermal equilibrium with the molecules of the fluid, then the quantity $\langle v^2 \rangle$ in this equation may be replaced by its *equipartition value* $3kT/M$. The equation is then readily integrated, with the result

$$\langle r^2 \rangle = \frac{6kT\tau^2}{M}\left\{\frac{t}{\tau} - (1 - e^{-t/\tau})\right\}, \tag{7}$$

where the constants of integration have been so chosen that at $t = 0$ both $\langle r^2 \rangle$ and its first time-derivative vanish. We observe that, for $t \ll \tau$,

$$\langle r^2 \rangle \simeq \frac{3kT}{M}t^2 = \langle v^2 \rangle t^2, \tag*{(8)^{12}}$$

which is consistent with the reversible equations of motion whereby one should simply have

$$\boldsymbol{r} = \boldsymbol{v}t. \tag{9}$$

On the other hand, for $t \gg \tau$,

$$\langle r^2 \rangle \simeq \frac{6kT\tau}{M}t = (6BkT)t, \tag*{(10)^{13}}$$

which is essentially the same as the Einstein–Smoluchowski result (14.3.22); incidentally, we obtain here a simple, but important, relationship between the coefficient of diffusion D and the mobility B, viz.

$$D = BkT, \tag{11}$$

which is generally referred to as the *Einstein relation*. The irreversible character of eqn. (10) is self-evident; it is also clear that it arises essentially from the viscosity of the medium. Moreover, the Einstein relation (11), which connects the coefficient of diffusion D with the mobility B of the system, tells us that the ultimate source of the viscosity of the medium (as well as of diffusion) lies in the random, fluctuating forces arising from the incessant motion of the fluid molecules; see also the fluctuation–dissipation theorem of Sec. 14.7.

In this context, if we consider a particle of charge e and mass M moving in a viscous fluid under the influence of an external electric field of intensity E, then the "coarse-grained" motion of the particle will be determined by the equation

$$M\frac{d}{dt}\langle \boldsymbol{v} \rangle = -\frac{1}{B}\langle \boldsymbol{v} \rangle + e\boldsymbol{E}; \tag{12}$$

cf. eqn. (3). The "terminal" drift velocity of the particle would now be given by the expression $(eB)\boldsymbol{E}$, which prompts one to define (eB) as the "mobility" of the system and denote it by the symbol μ. Consequently, one obtains, instead of (11),

$$D = \frac{kT}{e}\mu, \tag{13}$$

which is, in fact, the original version of the Einstein relation; sometimes this is also referred to as the *Nernst relation*.

So far we have not felt any *direct* influence of the rapidly fluctuating term $A(t)$ that appears in the equation of motion (5) of the Brownian particle. For this, let us try to evaluate the quantity $\langle v^2(t)\rangle$ which, in the preceding analysis, was assumed to have already attained its "limiting" value $3kT/M$. For this evaluation we replace the variable t in eqn. (5) by u, multiply both sides of the equation by exp (u/τ), rearrange and integrate over du between the limits $u = 0$ and $u = t$; we thus obtain the formal solution

$$v(t) = v(0)e^{-t/\tau} + e^{-t/\tau}\int_0^t e^{u/\tau}A(u)\,du. \tag{14}$$

Thus, the drift velocity $v(t)$ of the particle is also a fluctuating function of time; of course, since $\langle A(u)\rangle = 0$ for all u, the *average* drift velocity is given by the first term alone, viz.

$$\langle v(t)\rangle = v(0)e^{-t/\tau}, \tag{15}$$

which is the same as our earlier result (4). For the mean square velocity $\langle v^2(t)\rangle$, we now obtain from (14)

$$\langle v^2(t)\rangle = v^2(0)e^{-2t/\tau} + 2e^{-2t/\tau}\left[v(0)\cdot\int_0^t e^{u/\tau}\langle A(u)\rangle\,du\right]$$
$$+ e^{-2t/\tau}\int_0^t\int_0^t e^{(u_1+u_2)/\tau}\langle A(u_1)\cdot A(u_2)\rangle\,du_1\,du_2. \tag{16}$$

The second term on the right-hand side of this equation is identically zero, because $\langle A(u)\rangle$ vanishes for all u. In the third term, we have the quantity $\langle A(u_1)\cdot A(u_2)\rangle$, which is a measure of the "statistical correlation between the value of the fluctuating variable A at time u_1 and its value at time u_2"; we call it the *autocorrelation function* of the variable A and denote it by the symbol $K_A(u_1. u_2)$ or simply by $K(u_1, u_2)$. Before proceeding with (16) any further, we place on record some of the important properties of the function $K(u_1, u_2)$.

(i) In a stationary ensemble (i.e. one in which the overall macroscopic behavior of the systems does not change with time), the function $K(u_1, u_2)$ depends only on the time interval $(u_2 - u_1)$. Denoting this interval by the symbol s, we have

$$K(u_1, u_1 + s) \equiv \langle A(u_1)\cdot A(u_1 + s)\rangle = K(s), \text{ independently of } u_1. \tag{17}$$

(ii) The quantity $K(0)$, which is identically equal to the mean square value of the variable A at time u_1, must be *positive definite*. In a stationary ensemble, it would be a constant, independent of u_1:

$$K(0) = \text{const.} > 0. \tag{18}$$

(iii) For any value of s, the magnitude of the function $K(s)$ cannot exceed $K(0)$.

Proof. Since

$$\langle |A(u_1) \pm A(u_2)|^2\rangle = \langle A^2(u_1)\rangle + \langle A^2(u_2)\rangle \pm 2\langle A(u_1)\cdot A(u_2)\rangle$$
$$= 2\{K(0) \pm K(s)\} \geq 0,$$

the function $K(s)$ cannot go outside the limits $-K(0)$ and $+K(0)$; consequently,

$$|K(s)| \le K(0) \quad \text{for all } s. \tag{19}$$

(iv) The function $K(s)$ is symmetric about the value $s = 0$, i.e.

$$K(-s) = K(s) = K(|s|). \tag{20}$$

Proof.

$$\begin{aligned} K(s) &\equiv \langle A(u_1) \cdot A(u_1 + s)\rangle = \langle A(u_1 - s) \cdot A(u_1)\rangle^{14} \\ &= \langle A(u_1) \cdot A(u_1 - s)\rangle \equiv K(-s). \end{aligned}$$

(v) As s becomes large in comparison with the characteristic time τ^*, the values $A(u_1)$ and $A(u_1 + s)$ become *uncorrelated*, that is

$$K(s) \equiv \langle A(u_1) \cdot A(u_1 + s)\rangle \xrightarrow[s \gg \tau^*]{} \langle A(u_1)\rangle \cdot \langle A(u_1 + s)\rangle = 0. \tag{21}$$

In other words, the "memory" of the molecular impacts received during a given interval of time, say between u_1 and $u_1 + du_1$, is "completely lost" after a lapse of time large in comparison with τ^*. It follows that the magnitude of the function $K(s)$ is significant only so long as the variable s is of the same order of magnitude as τ^*.

Figures 14.7–14.9 show the s-dependence of certain typical correlation functions $K(s)$; they fully conform to the properties listed above.

We now evaluate the double integral appearing in (16):

$$I = \int_0^t \int_0^t e^{(u_1+u_2)/\tau} K(u_2 - u_1)\, du_1\, du_2. \tag{22}$$

Changing over to the variables

$$S = \tfrac{1}{2}(u_1 + u_2) \quad \text{and} \quad s = (u_2 - u_1), \tag{23}$$

the integrand becomes $\exp(2S/\tau)K(s)$, the element $(du_1\, du_2)$ gets replaced by the corresponding element $(dS\, ds)$ while the limits of integration, in terms of the variables S and s, can be read from Fig. 14.2; we find that, for $0 \le S \le t/2$, s goes from $-2S$ to $+2S$ while, for $t/2 \le S \le t$, it goes from $-2(t - S)$ to $+2(t - S)$. Accordingly,

$$I = \int_0^{t/2} e^{2S/\tau}\, dS \int_{-2S}^{+2S} K(s)\, dS + \int_{t/2}^{t} e^{2S/\tau}\, dS \int_{-2(t-S)}^{+2(t-S)} K(s)\, ds. \tag{24}$$

In view of property (v) of the function $K(s)$, see eqn. (21), the integrals over s draw significant contribution only from a very narrow region, of the order of τ^*, around the value $s = 0$, i.e. from the shaded region in Fig. 14.2; contributions from regions with larger values of $|s|$ are negligible. Thus, if $t \gg \tau^*$, the limits of integration for s may be replaced by $-\infty$ and $+\infty$, with the result

$$I \simeq C \int_0^t e^{2s/\tau}\, dS = C\frac{\tau}{2}(e^{2t/\tau} - 1), \tag{25}$$

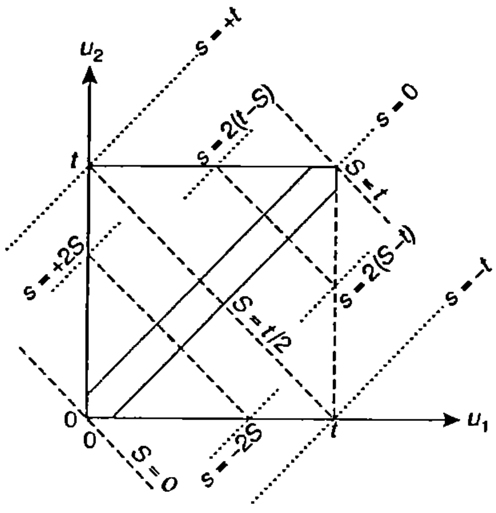


FIG. 14.2. Limits of integration, of the double integral I, in terms of the variables S and s.

where

$$C = \int_{-\infty}^{\infty} K(s)\,ds. \tag{26}$$

Substituting (25) into (16), we obtain

$$\langle v^2(t)\rangle = v^2(0)e^{-2t/\tau} + C\frac{\tau}{2}(1 - e^{-2t/\tau}). \tag{27}$$

Now, as $t \to \infty$, $\langle v^2(t)\rangle$ must tend to the equipartition value $3kT/M$; therefore,

$$C = 6kT/M\tau \tag{28}$$

and hence

$$\langle v^2(t)\rangle = v^2(0) + \left\{\frac{3kT}{M} - v^2(0)\right\}(1 - e^{-2t/\tau}). \qquad (29)^{15}$$

We note that if $v^2(0)$ were itself equal to the equipartition value $3kT/M$, then $\langle v^2(t)\rangle$ would always remain the same, which shows that statistical equilibrium, once attained, has a natural tendency to persist.

Substituting (29) into the right-hand side of (6), we obtain a more representative description of the manner in which the quantity $\langle r^2\rangle$ varies with t; we thus have

$$\frac{d^2}{dt^2}\langle r^2\rangle + \frac{1}{\tau}\frac{d}{dt}\langle r^2\rangle = 2v^2(0)e^{-2t/\tau} + \frac{6kT}{M}(1 - e^{-2t/\tau}), \tag{30}$$

with the solution

$$\langle r^2\rangle = v^2(0)\tau^2(1 - e^{-t/\tau})^2 - \frac{3kT}{M}\tau^2(1 - e^{-t/\tau})(3 - e^{-t/\tau}) + \frac{6kT\tau}{M}t. \tag{31}$$

Solution (31) satisfies the initial conditions that both $\langle r^2\rangle$ and its first time-derivative vanish at $t = 0$; moreover, if we put $v^2(0) = 3kT/M$, it reduces to solution (7) obtained earlier. Once again, we note the *reversible* nature of the motion for $t \ll \tau$, with $\langle r^2 \rangle \simeq v^2(0)t^2$, and its *irreversible* nature for $t \gg \tau$, with $\langle r^2 \rangle \simeq (6BkT)t$.

Figures 14.3 and 14.4 show the variation, with time, of the ensemble averages $\langle v^2(t) \rangle$ and $\langle r^2(t) \rangle$ of a Brownian particle, as given by eqns (29) and (31), respectively. All important features of our results are manifestly evident in these plots.

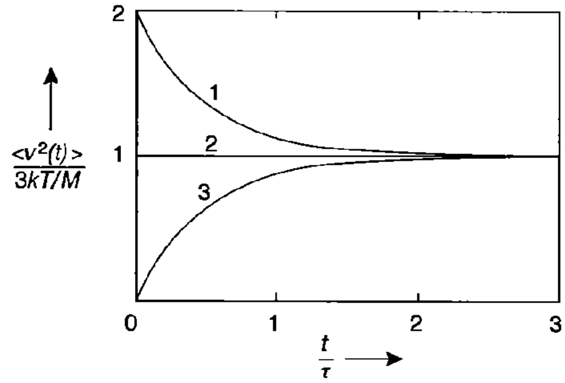


FIG. 14.3. The mean square velocity of a Brownian particle as a function of time. Curves 1, 2 and 3 correspond, respectively, to the initial conditions $v^2(0) = 6kT/M$, $3kT/M$ and 0.

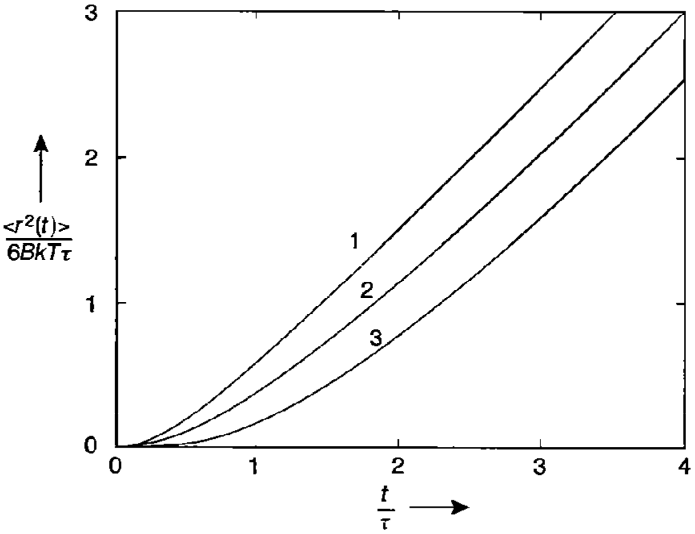


FIG. 14.4. The mean square displacement of a Brownian particle as a function of time. Curves 1, 2 and 3 correspond, respectively, to the initial conditions $v^2(0) = 6kT/M$, $3kT/M$ and 0.

14.5. Approach to equilibrium: the Fokker–Planck equation

In our analysis of the Brownian motion we have considered the behavior of a dynamical variable, such as the position $r(t)$ or the velocity $v(t)$ of a Brownian particle, from the point of view of fluctuations in the value of the variable.

To determine the average behavior of such a variable, we sometimes invoked an "ensemble" of Brownian particles immersed in identical environments and undergoing diffusion. A treatment along these lines was carried out towards the end of Sec. 14.3, and the most important results of that treatment are summarized in eqn. (14.3.20) for the density function $n(\boldsymbol{r}, t)$ and in eqn. (14.3.22) for the mean square displacement $\langle r^2(t)\rangle$.

A more generalized way of looking at "the manner in which, and the rate at which, a given distribution of Brownian particles approaches a state of thermal equilibrium" is provided by the so-called *Master Equation*, a simplified version of which is known as the Fokker–Planck equation. For illustration, we examine the displacement, $x(t)$, of the given set of particles along the x-axis. At any time t, let $f(x,t)\,dx$ be the probability that an arbitrary particle in the ensemble may have a displacement between x and $x + dx$. The function $f(x,t)$ must satisfy the normalization condition

$$\int_{-\infty}^{\infty} f(x,t)\,dx = 1. \tag{1}$$

The Master Equation then reads:

$$\frac{\partial f(x,t)}{\partial t} = \int_{-\infty}^{\infty} \{-f(x,t)W(x,x') + f(x',t)W(x',x)\}\,dx', \tag{2}$$

where $W(x,x')\,dx'\delta t$ denotes the probability that in a short interval of time δt a particle having displacement x makes a "transition" to having a displacement between x' and $x' + dx'$.[16] The first part of the integral corresponds to all those transitions that remove particles from the displacement x at time t to some other displacement x' and, hence, represent a net *loss* to the function $f(x,t)$; similarly, the second part of the integral corresponds to all those transitions that bring particles from some other displacement x' at time t to the displacement x and, hence, represent a net *gain* to the function $f(x,t)$.[17] The structure of the Master Equation is thus founded on very simple and straightforward premises. Of course, under certain conditions, this equation, or any generalization thereof (such as the one including velocity, or momentum, coordinates in the argument of f), can be reduced to the simple form

$$\frac{\partial f}{\partial t} = -\frac{f - f_0}{\tau}, \tag{3}$$

which has proved to be a very useful first approximation for studying problems related to *transport phenomena*. Here, f_0 denotes the *equilibrium distribution function* (for $\partial f/\partial t = 0$ when $f = f_0$), while τ is the *relaxation time* that determines the rate at which the fluctuations in the system drive it to a state of equilibrium.

In studying Brownian motion on the basis of eqn. (2), we can safely assume that it is only transitions between "closely neighboring" states x and x' that have an appreciable probability of occurring; in other words, the transition probability function $W(x,x')$ is sharply peaked around the value $x' = x$ and falls rapidly to zero away from x. Denoting the interval $(x' - x)$ by ξ, we may write

$$W(x,x') \to W(x;\ \xi),\ W(x',x) \to W(x';\ -\xi) \tag{4}$$

where $W(x;\,\xi)$ and $W(x';\,-\xi)$ have sharp peaks around the value $\xi = 0$ and fall rapidly to zero elsewhere.[18] This enables us to expand the right-hand side of (2) as a Taylor series around $\xi = 0$. Retaining terms up to second order only, we obtain

$$\frac{\partial f(x,t)}{\partial t} = -\frac{\partial}{\partial x}\{\mu_1(x)f(x,t)\} + \frac{1}{2}\frac{\partial^2}{\partial x^2}\{\mu_2(x)f(x,t)\}, \tag{5}$$

where

$$\mu_1(x) = \int_{-\infty}^{\infty} \xi W(x;\,\xi)\,d\xi = \frac{\langle \delta x\rangle_{\delta t}}{\delta t} = \langle v_x\rangle \tag{6}$$

and

$$\mu_2(x) = \int_{-\infty}^{\infty} \xi^2 W(x;\,\xi)\,d\xi = \frac{\langle(\delta x)^2\rangle_{\delta t}}{\delta t}. \tag{7}$$

Equation (5) is the so-called Fokker–Planck equation which occupies a classic place in the field of Brownian motion and fluctuations.

We now consider a specific system of Brownian particles (of negligible mass), each particle being acted upon by a linear restoring force, $F_x = -\lambda x$, and having mobility B in the surrounding medium; the assumption of negligible mass implies that the relaxation time $\tau(= MB)$ of eqn. (14.4.4) is very small, so the time t here may be regarded as very large in comparison with that τ. The mean viscous force, $-\langle v_x\rangle/B$, is then balanced by the linear restoring force, with the result that

$$-\frac{\langle v_x\rangle}{B} + F_x = 0 \tag{8}$$

and hence

$$\langle v_x\rangle \equiv \mu_1(x) = -\lambda Bx. \tag{9}$$

Next, in view of eqn. (14.4.10), we have

$$\frac{\langle(\delta x)^2\rangle}{\delta t} \equiv \mu_2(x) = 2BkT; \tag{10}$$

it will be noted that the influence of λ on this quantity is being neglected here. Substituting (9) and (10) into (5), we obtain

$$\frac{\partial f}{\partial t} = \lambda B\frac{\partial}{\partial x}(xf) + BkT\frac{\partial^2 f}{\partial x^2}. \tag{11}$$

We now apply eqn. (11) to an "ensemble" of particles, initially concentrated at the point $x = x_0$. To begin with, we note that, in the absence of the restoring force ($\lambda = 0$), eqn. (11) reduces to the one-dimensional *diffusion equation*

$$\frac{\partial f}{\partial t} = D\frac{\partial^2 f}{\partial x^2} \quad (D = BkT), \tag{12}$$

which conforms to our earlier results (14.3.19) and (14.4.11). The present derivation shows that the process of diffusion is essentially a "random walk, at the molecular level". In view of eqn. (14.3.20), the function $f(x,t)$ here would be

$$f(x,t) = \frac{1}{(4\pi Dt)^{1/2}}\exp\left\{-\frac{(x-x_0)^2}{4Dt}\right\}, \tag{13}$$

with

$$\bar{x} = x_0 \quad \text{and} \quad \overline{x^2} = x_0^2 + 2Dt, \tag{14}$$

the last result shows that the mean square distance traversed by the particle(s) increases *linearly* with time, without any upper limit on its value. The restoring force, on the other hand, puts a check on the diffusive tendency of the particles. For instance, in the presence of such a force ($\lambda \neq 0$), the terminal distribution f_∞ (for which $\partial f/\partial t = 0$) is determined by the equation

$$\frac{\partial}{\partial x}(xf_\infty) + \frac{kT}{\lambda}\frac{\partial^2 f_\infty}{\partial x^2} = 0, \tag{15}$$

whence

$$f_\infty(x) = \left(\frac{\lambda}{2\pi kT}\right)^{1/2} \exp\left(-\frac{\lambda x^2}{2kT}\right), \tag{16}$$

with

$$\bar{x} = 0 \quad \text{and} \quad \overline{x^2} = kT/\lambda. \tag{17}$$

The last result agrees with the fact that the mean square value of x must ultimately comply with the *equipartition theorem*, viz. $\overline{(\tfrac{1}{2}\lambda x^2)}_\infty = \tfrac{1}{2}kT$. From the point of view of equilibrium statistical mechanics, if we regard Brownian particles, with kinetic energy $p_x^2/2m$ and potential energy $\tfrac{1}{2}\lambda x^2$, as *loosely coupled* to a thermal environment at temperature T, then we may directly write

$$f_{eq}(x, p_x)\,dx\,dp_x \propto e^{-(p_x^2/2m + \lambda x^2/2)/kT}\,dx\,dp_x. \tag{18}$$

On integration over p_x, expression (18) leads directly to the distribution function (16).

The general solution of eqn. (11), relevant to the ensemble under consideration, is given by

$$f(x,t) = \left\{\frac{\lambda}{2\pi kT(1 - e^{-2\lambda Bt})}\right\}^{1/2} \exp\left\{-\frac{\lambda(x - x_0 e^{-\lambda Bt})^2}{2kT(1 - e^{-2\lambda Bt})}\right\}, \tag{19}$$

with

$$\bar{x} = x_0 e^{-\lambda Bt} \quad \text{and} \quad \overline{x^2} = x_0^2 e^{-2\lambda Bt} + \frac{kT}{\lambda}(1 - e^{-2\lambda Bt}); \tag{20}$$

in the limit $\lambda \to 0$, we recover the purely "diffusive" situation, as described by eqns (13) and (14), while for $t \gg (\lambda B)^{-1}$, we approach the "terminal" situation, as described by eqns (16) and (17). Figure 14.5 shows the manner in which an ensemble of Brownian particles approaches a state of equilibrium under the combined influence of the restoring force and the molecular bombardment; clearly, the relaxation time of the present process is $\sim (\lambda B)^{-1}$.

A physical system to which the foregoing theory is readily applicable is provided by the oscillating component of a moving-coil galvanometer. Here, we have a coil of wire and a mirror which are suspended by a fine fiber, so they can rotate about a vertical axis. Random, incessant collisions of air molecules with the suspended system produce a succession of torques of *fluctuating* intensity; as a result, the angular position θ of the system continually fluctuates and the system exhibits an

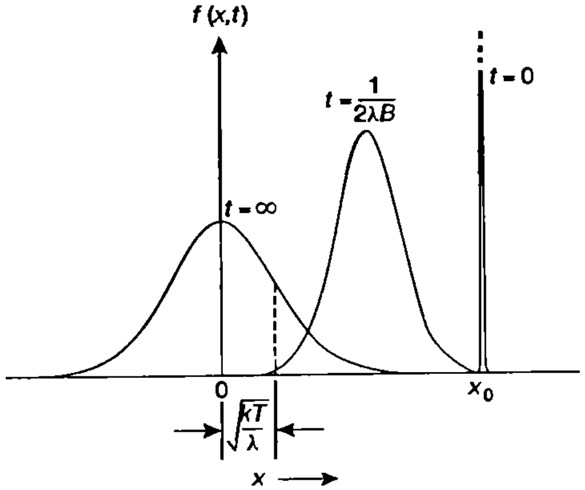


FIG. 14.5. The distribution function (19) at times $t = 0$, $t = 1/(2\lambda B)$ and $t = \infty$.

unsteady zero. This is clearly another example of the Brownian motion! The role of the viscous force in this case is played by the mechanism of air damping (or, else, electromagnetic damping) of the galvanometer, while the restoring torque, $N_\theta = -c\theta$, arises from the torsional properties of the fiber. In equilibrium, we expect that

$$\overline{\left(\frac{1}{2}c\theta^2\right)} = \frac{1}{2}kT, \quad \text{that is} \quad \overline{\theta^2} = \frac{kT}{c}; \tag{21}$$

cf. eqn. (17). An experimental determination of the mean square deflection, $\overline{\theta^2}$, of such a system was made by Kappler (1931) who, in turn, applied his results to derive, with the help of eqn. (21), an empirical value for the Boltzmann constant k (or, for that matter, the Avogadro number N_A). The system used by Kappler had a moment of inertia $I = 4.552 \times 10^{-4}$ g cm^2 and a time period of oscillation $\tau = 1379$ sec; accordingly, the constant c of the restoring torque had a value given by the formula $\tau = 2\pi(I/c)^{1/2}$, so that

$$c = 4\pi^2(I/\tau^2) = 9.443 \times 10^{-9} \text{ g cm}^2 \text{ sec}^{-2}/\text{rad}.$$

The observed value of $\overline{\theta^2}$, at a temperature of 287.1 K, was 4.178×10^{-6}. Substituting these numbers in (21), Kappler obtained: $k = 1.374 \times 10^{-16}$ erg K^{-1}. And, since the gas constant R is equal to 8.31×10^7 erg K^{-1} mole^{-1}, he obtained for the Avogadro number: $N_A = R/k = 6.06 \times 10^{23}$ mole^{-1}.

One might expect that by suspending the mirror system in an "evacuated" casing the fluctuations caused by the collisions of the air molecules could be severely reduced. This is not true because even at the lowest possible pressures there still remain a tremendously large number of molecules in the system which keep the Brownian motion "alive". The interesting part of the story, however, is that the mean square deflection of the system, caused by molecular bombardment, is not at all affected by the density of the molecules; for a given system *in equilibrium*, it is determined solely by the temperature. This situation is depicted, rather dramatically, in Fig. 14.6 where we have two traces of oscillations of the mirror system, the upper one having been taken at the atmospheric pressure and the lower one at a pressure of 10^{-4} mm of mercury. The root-mean-square deviation is very nearly the same in the two cases! Nevertheless, one does note a difference of "quality" between the two traces which relates to the "frequency spectrum" of the fluctuations and arises for the following reason. When the density of the surrounding gas is relatively high, the molecular impulses come in rapid succession, with the result that the *individual* deflections of the system are large in number but small in magnitude. As the pressure is lowered, the time intervals between successive impulses become longer, making the *individual* deflections smaller in number but larger in magnitude. However, the overall deflection, observed over a long interval of time, remains essentially the same.

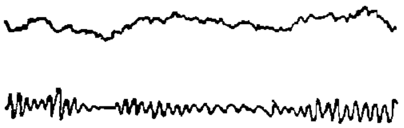


FIG. 14.6. Two traces of the thermal oscillations of a mirror system suspended in air; the upper trace was taken at the atmospheric pressure, the lower one at a pressure of 10^{-4} mm of mercury.

14.6. Spectral analysis of fluctuations: the Wiener–Khintchine theorem

We have already made reference to the (spectral) quality of a fluctuation pattern. Referring once again to the patterns shown in Fig. 14.6, we note that, even though the mean square fluctuation of the variable θ is the same in the two cases, the second pattern is far more "jagged" than the first: in other words, the high frequency components are far more prominent in the second pattern. At the same time, there is a lot more "predictability" in the first pattern (insofar as it is represented by a much smoother curve); in other words, the *correlation function*, or the *memory function*, $K(s)$ of the first pattern extends over much larger values of s. In fact, these two aspects of a fluctuation process, viz. its time-dependence on one hand and its frequency spectrum on the other, are very closely related to one another. And the most natural course for studying this relationship is to carry out a Fourier analysis of the given process.

For this study we consider only those variables, $y(t)$, whose *mean square value*, $\langle y^2(t)\rangle$, has already attained an equilibrium, or stationary, value:

$$\langle y^2(t)\rangle = \text{const.} \tag{1}$$

Such a variable is said to be *statistically stationary*. As an example of such a variable, we may recall the velocity $v(t)$ of a "free" Brownian particle at times t much larger than the relaxation time τ, see eqn. (14.4.29), or the displacement $x(t)$ of a Brownian particle moving under the influence of a restoring force ($F_x = -\lambda x$) at times t much larger than $(\lambda B)^{-1}$, see eqn. (14.5.20). Now, if the variable $y(t)$ were *strictly periodic* (and hence *completely predictable*), with a time period $T = 1/f_0$, then we could write

$$y(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos(2\pi n f_0 t) + \sum_{n=1}^{\infty} b_n \sin(2\pi n f_0 t), \tag{2}$$

where

$$a_0 = \frac{1}{T}\int_0^T y(t)\,dt, \tag{3}$$

$$a_n = \frac{2}{T}\int_0^T y(t)\cos(2\pi n f_0 t)\,dt \tag{4}$$

and

$$b_n = \frac{2}{T}\int_0^T y(t)\sin(2\pi n f_0 t)\,dt; \tag{5}$$

in this case, the coefficients a and b would be completely known and would define, with no uncertainty, the frequency spectrum of the variable $y(t)$. If, on the other hand, the given variable is *more* or *less* a random function of time, then the coefficients a and b would themselves be statistical in nature. To apply the concept of periodicity to such a function, we must take the "time interval of repetition" to be infinitely large, i.e. we let $f_0 \to 0$.

In the proposed limit, eqn. (3) would read

$$a_0 = \lim_{T\to\infty}\frac{1}{T}\int_0^T y(t)\,dt \equiv \langle y(t)\rangle; \tag{6}$$

thus, the coefficient a_0, which represents the mean (or d.c.) value of the variable y, may be determined *either* by taking a time average (over a sufficiently long interval) of the variable or by taking an ensemble average (at any instant of time t). For convenience, and without loss of generality, we take $a_0 = 0$; in other words, we assume that from the actual values of the variable $y(t)$ its mean value, $\langle y(t)\rangle$, has already been subtracted.[19] Taking the ensemble average of eqns (4) and (5), we obtain, for all n,

$$\langle a_n\rangle = \langle b_n\rangle = 0. \tag{7}$$

However, by taking the ensemble average of eqn. (2) *squared*, we obtain

$$\begin{aligned}\langle y^2(t)\rangle &= \sum_n \tfrac{1}{2}\langle a_n^2\rangle + \sum_n \tfrac{1}{2}\langle b_n^2\rangle \\ &= \sum_n \tfrac{1}{2}\left\{\langle a_n^2\rangle + \langle b_n^2\rangle\right\} = \text{const.}\end{aligned} \tag{8}$$

The term $\frac{1}{2}\left\{\langle a_n^2\rangle + \langle b_n^2\rangle\right\}$ represents the respective "share", belonging to the frequency nf_0, in the total, time-independent value of the quantity $\langle y^2(t)\rangle$. Now, in view of the randomness of the phases of the various components, we have, for all n, $\langle a_n^2\rangle = \langle b_n^2\rangle$; consequently, eqn. (8) may be written as

$$\langle y^2\rangle = \sum_n \langle a_n^2\rangle \simeq \int_0^\infty w(f)\,df, \tag{9}$$

where

$$\langle a_n^2\rangle = w(nf_0)\Delta(nf_0), \quad \text{that is} \quad w(nf_0) = \frac{1}{f_0}\langle a_n^2\rangle; \tag{10}$$

the function $w(f)$ defines the *power spectrum* of the variable $y(t)$.

We shall now show that the power spectrum $w(f)$ of the fluctuating variable $y(t)$ is completely determined by its auto-correlation function $K(s)$. For this, we make use of eqn. (4) which gives

$$\langle a_n^2\rangle = 4f_0^2 \int_0^{1/f_0}\int_0^{1/f_0} \langle y(t_1)y(t_2)\rangle \cos(2\pi n f_0 t_1)\cos(2\pi n f_0 t_2)\,dt_1\,dt_2. \tag{11}$$

Changing over to the variables

$$S = \tfrac{1}{2}(t_1 + t_2) \quad \text{and} \quad s = (t_2 - t_1),$$

and remembering that the interval T over which the integrations extend is much larger than the duration over which the "memory" of the variable y lasts, we obtain

$$\langle a_n^2\rangle \simeq 2f_0^2 \int_{S=0}^{1/f_0}\int_{s=-\infty}^{\infty} K(s)\{\cos(2\pi n f_0 s) + \cos(4\pi n f_0 S)\}\,dS\,ds; \tag{12}$$

cf. the steps leading from eqn. (14.4.22) to (14.4.25,26). The second part of the integral vanishes on integration over S; the first part gives

$$\langle a_n^2\rangle = 4f_0 \int_0^{\infty} K(s)\cos(2\pi n f_0 s)\,ds. \tag{13}$$

Comparing (13) with (10), we obtain the desired formula

$$w(f) = 4\int_0^{\infty} K(s)\cos(2\pi f s)\,ds. \tag{14}$$

Taking the inverse of (14), we obtain

$$K(s) = \int_0^{\infty} w(f)\cos(2\pi f s)\,df. \tag{15}$$

For $s = 0$, formula (15) yields the important relationship

$$K(0) = \int_0^{\infty} w(f)\,df = \langle y^2\rangle; \tag{16}$$

see eqn. (9) as well as the definition of the autocorrelation function of the variable y, viz. $K(s) = \langle y(t_1)y(t_1+s)\rangle$. Equations (14) and (15), which connect the complementary functions $w(f)$ and $K(s)$, constitute a theorem that goes after the names of Wiener (1930) and Khintchine (1934).

We shall now look at some special cases of the variable $y(t)$ to illustrate the use of the *Wiener–Khintchine theorem*.

(i) If the given variable $y(t)$ is extremely irregular, and hence unpredictable, then its correlation function $K(s)$ would extend over a negligibly small range of the time interval s.[20] We may then write

$$K(s) = c\delta(s). \tag{17a}$$

Equation (14) t' gives:

$$w(f) = 2c \quad \text{for all } f. \tag{17b}$$

A spectrum in which the distribution (of power) over different frequencies is uniform is known as a "flat" or a "white" spectrum. We note, however, that if the uniformity of distribution were literally true for all frequencies, from 0 to ∞, then the integral in (16), which is identically equal to $\langle y^2 \rangle$, would diverge! We therefore expect that, in any realistic situation, the correlation function $K(s)$ will not be as sharply peaked as in (17a). Typically, $K(s)$ will extend over a small range, $O(\sigma)$, of the variable s, which in turn will define a "frequency zone", with $f = O(1/\sigma)$, such that the function $w(f)$ would undergo a change of character as f passes through this zone; towards lower frequencies $w(f) \to$ const. $\neq 0$, while towards higher frequencies $w(f) \to$ const. $= 0$. One possible representation of this situation is shown in Fig. 14.7 where we have taken, rather arbitrarily,

$$K(s) = K(0)\frac{\sin(as)}{as} \quad (a > 0), \tag{18a}$$

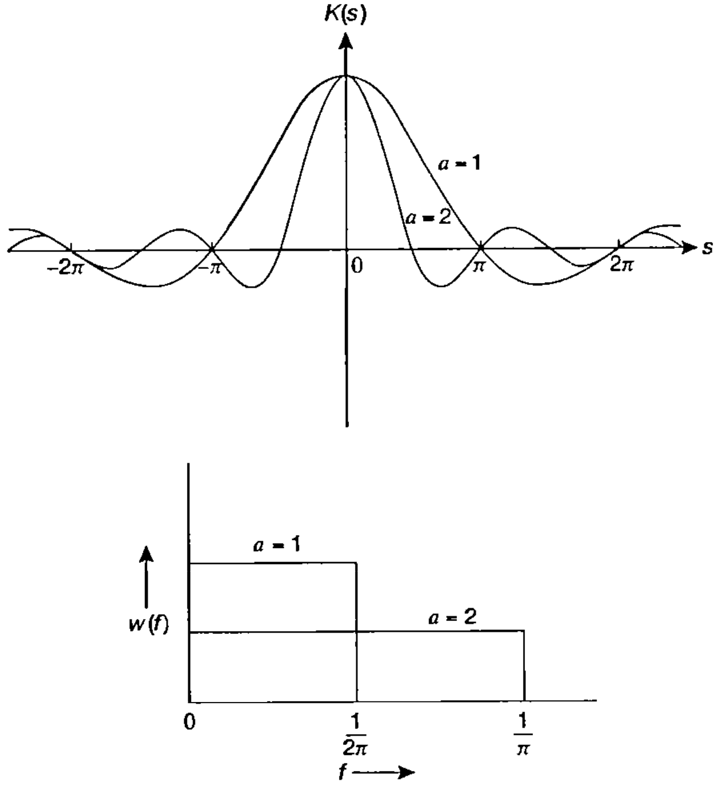


FIG. 14.7. The autocorrelation function $K(s)$ and the power distribution function $w(f)$ of a given variable $y(t)$; the parameter a appears in terms of an arbitrary unit of (time)$^{-1}$.

for which

$$w(f) = \begin{cases} \dfrac{2\pi}{a}K(0) & \text{for } f < \dfrac{a}{2\pi} \\ 0 & \text{for } f > \dfrac{a}{2\pi}. \end{cases} \tag{18b}$$

In the limit $a \to \infty$, eqns (18) reduce to (17), with $c = \pi a^{-1} K(0)$.

(ii) On the other hand, if the variable $y(t)$ is extremely regular, and hence predictable, then its correlation function would extend over large values of s; its power spectrum would then appear in the form of "peaks", located at certain "characteristic frequencies" of the variable. In the simplest case of a *monochromatic* variable, with characteristic frequency f^*, the correlation function would be

$$K(s) = K(0)\cos(2\pi f^* s), \tag{19a}$$

for which

$$w(f) = K(0)\delta(f - f^*); \tag{19b}$$

see Fig. 14.8. A very special case arises when $f^* = 0$; then, both $y(t)$ and $K(s)$ are constant in value, and the function $w(f)$ is peaked at the d.c. frequency $f = 0$.

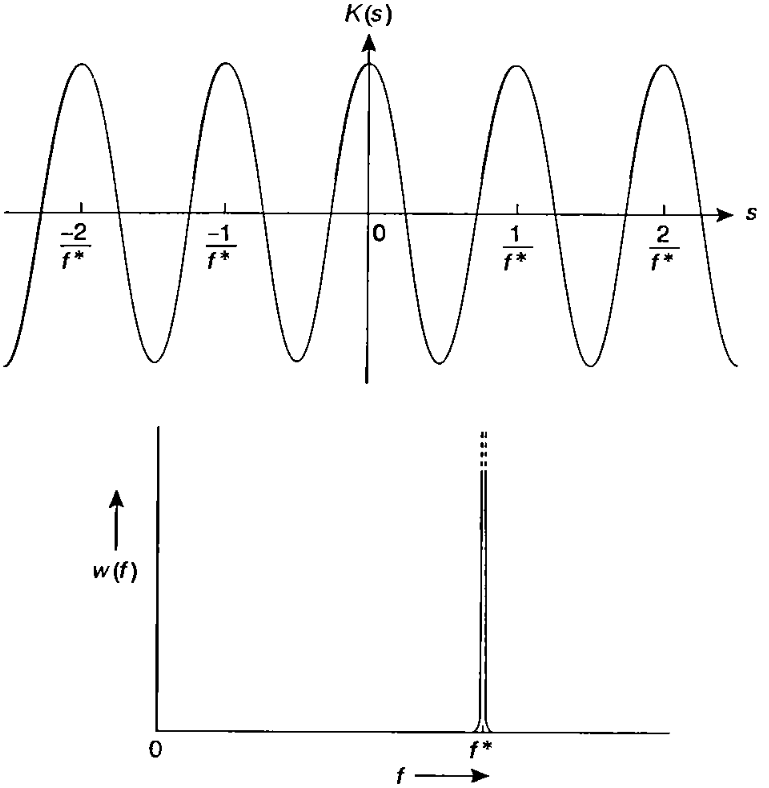


FIG. 14.8. The autocorrelation function $K(s)$ and the power distribution function $w(f)$ of a monochromatic variable $y(t)$, with characteristic frequency f^*.

(iii) If the variable $y(t)$ represents a signal that arises from, or has been filtered through, a lightly-damped tuned circuit (a "narrow-band" filter), then its power will be distributed over a "hump" around the mean frequency f^*. The function $K(s)$ will then appear in the nature of an "attenuated" function whose time scale, σ, is determined by the width, Δf, of the hump in the power spectrum. A situation of this kind is shown in Fig. 14.9.

The relevance of spectral analysis to the problem of the actual observation of a fluctuating variable is best brought out by examining power spectrum of the velocity $v(t)$ of a Brownian particle. Considering the x-component alone, the

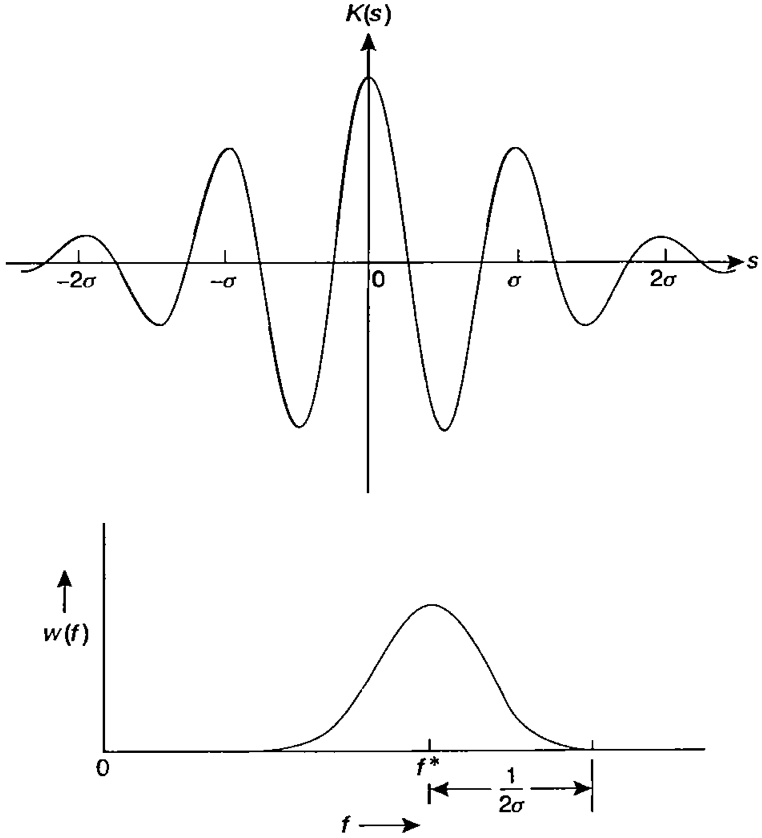


FIG. 14.9. The autocorrelation function $K(s)$ and the power distribution function $w(f)$ of a variable that has been filtered through a lightly-damped tuned circuit, with mean frequency f^* and width $\Delta f \sim (1/\sigma)$.

autocorrelation function $K_{v_x}(s)$, or simply $K(s)$, is given by

$$K(s) = \frac{kT}{M} e^{-|s|/\tau} \quad (\tau = MB); \tag{20}$$

see eqn. (14.7.10). The power spectrum $w(f)$ is then given by the expression

$$w(f) = \frac{4kT}{M} \int\limits_0^\infty e^{-s/\tau} \cos(2\pi f s)\, ds = \frac{4kT\tau}{M} \frac{1}{1 + (2\pi f \tau)^2}, \tag{21}$$

which indeed satisfies the relationship

$$\begin{aligned} \int\limits_0^\infty w(f)\, df &= \frac{2kT}{\pi M} \tan^{-1}(2\pi f \tau) \Big|_0^\infty \\ &= \frac{kT}{M} = \langle v_x^2 \rangle, \end{aligned} \tag{22}$$

in agreement with the *equipartition theorem* (as applied to a single component of the velocity v). For $f \ll \tau^{-1}$, the power distribution is practically independent of f, which implies a practically "white" spectrum, with

$$w(f) \simeq \frac{4kT\tau}{M} = 4BkT. \tag{23}$$

We can then write for the velocity fluctuations in the frequency range $(f, f+\Delta f)$, with $f \ll \tau^{-1}$,

$$\langle \Delta v_x^2 \rangle_{(f, f+\Delta f)} \simeq w(f)\Delta f \simeq (4BkT)\Delta f. \tag{24}$$

In general, our measuring instrument (or the eye, in the case of a visual examination of the particle) has a *finite* response time τ_0, as a consequence of which it is unable to respond to frequencies larger than, say, τ_0^{-1}. The *observed* fluctuation is then given by the "pruned" expression

$$\langle v_x^2 \rangle_{\text{obs}} \simeq \int_0^{1/\tau_0} w(f)\,df = \frac{2kT}{\pi M} \tan^{-1}\left(2\pi \frac{\tau}{\tau_0}\right), \tag{25}$$

instead of the "full" expression (22). In a typical case, the mass of the Brownian particle $M \sim 10^{-12}$ g, its diameter $2a \sim 10^{-4}$ cm and the coefficient of viscosity of the fluid $\eta \sim 10^{-2}$ poise, so that the relaxation time $\tau = M/(6\pi\eta a) \sim 10^{-7}$ sec. However, the response time τ_0, in the case of visual observation, is of the order of 10^{-1} sec; clearly, $\tau/\tau_0 \sim 10^{-6} \ll 1$. Equation (25) then reduces to[21]

$$\langle v_x^2 \rangle_{\text{obs}} \simeq \frac{4kT\tau}{M\tau_0} \ll \frac{kT}{M}; \tag{26}$$

thus, in view of the *finiteness* of the response time τ_0, the observed root-mean-square velocity of the Brownian particle will be down by a factor of $2(\tau/\tau_0)^{1/2} \sim 10^{-3}$; numerically, this takes us down from a root-mean-square value which, at room temperatures, is $\sim 10^{-1}$ cm/s to a value $\sim 10^{-4}$ cm/s. It is gratifying to note that the outcome of actual observations of Brownian particles is in complete agreement with the latter result; for a more detailed analysis of this question, see MacDonald (1950). The foregoing discussion highlights the fact that, in the process of observing a fluctuating variable, our measuring instrument picks up signals over only a *limited* range of frequencies (as determined by the response time of the instrument); signals belonging to higher frequencies are simply left out.

The theory of this section can be readily applied to fluctuations in the motion of electrons in an (L, R) circuit. Corresponding to eqns (21)–(24), we have for fluctuations in the electric current I

$$w(f) = \frac{4kT\tau'}{L} \frac{1}{1 + (2\pi f \tau')^2} \quad \left(\tau' = \frac{L}{R}\right), \tag{27}$$

so that

$$\int_0^\infty w(f)\,df = \frac{kT}{L} = \langle I^2 \rangle, \tag{28}$$

in agreement with the *equipartition* theorem: $\langle \frac{1}{2}LI^2 \rangle = \frac{1}{2}kT$. For $f \ll 1/\tau'$, eqn. (27) reduces to

$$w(f) \simeq \frac{4kT}{R}, \tag{29}$$

which, once again, implies "white" noise; accordingly, for low frequencies,

$$\langle \Delta I^2 \rangle_{(f, f+\Delta f)} \simeq w(f)\Delta f \simeq \frac{4kT}{R}\Delta f. \tag{30}$$

Equivalently obtain for fluctuations in the voltage

$$\langle \Delta V^2 \rangle_{(f,f+\Delta f)} \simeq (4RkT)\Delta f. \tag{31}$$

Equation (31) constitutes the so-called *Nyquist theorem*, which was first discovered empirically by Johnson (1927–8) and was later derived by Nyquist (1927–8) on the basis of an argument involving the second law of thermodynamics and the exchange of energy between two resistances in thermal equilibrium [22]

14.7. The fluctuation–dissipation theorem

In Sec. 14.4 we obtained a result of considerable importance, namely

$$\frac{1}{B} \equiv \frac{M}{\tau} = \frac{M^2}{6kT}C = \frac{M^2}{6kT}\int_{-\infty}^{\infty} K_A(s)\,ds$$

$$= \frac{1}{6kT}\int_{-\infty}^{\infty} K_F(s)\,ds; \tag{1}$$

see eqns (14.4.4, 26 and 28). Here, $K_A(s)$ and $K_F(s)$ are, respectively, the autocorrelation functions of the fluctuating acceleration $A(t)$ and the fluctuating force $F(t)$ experienced by the Brownian particle:

$$K_A(s) = \langle A(0)\cdot A(s)\rangle = \frac{1}{M^2}\langle F(0)\cdot F(s)\rangle = \frac{1}{M^2}K_F(s). \tag{2}$$[23]

Equation (1) establishes a fundamental relationship between the coefficient, $1/B$, of the "averaged-out" part of the total force $\mathcal{F}(t)$ experienced by the Brownian particle due to the impacts of the fluid molecules and the statistical character of the "fluctuating" part, $F(t)$, of that force; see Langevin's equation (14.4.2). In other words, it relates the coefficient of viscosity of the fluid, which represents *dissipative* forces operating in the system, with the temporal character of the molecular *fluctuations*; the content of eqn. (1) is, therefore, referred to as a *fluctuation–dissipation theorem*. The most striking feature of this theorem is that it relates, in a fundamental manner, the fluctuations of a physical quantity pertaining to the *equilibrium state* of the given system to a dissipative process which, in practice, is realized only when the system is subject to an external force that drives it *away from equilibrium*. Consequently, it enables us to determine the *non-equilibrium properties* of a given system on the basis of a knowledge of the thermal fluctuations occurring in the system when the system is in one of its *equilibrium states*! For an expository account of the fluctuation–dissipation theorem, the reader may refer to Kubo (1966).

At this stage we recall that in eqn. (14.4.11) we obtained a relationship between the *diffusion coefficient* D and the *mobility* B, viz. $D = BkT$. Combining this with (1), we get

$$\frac{1}{D} = \frac{1}{6(kT)^2}\int_{-\infty}^{\infty} K_F(s)\,ds. \tag{3}$$

Now, the diffusion coefficient D can be related directly to the autocorrelation function $K_v(s)$ of the fluctuating variable $\boldsymbol{v}(t)$. For this, one starts with the observation that, by definition,

$$\boldsymbol{r}(t) = \int_0^t \boldsymbol{v}(u)\, du, \tag{4}$$

whence

$$\langle r^2(t)\rangle = \int_0^t \int_0^t \langle \boldsymbol{v}(u_1) \cdot \boldsymbol{v}(u_2)\rangle\, du_1\, du_2. \tag{5}$$

Proceeding in the same manner as for the integral in eqn. (14.4.22), one obtains

$$\langle r^2(t)\rangle = \int_0^{t/2} dS \int_{-2S}^{+2S} K_v(s)\, ds + \int_{t/2}^{t} dS \int_{-2(t-S)}^{+2(t-S)} K_v(s)\, ds; \tag{6}$$

cf. eqn. (14.4.24).

The function $K_v(s)$ can be determined by making use of expression (14.4.14) for $\boldsymbol{v}(t)$ and following exactly the same procedure as was followed for determining the quantity $\langle v^2(t)\rangle$, which is nothing but the maximal value, $K_v(0)$, of the desired function. One obtains

$$K_v(s) = \begin{cases} v^2(0)e^{-(2t+s)/\tau} + \dfrac{3kT}{M} e^{-s/\tau}(1 - e^{-2t/\tau}) & \text{for} \quad s > 0 \qquad (7) \\[2ex] v^2(0)e^{-(2t+s)/\tau} + \dfrac{3kT}{M} e^{s/\tau}(1 - e^{-2(t+s)/\tau}) & \text{for} \quad s < 0; \qquad (8) \end{cases}$$

cf. eqn. (14.4.27). It is easily seen that formulae (7) and (8) can be combined into a single formula, viz.

$$K_v(s) = v^2(0)e^{-|s|/\tau} + \left\{\frac{3kT}{M} - v^2(0)\right\} (e^{-|s|/\tau} - e^{-(2t+s)/\tau}) \quad \text{for all } s; \tag{9}$$

cf. eqn. (14.4.29). In the case of a "stationary ensemble",

$$K_v(s) = \frac{3kT}{M} e^{-|s|/\tau}, \tag{10}$$

which is consistent with property (14.4.20). It should be noted that the time scale for the correlation function $K_v(s)$ is provided by the *relaxation time* τ of the (Brownian) motion, which is many orders of magnitude larger than the *characteristic time* τ^* that provides the time scale for correlation functions $K_A(s)$ and $K_F(s)$.

It is now instructive to verify that the substitution of expression (10) into (6) leads to formula (14.4.7) for $\langle r^2\rangle$, while the substitution of the more general expression (9) leads to formula (14.4.31); see Problem 14.20. In either case,

$$\langle r^2\rangle \xrightarrow[t \gg \tau]{} 6Dt. \tag{11}$$

In the same limit, eqn. (6) reduces to

$$\langle r^2\rangle \simeq \int_0^t dS \int_{-\infty}^{\infty} K_v(s)\, ds = t \int_{-\infty}^{\infty} K_v(s)\, ds. \tag{12}$$

Comparing the two results, we obtain the desired relationship:

$$D = \tfrac{1}{6} \int_{-\infty}^{\infty} K_v(s)\, ds. \tag{13}$$

In passing, we note, from eqns (3) and (13), that

$$\int_{-\infty}^{\infty} K_v(s)\, ds \int_{-\infty}^{\infty} K_F(s)\, ds = (6kT)^2; \tag{14}$$

see also Problem 14.10.

It is not surprising that the equations describing a fluctuation–dissipation theorem can be adapted to any situation that involves a dissipative mechanism. For instance, fluctuations in the motion of electrons in an electric resistor give rise to a "spontaneous" thermal e.m.f., which may be denoted as $\mathcal{E}(t)$. In the spirit of the Langevin theory, this e.m.f. may be split into two parts: (i) an "averaged-out" part, $-RI(t)$, which represents the resistive (or dissipative) aspect of the situation, and (ii) a "rapidly fluctuating" part, $V(t)$, which, over long intervals of time, averages out to zero. The "spontaneous" current in the resistor is then given by the equation

$$L\frac{dI}{dt} = -RI + V(t); \quad \langle V(t) \rangle = 0. \tag{15}$$

Comparing this with the *Langevin equation* (14.4.2) and pushing the analogy further, we infer that there exists a direct relationship between the resistance R and the temporal character of the fluctuations in the variable $V(t)$. In view of eqns (1) and (13), this relationship would be

$$R = \frac{1}{6kT} \int_{-\infty}^{\infty} \langle V(0) \cdot V(s) \rangle\, ds \tag{16}$$

or, equivalently,

$$\frac{1}{R} = \frac{1}{6kT} \int_{-\infty}^{\infty} \langle I(0) \cdot I(s) \rangle\, ds. \tag{17}$$

A generalization of the foregoing result has been given by Kubo (1957, 1959); see, for instance, Kubo (1965), problem 6.19, or Wannier (1966), sec. 23.2. On generalization, the *electric current density* $\boldsymbol{j}(t)$ is given by the expression

$$j_i(t) = \sum_l \int_{-\infty}^{t} E_l(t')\Phi_{li}(t - t')\, dt' \quad (i, l = x, y, z); \tag{18}$$

here, $\boldsymbol{E}(t)$ denotes the applied electric field while

$$\Phi_{li}(s) = \frac{1}{kT} \langle j_l(0) j_i(s) \rangle. \tag{19}$$

Clearly, the quantities $kT\Phi_{li}(s)$ are the components of the *autocorrelation tensor* of the fluctuating vector $\boldsymbol{j}(t)$. In particular, if we consider the static case $\boldsymbol{E} = (E, 0, 0)$, we obtain for the *conductivity* of the system

$$\sigma_{xx} \equiv \frac{j_x}{E} = \int_{-\infty}^{t} \Phi_{xx}(t-t')\,dt' = \int_{0}^{\infty} \Phi_{xx}(s)\,ds$$

$$= \frac{1}{2kT} \int_{-\infty}^{\infty} \langle j_x(0) j_x(s)\rangle\, ds, \tag{20}$$

which may be compared with eqn. (17). If, on the other hand, we take $\boldsymbol{E} = (E\cos\omega t, 0, 0)$, we obtain instead

$$\sigma_{xx}(\omega) = \frac{1}{2kT} \int_{-\infty}^{\infty} \langle j_x(0) j_x(s)\rangle e^{-i\omega s}\, ds. \tag{21}$$

Taking the inverse of (21), we get

$$\langle j_x(0) j_x(s)\rangle = \frac{kT}{\pi} \int_{-\infty}^{\infty} \sigma_{xx}(\omega) e^{i\omega s}\, d\omega. \tag{22}$$

If we now assume that $\sigma_{xx}(\omega)$ does not depend on ω (and may, therefore, be denoted by the simpler symbol σ), then

$$\langle j_x(0) j_x(s)\rangle = (2kT\sigma)\delta(s); \tag{23}$$

cf. note 23. A reference to eqns (14.6.17) shows that, in the present approximation, thermal fluctuations in the electric current are charaterized by a "white" noise.

14.8. The Onsager relations

Most physical phenomena exhibit a kind of symmetry, sometimes referred to as *reciprocity*, which arises from certain basic properties of the microscopic processes that operate behind the (observable) macroscopic situations. A notable example of this is met with in the *thermodynamics of irreversible processes* where one deals with a variety of flow processes, such as heat flow, electric current, mass transfer, etc. These flows (or "currents") are driven by "forces", such as a temperature difference, a potential difference, a pressure difference, etc., which come into play because of a natural tendency among physical systems, which happen to be out of equilibrium, to approach a state of equilibrium. If the given state (of the system) is not too far removed from a state of equilibrium, then one might assume a *linear* relationship between the forces X_i and the currents $\dot{x}_i$:

$$\dot{x}_i = \gamma_{ij} X_j, \tag{1}$$

where the γ_{ij} are referred to as the *kinetic coefficients* of the system.[24] Simple examples of such coefficients are thermal conductivity, electrical conductivity, diffusion coefficient, etc. There are, however, *nondiagonal elements*, $\gamma_{ij} (i \neq j)$, as well which may or may not vanish; they are responsible for the so-called *cross* *effects*. It is the symmetry properties of the matrix (γ_{ij}) that form the subject matter of this section.

The most obvious way to approach this problem is to consider the entropy, $S(x_i)$, of the system in the *disturbed state* in relation to its maximal value, $S(\tilde{x}_i)$, in the relevant *state of equilibrium*. It is the natural tendency of the function $S(x_i)$ to approach its maximal value $S(\tilde{x}_i)$ that brings into play the driving forces X_i; these forces give rise to currents $\dot{x}_i$, which take the "coordinates" x_i towards their equilibrium values $\tilde{x}_i$. If the deviations $(x_i - \tilde{x}_i)$ are small, then the function $S(x_i)$ may be expressed as a Taylor series around the values $x_i = \tilde{x}_i$; retaining terms up to the second order only, we have

$$S(x_i) = S(\tilde{x}_i) + \left(\frac{\partial S}{\partial x_i}\right)_{x_i=\tilde{x}_i} (x_i - \tilde{x}_i) + \frac{1}{2}\left(\frac{\partial^2 S}{\partial x_i \partial x_j}\right)_{x_{i,j}=\tilde{x}_{i,j}} (x_i - \tilde{x}_i)(x_j - \tilde{x}_j). \tag{2}$$

In view of the fact that the function $S(x_i)$ is *maximum* at $x_i = \tilde{x}_i$, its first derivatives vanish; we may, therefore, write

$$\Delta S \equiv S(x_i) - S(\tilde{x}_i) = -\tfrac{1}{2}\beta_{ij}(x_i - \tilde{x}_i)(x_j - \tilde{x}_j), \tag{3}$$

where

$$\beta_{ij} = -\left(\frac{\partial^2 S}{\partial x_i \partial x_j}\right)_{x_{i,j}=\tilde{x}_{i,j}} = \beta_{ji}. \tag{4}$$

The driving forces X_i may be defined in the spirit of the *second law of thermodynamics*, i.e.

$$X_i = \left(\frac{\partial S}{\partial x_i}\right) = -\beta_{ij}(x_j - \tilde{x}_j). \tag{5}$$

We note that in the present approximation the forces X_i depend linearly on the displacements $(x_i - \tilde{x}_i)$; in the state of equilibrium, they vanish. Now, in view of eqn. (14.1.2), the *ensemble average* of the quantity $x_i X_j$ is given by

$$\langle x_i X_j \rangle = \frac{\int\limits_{-\infty}^{\infty} (x_i X_j) \exp\left\{-\frac{1}{2k}\beta_{ij}(x_i - \tilde{x}_i)(x_j - \tilde{x}_j)\right\} \prod_i dx_i}{\int\limits_{-\infty}^{\infty} \exp\left\{-\frac{1}{2k}\beta_{ij}(x_i - \tilde{x}_j)(x_j - \tilde{x}_j)\right\} \prod_i dx_i}; \tag{6}$$

the limits of integration in (6) have been extended to $-\infty$ and $+\infty$ because the integrals here do not draw any significant contribution from large values of the variables involved. In the same way,

$$\langle x_i \rangle = \frac{\int\limits_{-\infty}^{\infty} x_i \exp\left\{-\frac{1}{2k}\beta_{ij}(x_i - \tilde{x}_i)(x_j - \tilde{x}_j)\right\} \prod_i dx_i}{\int\limits_{-\infty}^{\infty} \exp\left\{-\frac{1}{2k}\beta_{ij}(x_i - \tilde{x}_i)(x_j - \tilde{x}_j)\right\} \prod_i dx_i} = \tilde{x}_i. \tag{7}$$

Differentiating (7) with respect to $\tilde{x}_j$ (and remembering that the integral in the denominator is a constant, independent of the actual values of the quantities $\tilde{x}_i$), and comparing the resulting expression with (6), we obtain the remarkable result

$$\langle x_i X_j \rangle = -k\delta_{ij}. \tag{8}$$

We now proceed towards the key point of the argument. First of all, we note that, though eqns (1) are concerned with *irreversible* phenomena, the microscopic processes underlying these phenomena obey *time reversal*, which means that the temporal correlations of the relevant variables are the same whether measured forward or backward. Thus

$$\langle x_i(0)x_j(s)\rangle = \langle x_i(0)x_j(-s)\rangle; \tag{9}$$

also, by a shift in the zero of time,

$$\langle x_i(0)x_j(-s)\rangle = \langle x_i(s)x_j(0)\rangle. \tag{10}$$

Combining (9) and (10), we obtain

$$\langle x_i(0)x_j(s)\rangle = \langle x_i(s)x_j(0)\rangle. \tag{11}$$

If we now subtract, from both sides of this equation, the quantity $\langle x_i(0)x_j(0)\rangle$, divide the resulting equation by s and let $s \to 0$, we obtain

$$\langle x_i(0)\dot{x}_j(0)\rangle = \langle \dot{x}_i(0)x_j(0)\rangle. \tag{12}$$

Substituting from (1) and making use of (8), we obtain on the left-hand side

$$\langle x_i(0)\gamma_{jl}X_l(0)\rangle = -k\gamma_{jl}\delta_{il} = -k\gamma_{ji}$$

and on the right-hand side

$$\langle \gamma_{il}X_l(0)x_j(0)\rangle = -k\gamma_{il}\delta_{jl} = -k\gamma_{ij}.$$

It follows that

$$\gamma_{ij} = \gamma_{ji}. \tag{13}$$

Equations (13) constitute the *Onsager reciprocity relations*; they were first derived by Onsager in 1931 and have become an essential basis for the thermodynamics of irreversible phenomena.

In view of eqns (1) and (13), the currents $\dot{x}_i$ may be written as

$$\dot{x}_i = \frac{\partial f}{\partial X_i}, \tag{14}$$

where the *generating function* f is a quadratic function of the forces X_i:

$$f = \tfrac{1}{2}\gamma_{ij}X_iX_j. \tag{15}$$

The function f is especially important in that it determines directly the rate at which the entropy of the system changes with time:

$$\dot{S} = \frac{\partial S}{\partial x_i}\dot{x}_i = X_i\dot{x}_i = X_i\frac{\partial f}{\partial X_i} = 2f. \tag{16}$$

As the system approaches the state of equilibrium, its entropy must *increase* towards the equilibrium value $S(\tilde{x}_i)$. The function f must, therefore, be *positive definite*, which places certain restrictions on the coefficients γ_{ij}.

Analogous to eqn. (1), we could also write

$$\dot{X}_i = \zeta_{ij}(x_j - \tilde{x}_j), \tag{17}$$

the quantities ζ_{ij} being another set of coefficients pertaining to the system. From eqns (1) and (5), on the other hand, we obtain

$$\begin{aligned} \dot{X}_i &= -\beta_{ij}\dot{x}_j = -\beta_{ij}(\gamma_{jl}X_l) = -\beta_{ij}\gamma_{jl}\{-\beta_{lm}(x_m - \tilde{x}_m)\} \\ &= \beta_{ij}\gamma_{jl}\beta_{lm}(x_m - \tilde{x}_m). \end{aligned} \tag{18}$$

Comparing (17) and (18), we obtain a relationship between the new coefficients ζ_{ij} and the (kinetic) coefficients γ_{ij}:

$$\zeta_{im} = \beta_{ij}\gamma_{jl}\beta_{lm}. \tag{19}$$

Further, in view of the symmetry properties of the matrices β and γ, we get

$$\zeta_{im} = \zeta_{mi}; \tag{20}$$

thus, the coefficients ζ_{ij}, introduced through the phenomenological eqns (17), also obey the reciprocity relations. It then follows that the quantities $\dot{X}_i$ may be written as, cf. eqn. (14),

$$\dot{X}_i = \frac{\partial f'}{\partial x_i}, \tag{21}$$

where

$$f' = \tfrac{1}{2}\zeta_{ij}(x_i - \tilde{x}_i)(x_j - \tilde{x}_j). \tag{22}$$

The entropy change dS may now be written as

$$\begin{aligned} dS &= \frac{\partial S}{\partial x_j}\,dx_j = X_j\,dx_j = -\beta_{ji}(x_i - \tilde{x}_i)\,dx_j \\ &= (x_i - \tilde{x}_i)\,d\{-\beta_{ij}(x_j - \tilde{x}_j)\} = (x_i - \tilde{x}_i)\,dX_i, \end{aligned} \tag{23}$$

so that

$$\frac{\partial S}{\partial X_i} = (x_i - \tilde{x}_i); \tag{24}$$

clearly, the entropy S is now regarded as an *explicit* function of the forces X_i (rather than of the coordinates x_i). The time derivative of S now takes the form

$$\dot{S} = \frac{\partial S}{\partial X_i}\dot{X}_i = (x_i - \tilde{x}_i)\frac{\partial f'}{\partial x_i} = 2f' \tag{25}$$

Comparing (16) and (25), we conclude that the functions f and f' are, in fact, the same; they are only expressed in terms of two different sets of variables.

It seems important to mention here that Onsager's reciprocity relations have an intimate connection with the fluctuation–dissipation theorem of the preceding section. Following eqns (14.7.18) and (14.7.19), and adopting the summation convention, we have in the present context

$$\dot{x}_i(t) = \frac{1}{kT} \int_{-\infty}^{t} E_l(t') \langle \dot{x}_l(t') \dot{x}_i(t) \rangle \, dt' \tag{26}$$

or, setting $(t - t') = s$,

$$\dot{x}_i(t) = \frac{1}{kT} \int_{0}^{\infty} E_l(t - s) \langle \dot{x}_l(t - s) \langle \dot{x}_i(t) \rangle \, ds; \tag{27}$$

cf. eqn. (1). Interchanging the indices i and l, we obtain

$$\dot{x}_l(t) = \frac{1}{kT} \int_{0}^{\infty} E_i(t - s) \langle \dot{x}_i(t - s) \dot{x}_l(t) \rangle \, ds. \tag{28}$$

The crucial point now is that the correlation functions appearing in eqns (27) and (28) are identical in value:

$$\langle \dot{x}_l(t - s) \dot{x}_i(t) \rangle = \langle \dot{x}_l(0) \dot{x}_i(s) \rangle = \langle \dot{x}_l(0) \dot{x}_i(-s) \rangle = \langle \dot{x}_l(t) \dot{x}_i(t - s) \rangle; \tag{29}$$

in establishing (29), the first and third steps followed from "a shift in time" while the second step followed from the "principle of *dynamical reversibility* of microscopic processes". The equivalence depicted in eqn. (29) is, in essence, the content of Onsager's reciprocity relations. In particular, if the correlation functions appearing in (27) and (28) are sharply peaked at the value $s = 0$, then these equations reduce to the phenomenological eqns (1), and eqn. (29) becomes synonymous with the Onsager relations (13).

In the end, we make some further remarks concerning relations (13). We recall that, in arriving at these relations, we made an appeal to the invariance of the microscopic processes under time reversal. The situation is somewhat different in the case of a "system in rotation" (or a "system in an external magnetic field"), for then the invariance under time reversal holds only if there is also a simultaneous change of sign of the angular velocity Ω (or of the magnetic field $\boldsymbol{B}$). The kinetic coefficients, which in this case might depend upon the parameter Ω (or $\boldsymbol{B}$), will now satisfy the relations

$$\gamma_{ij}(\Omega) = \gamma_{ji}(-\Omega) \tag{13a}$$

and

$$\gamma_{ij}(\boldsymbol{B}) = \gamma_{ji}(-\boldsymbol{B}). \tag{13b}$$

Secondly, our proof rested on the *implicit* assumption that the quantities x_i themselves do not change under time reversal. If, however, these quantities are proportional to the velocities of a certain macroscopic motion, then they will also change their sign under time reversal. Now, if both x_i and x_j belong to this category, then eqn. (12), which is crucial to our proof, would remain unaltered; consequently, the coefficients γ_{ij} and γ_{ji} would continue to be equal. However, if only one of them belongs to this category while the other does not, then eqn. (12) would change to

$$\langle x_i(0) \dot{x}_j(0) \rangle = -\langle \dot{x}_i(0) x_j(0) \rangle; \tag{12'}$$

the coefficients γ_{ij} and γ_{ji} would then satisfy the relations

$$\gamma_{ij} = -\gamma_{ji}. \tag{13'}$$

For the application of Onsager's relations to different physical problems, reference may be made to the monographs by de Groot (1951), de Groot and Mazur (1962) and Prigogine (1967).

Problems

14.1. Making use of expressions (14.1.11) and (14.1.12) for ΔS and ΔP, and expressions (14.1.14) for $\overline{(\Delta T)^2}$, $\overline{(\Delta V)^2}$ and $\overline{(\Delta T \Delta V)}$, show that

(i) $\overline{(\Delta T \Delta S)} = kT$; (ii) $\overline{(\Delta P \Delta V)} = -kT$;

(iii) $\overline{(\Delta S \Delta V)} = kT(\partial V/\partial T)_P$; (iv) $\overline{(\Delta P \Delta T)} = kT^2 C_V^{-1}(\partial P/\partial T)_V$.

[Note that results (i) and (ii) give: $\overline{(\Delta T \Delta S - \Delta P \Delta V)} = 2kT$, which follows directly from the probability distribution function (14.1.8).]

14.2. Establish the probability distribution (14.1.15), which leads to expressions (14.1.16) for $\overline{(\Delta S)^2}$, $\overline{(\Delta P)^2}$ and $\overline{(\Delta S \Delta P)}$. Show that these results can also be obtained by following the procedure of the preceding problem.

14.3. If we choose the quantities E and V as "independent" variables, then the probability distribution function (14.1.8) does not reduce to a form as simple as (14.1.13) or (14.1.15); it is marked instead by the presence, in the exponent, of a cross term proportional to the product $\Delta E \Delta V$. Consequently, the variables E and V are not *statistically independent*: $\overline{(\Delta E \Delta V)} \neq 0$.

Show that

$$\overline{(\Delta E \Delta V)} = kT\left\{T\left(\frac{\partial V}{\partial T}\right)_P + P\left(\frac{\partial V}{\partial P}\right)_T\right\};$$

verify as well expressions (14.1.14) and (14.1.18) for $\overline{(\Delta V)^2}$ and $\overline{(\Delta E)^2}$.

[Note that in the case of a two-dimensional normal distribution, viz.

$$p(x, y) \propto \exp\{-\tfrac{1}{2}(ax^2 + 2bxy + cy^2)\},$$

the quantities $\langle x^2\rangle$, $\langle xy\rangle$ and $\langle y^2\rangle$ can be obtained in a straightforward manner by carrying out a logarithmic differentiation of the formula

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}(ax^2 + 2bxy + cy^2)\right\} dx\,dy = \frac{2\pi}{\sqrt{(ac - b^2)}}$$

with respect to the parameters a, b and c. This leads to the *covariance matrix* of the distribution, namely

$$\begin{pmatrix} \langle x^2\rangle & \langle xy\rangle \\ \langle yx\rangle & \langle y^2\rangle \end{pmatrix} = \frac{1}{(ac - b^2)}\begin{pmatrix} c & -b \\ -b & a \end{pmatrix}$$

If $b = 0$, then

$$\langle x^2\rangle = 1/a, \quad \langle xy\rangle = 0, \quad \langle y^2\rangle = 1/c.$$][25]

14.4. A string of length l is stretched, under a constant tension F, between two fixed points A and B. Show that the mean square (fluctuational) displacement $y(x)$ at the point P, distant x from A, is given by

$$\overline{\{y(x)\}^2} = \frac{kT}{Fl}x(l - x).$$

Further show that, for $x_2 \geq x_1$,

$$\overline{y(x_1)y(x_2)} = \frac{kT}{Fl}x_1(l - x_2).$$

[*Hint:* Calculate the energy, Φ, associated with the fluctuation in question; the desired probability distribution is then given by $p \propto \exp(-\Phi/kT)$, from which the required averages can be readily evaluated.]

14.5. How small must the volume, V_A, of a gaseous sub-system (at normal temperature and pressure) be, so that the root-mean-square deviation in the number, N_A, of particles occupying this volume be 1 per cent of the mean value $\overline{N}_A$?

14.6. Consider a gas, of infinite extent, divided into regions A and B by an imaginary sheet running through the system. The molecules of the gas are supposed to be interacting through a potential energy function $u(r)$. Show that the net force $\boldsymbol{F}$ experienced by all the molecules on the A-side of the sheet due to all the molecules on the B-side is *perpendicular* to the plane of the sheet and its magnitude (per unit area) is given by

$$\frac{F}{A} = -\frac{2\pi n^2}{3} \int_0^\infty \left(\frac{du}{dr}\right) g(r) r^3 \, dr,$$

where the quantity $n^2 g(r)$ is defined in eqn. (14.2.3). Compare this result with eqn. (3.7.17), which was derived on the basis of the *virial theorem.*

14.7. Show that for a gas of non-interacting bosons, or fermions, the *correlation function* $\nu(r)$ is given by the expression

$$\nu(r) = \pm \frac{g}{n^2 h^6} \left| \int_{-\infty}^{\infty} \frac{e^{i(\boldsymbol{p}\cdot\boldsymbol{r})/\hbar} \, d^3p}{e^{(p^2/2m-\mu)/kT} \mp 1} \right|^2$$

where $g\,(= 2s + 1)$ is the spin multiplicity factor while the other symbols have their usual meanings; note that the upper sign here applies to bosons, the lower one to fermions.[26]

[*Hint:* To solve this problem, one may use the method of second quantization, as developed in Chapter 10. The *particle density operator* $\hat{n}$ is then given by the sum, see eqn. (10.1.25),

$$\sum_{\alpha,\beta} a_\alpha^\dagger a_\beta u_\alpha^*(\boldsymbol{r}) u_\beta(\boldsymbol{r}),$$

whose diagonal terms are directly related to the mean particle density n in the system. The nondiagonal terms give the *density fluctuation operator* $(\hat{n} - n)$, etc.]

14.8. Show that, in the case of a degenerate gas of fermions ($T \ll T_F$), the correlation function $\nu(r)$, for $r \gg \hbar/p_F$, reduces to the expression

$$\nu(r) = -\frac{3(mkT)^2}{4 p_F^3 \hbar r^2} \left\{ \sinh\left(\frac{\pi m k T r}{p_F \hbar}\right) \right\}^{-2}$$

Note that, as $T \to 0$, this expression tends to the limiting form

$$\nu(r) = -\frac{3\hbar}{4\pi^2 p_F r^4} \propto \frac{1}{r^4}.$$

14.9. Pospišil (1927) observed the Brownian motion of soot particles, of radii 0.4×10^{-4} cm, immersed in a water–glycerine solution, of viscosity 0.0278 poise, at a temperature of 18.8°C. The observed value of $\overline{x^2}$, in a 10-second time interval, was 3.3×10^{-8} cm^2. Making use of these data, determine the Boltzmann constant k.

14.10. In the notation of Sec. 14.4, show that for a Brownian particle

$$\langle \boldsymbol{v}(t) \cdot \boldsymbol{F}(t) \rangle = 3kT/\tau, \quad \text{while} \quad \langle \boldsymbol{v}(t) \cdot \boldsymbol{\mathcal{F}}(t) \rangle = 0.$$

On the other hand,

$$\langle \boldsymbol{r}(t) \cdot \boldsymbol{\mathcal{F}}(t) \rangle = -3kT, \quad \text{while} \quad \langle \boldsymbol{r}(t) \cdot \boldsymbol{F}(t) \rangle = 0.$$

14.11. Integrate eqn. (14.4.14) to obtain

$$\boldsymbol{r}(t) = \boldsymbol{v}(0)\tau(1 - e^{-t/\tau}) + \tau \int_0^t \{1 - e^{(u-t)/\tau}\} \boldsymbol{A}(u) \, du,$$

so that $r(0) = 0$. Taking the square of this expression and making use of the autocorrelation function $K_A(s)$, derive formula (14.4.31) for $\langle r^2(t)\rangle$.

14.12. While detecting a very feeble current with the help of a moving-coil galvanometer, one must ensure that an observed deflection is not just a stray kick arising from the Brownian motion of the suspended system. If we agree that a deflection θ, whose magnitude exceeds $4\theta_{\text{r.m.s.}}[= 4(kT/c)^{1/2}]$, is highly unlikely to be due to the Brownian motion, we obtain a *lower* limit to the magnitude of the current that can be reliably recorded with the help of the given galvanometer. Express this limiting current in terms of the time period τ and the critical damping resistance R_c of the galvanometer.

14.13. (a) Integrate Langevin's equation (14.4.5), for the velocity component v_x, over a *small* interval of time δt, and show that

$$\frac{\langle \delta v_x\rangle}{\delta t} = -\frac{v_x}{\tau} \quad \text{and} \quad \frac{\langle(\delta v_x)^2\rangle}{\delta t} = \frac{2kT}{M\tau}.$$

(b) Now, set up the Fokker–Planck equation for the *distribution function* $f(v_x, t)$ and, making use of the foregoing results for $\mu_1(v_x)$ and $\mu_2(v_x)$, derive an explicit expression for this function. Study the various cases of interest, especially the one for which $t \gg \tau$.

14.14. The Langevin equation for a particle, of mass M, executing Brownian motion, under a restoring force $-\lambda x$, would be (Kappler, 1938)

$$M\frac{d^2x}{dt^2} + \frac{1}{B}\frac{dx}{dt} + \lambda x = F(t);$$

cf. eqn. (14.4.2) for which $\lambda = 0$. Derive, on the basis of this equation, expressions for the quantities $\langle x^2(t)\rangle$ and $\langle v_x^2(t)\rangle$, and show that, in the limit $\lambda \to 0$, these expressions reproduce eqns (14.4.29) and (14.4.31) while, in the limit $M \to 0$, they lead to the relevant results of Sec. 14.5.

14.15. Generalize the Fokker–Planck equation to the case of a particle executing Brownian motion in *three* dimensions. Determine the general solution of this equation and study its important features

14.16. The autocorrelation function $K(s)$ of a certain *statistically stationary* variable $y(t)$ is given by

(i) $K(s) = K(0)e^{-\alpha s^2}\cos(2\pi f^* s)$

or by

(ii) $K(s) = K(0)e^{-\alpha|s|}\cos(2\pi f^* s)$,

where $\alpha > 0$. Determine, and discuss the nature of, the power spectrum $w(f)$ in both these cases, and investigate its behavior in the limits (a) $\alpha \to 0$, (b) $f^* \to 0$ and (c) both α and $f^* \to 0$.

14.17. Show that if the autocorrelation function $K(s)$ of a certain *statistically stationary* variable $y(t)$ is given by

$$K(s) = K(0)\frac{\sin(as)}{as}\frac{\sin(bs)}{bs} \quad (a > b > 0),$$

then the power spectrum $w(f)$ of that variable is given by

$$\begin{aligned} w(f) &= \frac{2\pi}{a}K(0) && \text{for} \quad 0 < f \le \frac{a-b}{2\pi}, \\ &\quad \frac{2\pi}{ab}K(0)\left\{\frac{a+b}{2} - \pi f\right\} && \text{for} \quad \frac{a-b}{2\pi} \le f \le \frac{a+b}{2\pi}, \\ &\quad 0 && \text{for} \quad \frac{a+b}{2\pi} \le f < \infty. \end{aligned}$$

Verify that the function $w(f)$ satisfies the requirement (14.6.16).

[Note that, in the limit $b \to 0$, we recover the situation pertaining to eqns (14.6.18).]

14.18. (a) Show that the mean square value of the variable $Y(t)$, defined by the formula

$$Y(t) = \int_u^{u+t} y(u)\,du,$$

where $y(u)$ is a *statistically stationary* variable with power spectrum $w(f)$, is given by

$$\langle Y^2(t)\rangle = \frac{1}{2\pi^2}\int_0^\infty \frac{w(f)}{f^2}\{1 - \cos(2\pi f t)\}\,df;$$

and, accordingly,

$$w(f) = 4\pi f \int_0^\infty \frac{\partial}{\partial t}\langle Y^2(t)\rangle \sin(2\pi f t)\, dt$$

$$= 2\int_0^\infty \frac{\partial^2}{\partial t^2}\langle Y^2(t)\rangle \cos(2\pi f t)\, dt.$$

For details, see MacDonald (1962), sec. 2.2.1. A comparison of the last result with eqn. (14.6.14) suggests that

$$K_y(s) = \frac{1}{2}\frac{\partial^2}{\partial s^2}\langle Y^2(s)\rangle.$$

(b) Apply the foregoing analysis to the motion of a Brownian particle, taking y to be the velocity of the particle and Y its displacement.

14.19. Show that the power spectra $w_v(f)$ and $w_A(f)$ of the fluctuating variables $v(t)$ and $A(t)$ that appear in the Langevin equation (14.4.5) are connected by the relation

$$w_v(f) = w_A(f)\frac{\tau^2}{1+(2\pi f\tau)^2},$$

τ being the relaxation time of the problem. Hence, $w_A(f) = 12kT/M\tau$.

14.20. (a) Verify eqns (14.7.7)–(14.7.9).

(b) Substituting expression (14.7.9) for $K_v(s)$ into eqn. (14.7.6), derive formula (14.4.31) for $\langle r^2(t)\rangle$.

Notes

[1] Some of the statements made here are true only for $\boldsymbol{r} \neq 0$; for $\boldsymbol{r} = 0$, they have to be modified. This is due to the fact that "the possibility of finding a particle at the point $\boldsymbol{r} \to 0$, when we already know that there is a particle at the point $\boldsymbol{r} = 0$, tends to be rather a certainty". This (*singular*) situation may be taken care of by including in the expression for $g(\boldsymbol{r})$, and hence in the expression for $\nu(\boldsymbol{r})$, a term with a delta function at $\boldsymbol{r} = 0$. Now, eqn. (3) tells us that, since the integral $\int_V F_2(\boldsymbol{r}_1, \boldsymbol{r}_2)\, d^3r_1\, d^3r_2$ is equal to $N(N-1)$, the integral $\int_V g(\boldsymbol{r})\, d^3r$ must be equal to $V(1-1/N)$. Therefore, in the absence of spatial correlations, we should have

$$g(\boldsymbol{r}) = 1 - n^{-1}\delta(\boldsymbol{r}),$$

rather than $g(\boldsymbol{r}) = 1$; correspondingly, the correlation function $\nu(\boldsymbol{r})$ should be $-n^{-1}\delta(\boldsymbol{r})$, rather than zero.

It follows that the statements made in the text become applicable to all $\boldsymbol{r}$ (including $\boldsymbol{r} = 0$) if we add to the functions $g(\boldsymbol{r})$ and $\nu(\boldsymbol{r})$, as defined in eqns (3) and (4), an *ad hoc* term $n^{-1}\delta(\boldsymbol{r})$.

[2] Compare eqn. (10) with eqns (3.7.17) and (3.7.18), which connect the pressure P and the internal energy U of a system of interacting particles with the pair distribution function $g(\boldsymbol{r})$. See also Problem 14.6.

[3] We note that for light waves, since λ is relatively large, for *all* values of r for which the pair distribution function $g(r)$ is significantly different from unity, the product $(sr) \ll 1$; consequently, $\sin(sr)/(sr) \simeq 1$. Expression (13) then reduces to (10).

[4] For a discussion of *critical fluctuations*, see Klein and Tisza (1949).

[5] Since the quantities x and t are *macroscopic* in nature while l and τ^* are microscopic, the numbers n and m are much larger than unity; consequently, it is safe to assume that they are *integral* as well.

[6] The term "reversible" here is related to the fact that the Newtonian equations of motion, which govern this class of phenomena, preserve their form if the direction of time is reversed (i.e. if we change t to $-t$, etc.); alternatively, if at any instant of time we reverse the velocities of the particles in a given mechanical system, the system would "retrace" its path exactly. This is not true of equations

describing "irreversible" phenomena, such as the *diffusion equation* (19), with which the phenomenon of Brownian motion is intimately related.

[7] It is easy to recognize the additional result that if n is even, then $p_n(m) \equiv 0$ for odd m, and if n is odd, then $p_n(m) \equiv 0$ for even m.

[8] In the next section we shall see that, for a *spherical* particle, $D = kT/6\pi\eta a$ where η is the coefficient of viscosity of the medium and a the radius of the Brownian particle. In the case under study, $T \simeq 300$ K, $\eta \simeq 10^{-2}$ poise and $a \simeq 4 \times 10^{-5}$ cm. Substituting these values, we obtain for the Boltzmann constant: $k \simeq 1.3 \times 10^{-16}$ erg/K.

[9] If Stokes's law is applicable, then $B = 1/(6\pi\eta a)$, where η is the coefficient of viscosity of the fluid and a the radius of the particle (assumed spherical).

[10] The process of "averaging over an ensemble" implies that we are imagining a large number of systems similar to the one originally under consideration and are taking an average over this collection at *any* time t. By the very nature of the function $F(t)$, the ensemble average $\langle F(t)\rangle$ must be zero at all times.

[11] This is so because we have no reason to expect a statistical correlation between the position $r(t)$ of the Brownian particle and the force $F(t)$ exerted on it by the molecules of the fluid; see, however, Manoliu and Kittel (1979). Of course, we do expect a correlation between the variables $v(t)$ and $F(t)$; consequently, $\langle v \cdot F\rangle \neq 0$ (see Problem 14.10).

[12] Note that the limiting solution (8) corresponds to "dropping out" the second term on the left-hand side of eqn. (6).

[13] Note that the limiting solution (10) corresponds to "dropping out" the first term on the left-hand side of eqn. (6).

[14] This is the only crucial step in the proof. It involves a "shift", by an amount s, in both instants of the measurement process; the equality results from the fact that the ensemble is supposed to be stationary.

[15] One may check that

$$\frac{d}{dt}\langle v^2(t)\rangle = \frac{2}{\tau}\left[v^2(\infty) - \langle v^2(t)\rangle\right] = -\frac{2}{\tau}\Delta\langle v^2(t)\rangle,$$

where $v^2(\infty) = 3kT/M$ and $\Delta\langle v^2(t)\rangle$ is the "deviation of the quantity concerned from its equilibrium value". In this form of the equation, we have a typical example of a "relaxation phenomenon", with *relaxation time* $\tau/2$.

[16] We are tacitly assuming here a "Markovian" situation where the *transition probability function* $W(x, x')$ depends *only* on the present position x (and, of course, the subsequent position x') of the particle but *not* on the previous history of the particle.

[17] In the case of fermions, an account must be taken of the *Pauli exclusion principle* which controls the "occupation of single-particle states in the system"; for instance, we cannot, in that case, consider a transition that tends to transfer a particle to a state which is already occupied. This requires an appropriate modification of the Master Equation.

[18] Clearly, this assumption limits our analysis to what may be called the "Brownian motion approximation", in which the object under consideration is presumed to be on a very different scale of magnitude than the molecules constituting the environment. It is obvious that if one tries to apply this sort of analysis to "understand" the behaviour of molecules *themselves*, one cannot hope for anything but a "crude, semi-quantitative" outcome.

[19] Obviously, this does not affect the spectral quality of the fluctuations, except that now we do not have a component with frequency zero. To represent the actual situation, one may have to add, to the resulting spectrum, a suitably weighted $\delta(f)$-term.

[20] This is essentially true of the rapidly fluctuating force $F(t)$ experienced by a Brownian particle due to the incessant molecular impulses received by it.

[21] It will be noted that the fluctuations constituting this result belong *entirely* to the region of the "white" noise, with $\Delta f = 1/\tau_0$; see eqn. (24), with $B = \tau/M$.

[22] We note that the foregoing results are essentially equivalent to Einstein's original result for charge fluctuations in a conductor, viz.

$$\langle \delta q^2\rangle_t = \frac{2kT}{R}t;$$

compare, as well, the Brownian-particle result: $\langle x^2\rangle_t = 2BkTt$.

[23] We note that the functions $K_A(s)$ and $K_F(s)$, which are nonzero only for $s = O(\tau^*)$, see eqn. (14.4.21), may, for certain purposes, be written as

$$K_A(s) = \frac{6kT}{M^2B}\delta(s) \quad \text{and} \quad K_F(s) = \frac{6kT}{B}\delta(s).$$

In this form, the functions are nonzero only for $s = 0$.

[24] In writing eqn. (1), and other subsequent equations, we follow the summation convention which implies *an automatic summation over a repeated index.*

[25] For the covariance matrix of an n-dimensional normal distribution, see Landau and Lifshitz (1958), sec. 110.

[26] Note that, in the classical limit ($\hbar \to 0$), the infinitely rapid oscillations of the factor $\exp\{i(\boldsymbol{p} \cdot \boldsymbol{r})/\hbar\}$ make the integral vanish. Consequently, for an ideal classical gas, the function $\nu(r)$ is identically zero.

It is not difficult to see that, for $n\lambda^3 \ll 1$ where $\lambda = h/\sqrt{(2\pi mkT)}$,

$$\nu(r) \simeq \pm\frac{1}{g}\exp(-2\pi r^2/\lambda^2);$$

cf. eqn. (5.5.27). We also note that $\nu(0)$ is *identically* equal to $\pm 1/g$, which is ± 1 if $s = 0$.

APPENDIXES

A. Influence of boundary conditions on the distribution of quantum states

In this appendix we examine, under different boundary conditions, the asymptotic distribution of single-particle states in a bounded continuum. For simplicity, we consider a cuboidal enclosure of sides a, b and c. The admissible solutions of the free-particle Schrödinger equation

$$\nabla^2\psi + k^2\psi = 0 \quad (k = p\hbar^{-1}), \tag{1}$$

which satisfy Dirichlet boundary conditions (namely, $\psi = 0$ everywhere at the boundary), are then given by

$$\psi_{lmn}(\boldsymbol{r}) \propto \sin\left(\frac{l\pi x}{a}\right)\sin\left(\frac{m\pi y}{b}\right)\sin\left(\frac{n\pi z}{c}\right) \tag{2}$$

with

$$k = \pi\left(\frac{l^2}{a^2} + \frac{m^2}{b^2} + \frac{n^2}{c^2}\right)^{1/2}; \quad l, m, n = 1, 2, 3, \ldots \tag{3}$$

Note that in this case none of the quantum numbers l, m or n can be zero, for otherwise our wave function would identically vanish. If, on the other hand, we impose Neumann boundary conditions (namely, $\partial\psi/\partial n = 0$ everywhere at the boundary), the desired solutions turn out to be

$$\psi_{lmn}(\boldsymbol{r}) \propto \cos\left(\frac{l\pi x}{a}\right)\cos\left(\frac{m\pi y}{b}\right)\cos\left(\frac{n\pi z}{c}\right), \tag{4}$$

with

$$l, m, n = 0, 1, 2, \ldots ; \tag{5}$$

clearly, the value zero of the quantum numbers is now allowed! In each case, however, the negative-integral values of the quantum numbers do not lead to any new wave functions.

The total number $g(K)$ of distinct wave functions ψ, with wave number k not exceeding a given value K, may be written as

$$g(K) = \sum_{l,m,n}{}' f(l, m, n), \tag{6}$$

where $f(l, m, n) = 1$ for the numbers (l, m, n) belong to the set (3) or (5), as the case may be; the summation $\sum'$ in each case is restricted by the condition

$$\left(\frac{l^2}{a^2} + \frac{m^2}{b^2} + \frac{n^2}{c^2}\right) \leq \frac{K^2}{\pi^2}. \tag{7}$$

We now define a sum

$$G(K) = \sum_{l,m,n}{}' f^*(l, m, n), \tag{8}$$

where $f^*(l, m, n) = 1$ for all integral values of l, m and n (positive, negative or zero), the restriction on the numbers (l, m, n) being the same as stated in (7). One can then show, by setting up correspondence of terms, that

$$\begin{aligned}\sum_{l,m,n}{}' f(l, m, n) = \tfrac{1}{8}\Bigg[&\sum_{l,m,n}{}' f^*(l, m, n)\\ &\mp \left\{\sum_{l,m}{}' f^*(l, m, 0) + \sum_{l,n}{}' f^*(l, 0, n) + \sum_{m,n}{}' f^*(0, m, n)\right\}\\ &+ \left\{\sum_{l}{}' f^*(l, 0, 0) + \sum_{m}{}' f^*(0, m, 0) + \sum_{n}{}' f^*(0, 0, n)\right\} \mp 1\Bigg];\end{aligned} \tag{9}$$

the upper and the lower signs correspond, respectively, to the Dirichlet and the Neumann boundary conditions.

Clearly, the first sum on the right-hand side of (9) denotes the number of lattice points in the ellipsoid $(X^2/a^2 + Y^2/b^2 + Z^2/c^2) = K^2/\pi^2$,[1] the next three sums denote the numbers of lattice points in the ellipses which are cross-sections of this ellipsoid with the Z-, Y- and X-planes, while the last three sums denote the numbers of lattice points on the principal axes of the ellipsoid. Now, if a, b and c are sufficiently large in comparison with π/K, one may replace these numbers by the corresponding volume, areas and lengths, respectively, with the result

$$g(K) = \frac{K^3}{6\pi^2}(abc) \mp \frac{K^2}{8\pi}(ab + ca + bc) + \frac{K}{4\pi}(a + b + c) \mp \frac{1}{8} + E(K); \tag{10}$$

the term $E(K)$ here denotes the net error committed in making the aforementioned replacements. We thus find that the main term of our result is directly proportional to the volume of the enclosure while the first correction term is proportional to its surface area (and, hence, represents the "surface effect"); the next-order term(s) appear in the nature of an "edge effect" and a "corner effect".

Now, a reference to the literature dealing with the determination of the "number of lattice points in a given domain" reveals that the error term $E(K)$ in eqn. (10) is $O(K^\alpha)$ where $1 < \alpha < 1.4$; hence, expression (10) for $g(K)$ is safe only up to the surface term (inclusive). In view of this, we may write

$$g(K) = \frac{K^3}{6\pi^2}V \mp \frac{K^2}{16\pi}S + \text{a lower-order remainder}; \tag{11}$$

in terms of ε^* where

$$\varepsilon^* = \frac{8mL^2}{h^2}\varepsilon = \frac{4L^2}{h^2}P^2 = \frac{L^2}{\pi^2}K^2, \tag{12}$$

eqn. (11) reduces to eqns (1.4.15) and (1.4.16) of the text.

In the case of *periodic* boundary conditions, namely

$$\psi(x, y, z) = \psi(x + a, y, z) = \psi(x, y + b, z) = \psi(x, v, z + c), \tag{13}$$

the appropriate wave functions are

$$\psi_{lmn}(\boldsymbol{r}) \propto \exp\{i(\boldsymbol{k} \cdot \boldsymbol{r})\}, \tag{14}$$

with

$$\boldsymbol{k} = 2\pi\left(\frac{l}{a}, \frac{m}{b}, \frac{n}{c}\right); \quad l, m, n = 0, \pm 1, \pm 2, \dots . \tag{15}$$

The number of free-particle states $g(K)$ is now given by

$$g(K) = \sum_{l,m,n}{}' f^*(l, m, n), \tag{16}$$

such that

$$(l^2/a^2 + m^2/b^2 + n^2/c^2) \le K^2/(4\pi^2). \tag{17}$$

This is precisely the number of lattice points in the ellipsoid with semiaxes $Ka/2\pi$, $Kb/2\pi$ and $Kc/2\pi$, which, allowing for the approximation made in the earlier cases, is just equal to the volume term in (11). Thus, in the case of periodic boundary conditions, we do not obtain a surface term in the expression for the density of states.

For further information on this topic, see Fedosov (1963, 1964), Pathria (1966), Chaba and Pathria (1973) and Baltes and Hilf (1976).

B. Certain mathematical functions

In this appendix we outline the main properties of certain mathematical functions which are of special importance to the subject matter of this text.

We start with the *gamma function* $\Gamma(\nu)$, which is identical with the *factorial function* $(\nu - 1)!$ and is defined by the integral

$$\Gamma(\nu) \equiv (\nu - 1)! = \int_0^\infty e^{-x}x^{\nu-1}\,dx; \quad \nu > 0. \tag{1}$$

First of all we note that

$$\Gamma(1) \equiv 0! = 1. \tag{2}$$

Next, integrating by parts, we obtain the recurrence formula

$$\Gamma(\nu) = \frac{1}{\nu}\Gamma(\nu + 1), \tag{3}$$

whence it follows that

$$\Gamma(\nu + 1) = \nu(\nu - 1)\dots(1 + p)p\Gamma(p), \quad 0 < p \le 1, \tag{4}$$

p being the fractional part of ν. For integral values of ν ($\nu = n$, say), we have the familiar representation,

$$\Gamma(n+1) \equiv n! = n(n-1)\ldots 2 \cdot 1; \tag{5}$$

on the other hand, if ν is half-odd integral ($\nu = m + \frac{1}{2}$, say), then

$$\begin{aligned} \Gamma\left(m+\tfrac{1}{2}\right) &\equiv \left(m-\tfrac{1}{2}\right)! = \left(m-\tfrac{1}{2}\right)\left(m-\tfrac{3}{2}\right)\cdots\tfrac{3}{2}\cdot\left(\tfrac{1}{2}\right)\Gamma\left(\tfrac{1}{2}\right) \\ &= \frac{(2m-1)(2m-3)\ldots 3\cdot 1}{2^m}\pi^{1/2}, \end{aligned} \tag{6}$$

where use has been made of the fact that, see eqn. (21),

$$\Gamma\left(\tfrac{1}{2}\right) \equiv \left(-\tfrac{1}{2}\right)! = \pi^{1/2}. \tag{7}$$

By repeated application of the recurrence formula (3), the definition of the function $\Gamma(\nu)$ can be extended to all ν, except for $\nu = 0, -1, -2, \ldots$ where the singularities of the function lie. The behavior of $\Gamma(\nu)$ in the neighborhood of a singularity can be determined by setting $\nu = -n + \varepsilon$, where $n = 0, 1, 2, \ldots$ and $|\varepsilon| \ll 1$, and using formula (3) $n+1$ times, with the result

$$\begin{aligned} \Gamma(-n+\varepsilon) &= \frac{1}{(-n+\varepsilon)(-n+1+\varepsilon)\ldots(-1+\varepsilon)\varepsilon}\Gamma(1+\varepsilon) \\ &\approx \frac{(-1)^n}{n!\varepsilon}. \end{aligned} \tag{8}$$

Replacing x by αy^2, eqn. (1) takes the form

$$\Gamma(\nu) = 2\alpha^{\nu}\int_0^{\infty} e^{-\alpha y^2} y^{2\nu-1}\,dy, \qquad \nu > 0. \tag{9}$$

Thus, we obtain another closely related integral, namely

$$I_{2\nu-1} \equiv \int_0^{\infty} e^{-\alpha y^2} y^{2\nu-1}\,dy = \frac{1}{2\alpha^{\nu}}\Gamma(\nu), \qquad \nu > 0; \tag{10}$$

by a change of notation, this can be written as

$$I_{\nu} \equiv \int_0^{\infty} e^{-\alpha y^2} y^{\nu}\,dy = \frac{1}{2\alpha^{(\nu+1)/2}}\Gamma\left(\frac{\nu+1}{2}\right), \qquad \nu > -1. \tag{11}$$

One can easily see that these integrals satisfy the relationship

$$I_{\nu+2} = -\frac{\partial}{\partial\alpha}I_{\nu}. \tag{12}$$

The integrals I_{ν} appear so frequently in our study that we write down the values of some of them explicitly:

$$I_0 = \frac{1}{2}\left(\frac{\pi}{\alpha}\right)^{1/2}, \quad I_2 = \frac{1}{4}\left(\frac{\pi}{\alpha^3}\right)^{1/2}, \quad I_4 = \frac{3}{8}\left(\frac{\pi}{\alpha^5}\right)^{1/2}\cdots, \tag{13a}$$

while

$$I_1 = \frac{1}{2\alpha}, \quad I_3 = \frac{1}{2\alpha^2}, \quad I_5 = \frac{1}{\alpha^3}, \cdots \tag{13b}$$

In connection with these integrals, it may also be noted that

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-\alpha y^2} y^\nu \, dy &= 0 \qquad \text{if } \nu \text{ is an odd integer} \\ &= 2I_\nu \quad \text{if } \nu \text{ is an even integer.} \end{aligned} \tag{14}$$

Next we consider the product of two gamma functions, say $\Gamma(\mu)$ and $\Gamma(\nu)$. Using representation (9), with $\alpha = 1$, we have

$$\Gamma(\mu)\Gamma(\nu) = 4 \int_0^\infty \int_0^\infty e^{-(x^2+y^2)} x^{2\mu-1} y^{2\nu-1} \, dx \, dy, \qquad \mu > 0, \nu > 0. \tag{15}$$

Changing over to the polar coordinates (r, θ), eqn. (15) becomes

$$\begin{aligned} \Gamma(\mu)\Gamma(\nu) &= 4 \int_0^\infty e^{-r^2} r^{2(\mu+\nu)-1} \, dr \int_0^{\pi/2} \cos^{2\mu-1}\theta \sin^{2\nu-1}\theta \, d\theta \\ &= 2\Gamma(\mu+\nu) \int_0^{\pi/2} \cos^{2\mu-1}\theta \sin^{2\nu-1}\theta \, d\theta. \end{aligned} \tag{16}$$

Defining the *beta function* $B(\mu, \nu)$ by the integral

$$B(\mu, \nu) = 2 \int_0^{\pi/2} \cos^{2\mu-1}\theta \sin^{2\nu-1}\theta \, d\theta, \qquad \mu > 0, \nu > 0, \tag{17}$$

we obtain an important relationship:

$$B(\mu, \nu) = \frac{\Gamma(\mu)\Gamma(\nu)}{\Gamma(\mu+\nu)} = B(\nu, \mu). \tag{18}$$

Substituting $\cos^2\theta = \eta$, eqn. (17) takes the standard form

$$B(\mu, \nu) = \int_0^1 \eta^{\mu-1}(1-\eta)^{\nu-1} \, d\eta, \qquad \mu > 0, \nu > 0, \tag{19}$$

while the special case $\mu = \nu = \frac{1}{2}$ gives

$$B\left(\tfrac{1}{2}, \tfrac{1}{2}\right) = 2 \int_0^{\pi/2} d\theta = \pi; \tag{20}$$

coupled with eqns (2) and (18), this yields

$$\Gamma\left(\tfrac{1}{2}\right) = \pi^{1/2}. \tag{21}$$

(i). Stirling's formula for $\nu!$

We now derive an asymptotic expression for the factorial function

$$\nu! = \int_0^\infty e^{-x} x^\nu \, dx \tag{22}$$

for $\nu \gg 1$. It is not difficult to see that, for $\nu \gg 1$, the main contribution to this integral comes from that region of x which lies around the point $x = \nu$ and has a width of order $\sqrt{\nu}$. In view of this, we invoke the substitution

$$x = \nu + (\sqrt{\nu})u, \tag{23}$$

whereby eqn. (22) takes the form

$$\nu! = \sqrt{\nu}\left(\frac{\nu}{e}\right)^{\nu} \int\limits_{-\sqrt{\nu}}^{\infty} e^{-(\sqrt{\nu})u}\left(1 + \frac{u}{\sqrt{\nu}}\right)^{\nu} du. \tag{24}$$

The integrand of (24) attains its maximum value, unity, at $u = 0$ and on both sides of the maximum it falls rapidly to zero, which suggests that it may be approximated by a Gaussian. We therefore expand the logarithm of the integrand around $u = 0$ and then reconstruct the integrand by taking the exponential of the resulting expression; this gives

$$\nu! = \sqrt{\nu}\left(\frac{\nu}{e}\right)^{\nu} \int\limits_{-\sqrt{\nu}}^{\infty} \exp\left\{-\frac{u^2}{2} + \frac{u^3}{3\sqrt{\nu}} - \frac{u^4}{4\nu} + \cdots\right\} du. \tag{25}$$

If ν is sufficiently large, the integrand in (25) may be replaced by the single factor $\exp(-u^2/2)$; moreover, since the major contribution to this integral comes only from that range of u for which $|u|$ is of order unity, the lower limit of integration may be replaced by $-\infty$. We thus obtain the *Stirling formula*

$$\nu! \approx \sqrt{(2\pi\nu)}(\nu/e)^{\nu}, \qquad \nu \gg 1. \tag{26}$$

A more detailed analysis leads to the Stirling series

$$\nu! = \sqrt{(2\pi\nu)}\left(\frac{\nu}{e}\right)^{\nu}\left[1 + \frac{1}{12\nu} + \frac{1}{288\nu^2} - \frac{139}{51840\nu^3} - \frac{571}{2488320\nu^4} + \cdots\right]. \tag{27}$$

Next, we consider the function $\ln(\nu!)$. Corresponding to formula (27), we have, for large ν,

$$\begin{aligned}\ln(\nu!) = {} & \left(\nu + \frac{1}{2}\right)\ln\nu - \nu + \frac{1}{2}\ln(2\pi) \\ & + \left[\frac{1}{12\nu} - \frac{1}{360\nu^3} + \frac{1}{1260\nu^5} - \frac{1}{1680\nu^7} + \cdots\right].\end{aligned} \tag{28}$$

For most purposes, we may write

$$\ln(\nu!) \approx (\nu\ln\nu - \nu). \tag{29}$$

We note that formula (29) can be obtained very simply by an application of the Euler–Maclaurin formula. Since ν is large, we may consider its integral values only; then, by definition,

$$\ln(n!) = \sum_{i=1}^{n}(\ln i).$$

Replacing summation by an integration, we obtain

$$\ln(n!) \simeq \int_1^n (\ln x)\,dx = (x\ln x - x)|_{x=1}^{x=n}$$
$$\approx (n\ln n - n),$$

which is identical with (29).

We must, however, be warned that whereas approximation (29) is fine as it is, it would be wrong to take its exponential and write $\nu! \approx (\nu/e)^\nu$, for that would affect the evaluation of $\nu!$ by a factor $O(\nu^{1/2})$, which can be considerably large; see (26). In the expression for $\ln(\nu!)$, the corresponding term is indeed negligible.

(ii). The Dirac δ-function

We start with the Gaussian distribution function

$$p(x, x_0, \sigma) = \frac{1}{\sqrt{(2\pi)}\sigma} e^{-(x-x_0)^2/2\sigma^2}, \tag{30}$$

which satisfies the normalization condition

$$\int_{-\infty}^{\infty} p(x, x_0, \sigma)\,dx = 1. \tag{31}$$

The function $p(x)$ is symmetric about the value x_0 where it has a maximum; the height of the maximum is inversely proportional to the parameter σ while its width is directly proportional to σ, the total area under the curve being a constant. As σ becomes smaller and smaller, the function $p(x)$ becomes narrower and narrower in width and grows higher and higher at the central point x_0, condition (31) being satisfied at all σ; see Fig. B.1.

In the limit $\sigma \to 0$, we obtain a function whose value at $x = x_0$ is infinitely large while at $x \neq x_0$ it is vanishingly small, the area under the curve being still equal to unity. This limiting form of the function $p(x, x_0, \sigma)$ defines the δ-function

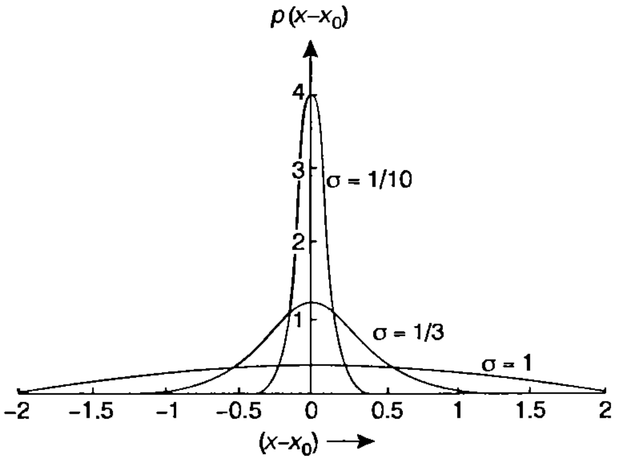


Fig. B.1.

of Dirac. Thus, we may define this function as the one satisfying the following properties:

$$\text{(i)}\ \delta(x-x_0)=0 \quad \text{for all } x\neq x_0 \tag{32}$$

and

$$\text{(ii)}\ \int_{-\infty}^{\infty}\delta(x-x_0)\,dx=1. \tag{33}$$

Conditions (32) and (33) inherently imply that, at $x=x_0$, $\delta(x-x_0)=\infty$ and that the range of integration in (33) need not extend all the way from $-\infty$ to $+\infty$; in fact, any range that includes the point $x=x_0$ would suffice. Thus, we may rewrite (33) as

$$\int_{A}^{B}\delta(x-x_0)\,dx=1 \quad \text{if } A<x_0<B; \tag{34}$$

it follows that for any well-behaved function $f(x)$

$$\int_{A}^{B}f(x)\delta(x-x_0)\,dx=f(x_0) \quad \text{if } A<x_0<B. \tag{35}$$

Another limiting process frequently employed to represent the δ-function is the following:

$$\delta(x-x_0)=\operatorname*{Lim}_{\gamma\to 0}\frac{\gamma}{\pi\{(x-x_0)^2+\gamma^2\}}. \tag{36}$$

To see the appropriateness of this representation, we note that, for $x\neq x_0$, this function vanishes (like γ) while, for $x=x_0$, it diverges (like γ^{-1}); moreover, for all γ,

$$\int_{-\infty}^{\infty}\frac{\gamma}{\pi\{(x-x_0)^2+\gamma^2\}}\,dx=\frac{1}{\pi}\left[\tan^{-1}\frac{(x-x_0)}{\gamma}\right]_{-\infty}^{\infty}=1. \tag{37}$$

An integral representation of the δ-function is

$$\delta(x-x_0)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{ik(x-x_0)}\,dk, \tag{38}$$

which means that the δ-function is the "Fourier transform of a constant". We note that, for $x=x_0$, the integrand in (38) is unity throughout, so the function diverges. On the other hand, for $x\neq x_0$, the oscillatory character of the integrand is such that it makes the integral vanish. And, finally, the integration of this function, over a range of x that includes the point $x=x_0$, gives

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{x_0-L}^{x_0+L}e^{ik(x-x_0)}\,dx\right]dk=\int_{-\infty}^{\infty}\frac{e^{ikL}-e^{-ikL}}{2\pi(ik)}\,dk$$

$$=\int_{-\infty}^{\infty}\frac{\sin(kL)}{\pi k}\,dk=1, \tag{39}$$

independently of the choice of L.

It is instructive to see how the integral representation of the δ-function is related to its previous representations. For this, we introduce into the integrand of (38) a factor $\exp(-\gamma k^2)$, where γ is a small, positive number. The resulting function, in the limit $\gamma \to 0$, should reproduce the δ-function; thus, we expect that

$$\delta(x - x_0) = \frac{1}{2\pi} \lim_{\gamma \to 0} \int_{-\infty}^{\infty} e^{ik(x-x_0)-\gamma k^2}\, dk. \tag{40}$$

The integral in (40) is easy to evaluate when we recall that

$$\int_{-\infty}^{\infty} \cos(kx) e^{-\gamma k^2}\, dk = 2 \int_{0}^{\infty} \cos(kx) e^{-\gamma k^2}\, dk = \sqrt{\left(\frac{\pi}{\gamma}\right)}\, e^{-x^2/4\gamma}, \tag{41}$$

while

$$\int_{-\infty}^{\infty} \sin(kx) e^{-\gamma k^2}\, dk = 0. \tag{42}$$

Accordingly, eqn. (40) becomes

$$\delta(x - x_0) = \lim_{\gamma \to 0} \frac{1}{\sqrt{(4\pi\gamma)}} e^{-(x-x_0)^2/4\gamma}, \tag{43}$$

which is precisely the representation we started with.[2]

Finally, the notation of the δ-function can be readily extended to spaces with more than one dimension. For instance, in n dimensions,

$$\delta(\boldsymbol{r}) = \delta(x_1) \ldots \delta(x_n), \tag{44}$$

so that

$$\text{(i)}\ \delta(\boldsymbol{r}) = 0 \quad \text{for all } \boldsymbol{r} \neq 0 \tag{45}$$

and

$$\text{(ii)}\ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \delta(\boldsymbol{r})\, dx_1 \ldots dx_n = 1. \tag{46}$$

The integral representation of $\delta(\boldsymbol{r})$ is

$$\delta(\boldsymbol{r}) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{i(\boldsymbol{k}\cdot\boldsymbol{r})}\, d^n k. \tag{47}$$

Once again, we may write

$$\delta(\boldsymbol{r}) = \frac{1}{(2\pi)^n} \lim_{\gamma \to 0} \int_{-\infty}^{\infty} \cdot \int_{-\infty}^{\infty} e^{i(\boldsymbol{k}\cdot\boldsymbol{r})-\gamma k^2}\, d^n k \tag{48}$$

$$= \lim_{\gamma \to 0} \left(\frac{1}{4\pi\gamma}\right)^{n/2} e^{-r^2/4\gamma}. \tag{49}$$

C. "Volume" and "surface area" of an n-dimensional sphere of radius R

Consider an n-dimensional space in which the position of a point is denoted by the vector $\boldsymbol{r}$, with Cartesian components $(x_1, \dots, x_n)$. The "volume element" dV_n in this space would be

$$d^n r = \prod_{i=1}^{n} (dx_i);$$

accordingly, the "volume" V_n of a sphere of radius R would be given by

$$V_n(R) = \underset{0 \le \sum_{i=1}^{n} x_i^2 \le R^2}{\int \cdots \int} \prod_{i=1}^{n} (dx_i). \tag{1}$$

Obviously, V_n will be proportional to R^n, so let it be written as

$$V_n(R) = C_n R^n, \tag{2}$$

where C_n is a constant which depends only on the dimensionality of the space. Clearly, the "volume element" dV_n can also be written as

$$dV_n = S_n(R)\, dR = n C_n R^{n-1}\, dR, \tag{3}$$

where $S_n(R)$ denotes the "surface area" of the sphere.

To evaluate C_n, we make use of the formula

$$\int_{-\infty}^{\infty} \exp(-x^2)\, dx = \pi^{1/2}. \tag{4}$$

Multiplying n such integrals, one for each x_i, we obtain

$$\begin{aligned} \pi^{n/2} &= \int_{x_i=-\infty}^{\infty} \cdots \int \exp\left(-\sum_{i=1}^{n} x_i^2\right) \prod_{i=1}^{n} (dx_i) \\ &= \int_0^{\infty} \exp(-R^2) n C_n R^{n-1}\, dR \\ &= n C_n \cdot \frac{1}{2} \Gamma\left(\frac{n}{2}\right) = \left(\frac{n}{2}\right)! C_n; \end{aligned} \tag{5}$$

here, use has been made of formula (B.11), with $\alpha = 1$. Thus

$$C_n = \pi^{n/2} \Big/ \left(\frac{n}{2}\right)!, \tag{6}$$

whence

$$V_n(R) = \frac{\pi^{n/2}}{(n/2)!} R^n \quad \text{and} \quad S_n(R) = \frac{2\pi^{n/2}}{\Gamma(n/2)} R^{n-1}, \tag{7a,b}$$

which are the desired results.

Alternatively, one may prefer to use spherical polar coordinates right from the beginning–as, for instance, in the evaluation of the Fourier transform

$$I(\boldsymbol{k}) = \int f(r) e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\, d^n r. \tag{8}$$

In that case

$$d^n r = r^{n-1}(\sin\theta_1)^{n-2}\dots(\sin\theta_{n-2})^1\,dr\,d\theta_1\dots d\theta_{n-2}\,d\phi, \tag{9}$$

where the θ_i range from 0 to π while ϕ ranges from 0 to 2π. Choosing our polar axis to be in the direction of $\boldsymbol{k}$, eqn. (8) takes the form

$$I(k) = \int f(r)e^{ikr\cos\theta_1}r^{n-1}(\sin\theta_1)^{n-2}\dots(\sin\theta_{n-2})^1\,dr\,d\theta_1\dots\,d\theta_{n-2}\,d\phi. \tag{10}$$

Integration over the angular coordinates $\theta_1, \theta_2, \theta_3, \dots, \theta_{n-2}$ and ϕ yields factors

$$\pi^{1/2}\Gamma\left(\frac{n-1}{2}\right)\left(\frac{2}{kr}\right)^{(n-2)/2} J_{(n-2)/2}(kr)\times$$
$$B\left(\frac{n-2}{2},\frac{1}{2}\right)\cdot B\left(\frac{n-3}{2},\frac{1}{2}\right)\dots B\left(1,\frac{1}{2}\right)\cdot 2\pi,$$

where $J_\nu(x)$ is the ordinary Bessel function while $B(\mu, \nu)$ is the beta function; see eqns (B.17) and (B.18). Equation (10) now becomes

$$I(k) = (2\pi)^{n/2}\int_0^\infty f(r)\left(\frac{1}{kr}\right)^{(n-2)/2} J_{(n-2)/2}(kr)r^{n-1}\,dr, \tag{11}$$

which is our main result.

In the limit $k \to 0$, $J_\nu(kr) \to \left(\frac{1}{2}kr\right)^\nu/\Gamma(\nu+1)$, so that

$$I(0) = \frac{2\pi^{n/2}}{\Gamma(n/2)}\int_0^\infty f(r)r^{n-1}\,dr, \tag{12}$$

consistent with (3) and (7b). On the other hand, if we take $f(r)$ to be a constant, say $1/(2\pi)^n$, we should obtain another representation of the Dirac δ-function in n dimensions; see eqns (8) and (B.47). We thus have, from (11),

$$\delta(\boldsymbol{k}) = \frac{1}{(2\pi)^{n/2}}\int_0^\infty \left(\frac{1}{kr}\right)^{(n-2)/2} J_{(n-2)/2}(kr)r^{n-1}\,dr. \tag{13}$$

As a check, we introduce a factor $\exp(-\alpha r^2)$ in the integrand of (13) and obtain

$$\delta(\boldsymbol{k}) = \operatorname*{Lim}_{\alpha\to 0}\frac{1}{(2\pi)^{n/2}}\int_0^\infty e^{-\alpha r^2}\left(\frac{1}{kr}\right)^{(n-2)/2} J_{(n-2)/2}(kr)r^{n-1}\,dr$$
$$= \operatorname*{Lim}_{\alpha\to 0}\left(\frac{1}{4\pi\alpha}\right)^{n/2} e^{-k^2/4\alpha}, \tag{14}$$

in complete agreement with (B.49). If, on the other hand, we use the factor $\exp(-\alpha r)$ rather than $\exp(-\alpha r^2)$, we get

$$\delta(\boldsymbol{k}) = \operatorname*{Lim}_{\alpha\to 0}\Gamma\left(\frac{n+1}{2}\right)\frac{\alpha}{\{\pi(k^2+\alpha^2)\}^{(n+1)/2}}, \tag{15}$$

which generalizes (B.36).

D. On Bose–Einstein functions

In the theory of Bose–Einstein systems we come across integrals of the type

$$G_\nu(z) = \int_0^\infty \frac{x^{\nu-1}\,dx}{z^{-1}e^x - 1} \quad (0 \le z < 1, \nu > 0; z = 1, \nu > 1). \tag{1}$$

In this appendix we study the behavior of $G_\nu(z)$ over the stated range[3] of the parameter z. First of all, we note that

$$\operatorname*{Lim}_{z\to 0} G_\nu(z) = \int_0^\infty z e^{-x} x^{\nu-1}\,dx = z\Gamma(\nu). \tag{2}$$

Hence, it appears useful to introduce another function, $g_\nu(z)$, such that

$$g_\nu(z) \equiv \frac{1}{\Gamma(\nu)} G_\nu(z) = \frac{1}{\Gamma(\nu)} \int_0^\infty \frac{x^{\nu-1}\,dx}{z^{-1}e^x - 1}. \tag{3}$$

For small z, the integrand in (3) may be expanded in powers of z, with the result

$$\begin{aligned} g_\nu(z) &= \frac{1}{\Gamma(\nu)} \int_0^\infty x^{\nu-1} \sum_{l=1}^\infty (ze^{-x})^l\,dx \\ &= \sum_{l=1}^\infty \frac{z^l}{l^\nu} = z + \frac{z^2}{2^\nu} + \frac{z^3}{3^\nu} + \cdots; \end{aligned} \tag{4}$$

thus, for $z \ll 1$, the function $g_\nu(z)$, *for all* ν, behaves like z itself. Moreover, $g_\nu(z)$ is a monotonically increasing function of z whose largest value in the physical range of interest obtains when $z \to 1$; then, for $\nu > 1$, $g_\nu(z)$ approaches the Riemann *zeta function* $\zeta(\nu)$:

$$g_\nu(1) = \sum_{l=1}^\infty \frac{1}{l^\nu} = \zeta(\nu) \quad (\nu > 1). \tag{5}$$

The numerical values of some of the $\zeta(\nu)$ are

$$\zeta(2) = \frac{\pi^2}{6} \simeq 1.645, \quad \zeta(4) = \frac{\pi^4}{90} \simeq 1.082, \quad \zeta(6) = \frac{\pi^6}{945} \simeq 1.017,$$
$$\zeta\left(\tfrac{3}{2}\right) \simeq 2.612, \quad \zeta\left(\tfrac{5}{2}\right) \simeq 1.341, \quad \zeta\left(\tfrac{7}{2}\right) \simeq 1.127,$$

and, finally,

$$\zeta(3) \simeq 1.202, \quad \zeta(5) \simeq 1.037, \quad \zeta(7) \simeq 1.008.$$

For $\nu \le 1$, the function $g_\nu(z)$ diverges as $z \to 1$. The case $\nu = 1$ is rather simple, for the function $g_\nu(z)$ now assumes a closed form:

$$g_1(z) = \int_0^\infty \frac{dx}{z^{-1}e^x - 1} = \ln(1 - ze^{-x})\Big|_0^\infty = -\ln(1 - z) \tag{6}$$

As $z \to 1$, $g_1(z)$ diverges logarithmically. Setting $z = e^{-\alpha}$, we have

$$g_1(e^{-\alpha}) = -\ln(1 - e^{-\alpha}) \xrightarrow[\alpha \to 0]{} \ln(1/\alpha). \tag{7}$$

For $\nu < 1$, the behavior of $g_\nu(e^{-\alpha})$, as $\alpha \to 0$, can be determined as follows:

$$g_\nu(e^{-\alpha}) = \frac{1}{\Gamma(\nu)} \int_0^\infty \frac{x^{\nu-1}\,dx}{e^{\alpha+x} - 1} \approx \frac{1}{\Gamma(\nu)} \int_0^\infty \frac{x^{\nu-1}\,dx}{\alpha + x}.$$

Setting $x = \alpha \tan^2\theta$ and making use of eqn. (B.17), we obtain

$$g_\nu(e^{-\alpha}) \approx \frac{\Gamma(1-\nu)}{\alpha^{1-\nu}} \quad (0 < \nu < 1). \tag{8}$$

Expression (8) isolates the singularity of the function $g_\nu(e^{-\alpha})$ at $\alpha = 0$; the remainder of the function can be expanded in powers of α, with the result (see Robinson, 1951)

$$g_\nu(e^{-\alpha}) = \frac{\Gamma(1-\nu)}{\alpha^{1-\nu}} + \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} \zeta(\nu - i)\alpha^i, \tag{9}$$

$\zeta(s)$ being the Riemann zeta function analytically continued to all $s \neq 1$.

A simple differentiation of $g_\nu(z)$ brings out the recurrence relation

$$z\frac{\partial}{\partial z}[g_\nu(z)] \equiv \frac{\partial}{\partial(\ln z)} g_\nu(z) = g_{\nu-1}(z). \tag{10}$$

This relation follows readily from the series expansion (4) but can also be derived from the defining integral (3). We thus have

$$z\frac{\partial}{\partial z}[g_\nu(z)] = \frac{z}{\Gamma(\nu)} \int_0^\infty \frac{e^x x^{\nu-1}\,dx}{(e^x - z)^2}.$$

Integrating by parts, we get

$$z\frac{\partial}{\partial z}[g_\nu(z)] = \frac{z}{\Gamma(\nu)} \left[-\frac{x^{\nu-1}}{e^x - z}\bigg|_0^\infty + (\nu - 1) \int_0^\infty \frac{x^{\nu-2}\,dx}{e^x - z} \right].$$

The integrated part vanishes at both limits (provided that $\nu > 1$), while the part yet to be integrated yields precisely $g_{\nu-1}(z)$. The validity of the recurrence relation (10) is thus established for all $\nu > 1$.

Adopting (10) as a part of the definition of the function $g_\nu(z)$, the notion of this function may be extended to *all* ν, including $\nu \leq 0$. Proceeding in this manner, Robinson showed that eqn. (9) applied to all $\nu < 1$ and to all *non-integral* $\nu > 1$. For $\nu = m$, a positive integer, we have instead

$$g_m(e^{-\alpha}) = \frac{(-1)^{m-1}}{(m-1)!} \left[\sum_{i=1}^{m-1} \frac{1}{i} - \ln\alpha \right] \alpha^{m-1} + \sum_{\substack{i=0 \\ i \neq m-1}}^{\infty} \frac{(-1)^i}{i!} \zeta(m - i)\alpha^i. \tag{11}$$

Equations (9) and (11) together provide a complete description of the function $g_\nu(e^{-\alpha})$ for small α; it may be checked that both these expressions conform to the recurrence relation

$$\frac{\partial}{\partial\alpha} g_\nu(e^{-\alpha}) = -g_{\nu-1}(e^{-\alpha}). \tag{12}$$

For the special cases $\nu = \frac{5}{2}, \frac{3}{2}$ and $\frac{1}{2}$, we obtain from (9)

$$g_{5/2}(\alpha) = 2.36\alpha^{3/2} + 1.34 - 2.61\alpha - 0.730\alpha^2 + 0.0347\alpha^3 + \cdots, \tag{13a}$$

$$g_{3/2}(\alpha) = -3.54\alpha^{1/2} + 2.61 + 1.46\alpha - 0.104\alpha^2 + 0.00425\alpha^3 + \cdots, \tag{13b}$$

$$g_{1/2}(\alpha) = 1.77\alpha^{-1/2} - 1.46 + 0.208\alpha - 0.0128\alpha^2 + \ldots \tag{13c}$$

The terms quoted here are sufficient to yield a better than 1 per cent accuracy for all $\alpha \leq 1$. The numerical values of these functions have been tabulated by London (1954) over the range $0 \leq \alpha \leq 2$.

E. On Fermi–Dirac functions

In the theory of Fermi–Dirac systems we come across integrals of the type

$$F_\nu(z) = \int_0^\infty \frac{x^{\nu-1}\,dx}{z^{-1}e^x + 1} \qquad (0 \leq z < \infty,\ \nu > 0). \tag{1}$$

In this appendix we study the behavior of $F_\nu(z)$ over the entire range of the parameter z. For the same reason as in the case of Bose–Einstein integrals, we introduce here another function, $f_\nu(z)$, such that

$$f_\nu(z) \equiv \frac{1}{\Gamma(\nu)} F_\nu(z) = \frac{1}{\Gamma(\nu)} \int_0^\infty \frac{x^{\nu-1}\,dx}{z^{-1}e^x + 1}. \tag{2}$$

For small z, the integrand in (2) may be expanded in powers of z, with the result

$$\begin{aligned} f_\nu(z) &= \frac{1}{\Gamma(\nu)} \int_0^\infty x^{\nu-1} \sum_{l=1}^\infty (-1)^{l-1} (ze^{-x})^l\,dx \\ &= \sum_{l=1}^\infty (-1)^{l-1} \frac{z^l}{l^\nu} = z - \frac{z^2}{2^\nu} + \frac{z^3}{3^\nu} - \cdots; \end{aligned} \tag{3}$$

thus, for $z \ll 1$, the function $f_\nu(z)$, *for all* ν, behaves like z itself.

The functions $f_\nu(z)$ and $f_{\nu-1}(z)$ are connected through the recurrence relation

$$z\frac{\partial}{\partial z}[f_\nu(z)] \equiv \frac{\partial}{\partial(\ln z)} f_\nu(z) = f_{\nu-1}(z); \tag{4}$$

this relation follows readily from the series expansion (3) but can also be derived from the defining integral (2).

To study the behavior of Fermi–Dirac integrals for large z, we introduce the variable

$$\xi = \ln z, \tag{5}$$

so that

$$F_\nu(e^\xi) \equiv \Gamma(\nu) f_\nu(e^\xi) = \int_0^\infty \frac{x^{\nu-1}dx}{e^{x-\xi}+1}. \tag{6}$$

For large ξ, the situation in (6) is primarily controlled by the factor $(e^{x-\xi}+1)^{-1}$, whose departure from its limiting values–namely, zero (as $x \to \infty$) and almost unity (as $x \to 0$)–is significant only in the neighborhood of the point $x = \xi$; see Fig. E.1. The width of this "region of significance" is $O(1)$ and hence *much smaller* than the total, effective range of integration, which is $O(\xi)$. Therefore, in the lowest approximation, we may replace the actual curve of Fig. E.1 by a step function, as shown by the dotted line. Eqn. (6) then reduces to

$$F_\nu(e^\xi) \approx \int_0^\xi x^{\nu-1}\,dx = \frac{\xi^\nu}{\nu} \tag{7}$$

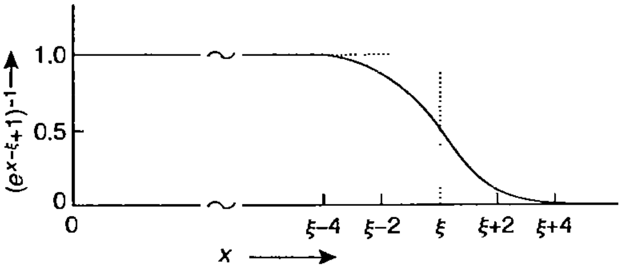


FIG. E.1.

and, accordingly,

$$f_\nu(e^\xi) \approx \frac{\xi^\nu}{\Gamma(\nu+1)}. \tag{8}$$

For a better approximation, we rewrite (6) as

$$F_\nu(e^\xi) = \int_0^\xi x^{\nu-1}\left[1 - \frac{1}{e^{\xi-x}+1}\right] dx + \int_\xi^\infty x^{\nu-1}\frac{1}{e^{x-\xi}+1}\,dx \tag{9}$$

and substitute in the respective integrals

$$x = \xi - \eta_1 \quad \text{and} \quad x = \xi + \eta_2, \tag{10}$$

with the result

$$F_\nu(e^\xi) = \frac{\xi^\nu}{\nu} - \int_0^\xi \frac{(\xi-\eta_1)^{\nu-1}\,d\eta_1}{e^{\eta_1}+1} + \int_0^\infty \frac{(\xi+\eta_2)^{\nu-1}\,d\eta_2}{e^{\eta_2}+1} \tag{11}$$

Since $\xi \gg 1$ while our integrands are significant only for η of order unity, the upper limit in the first integral may be replaced by ∞. Moreover, one may use the same variable η in both the integrals, with the result

$$F_\nu(e^\xi) \approx \frac{\xi^\nu}{\nu} + \int_0^\infty \frac{(\xi+\eta)^{\nu-1} - (\xi-\eta)^{\nu-1}}{e^\eta + 1} \, d\eta \tag{12}$$

$$= \frac{\xi^\nu}{\nu} + 2 \sum_{j=1,3,5,\ldots} \binom{\nu-1}{j} \left[\xi^{\nu-1-j} \int_0^\infty \frac{\eta^j}{e^\eta + 1} \, d\eta \right]; \tag{13}$$

in the last step the numerator in the integrand of (12) has been expanded in powers of η. Now,

$$\frac{1}{\Gamma(j+1)} \int_0^\infty \frac{\eta^j}{e^\eta + 1} \, d\eta = 1 - \frac{1}{2^{j+1}} + \frac{1}{3^{j+1}} - \cdots$$

$$= \left(1 - \frac{1}{2^j}\right) \zeta(j+1); \tag{14}$$

see eqns (2) and (3), with $\nu = j + 1$ and $z = 1$. Substituting (14) into (13), we obtain

$$f_\nu(e^\xi) = \frac{\xi^\nu}{\Gamma(\nu+1)} \left[1 + 2\nu \sum_{j=1,3,5,\ldots} \left\{ (\nu-1)\ldots(\nu-j) \left(1 - \frac{1}{2^j}\right) \frac{\zeta(j+1)}{\xi^{j+1}} \right\} \right]$$

$$= \frac{\xi^\nu}{\Gamma(\nu+1)} \left[1 + \nu(\nu-1)\frac{\pi^2}{6}\frac{1}{\xi^2} + \nu(\nu-1)(\nu-2)(\nu-3)\frac{7\pi^4}{360}\frac{1}{\xi^4} + \ldots \right], \tag{15}$$

which is the desired asymptotic formula–commonly known as Sommerfeld's lemma (see Sommerfeld, 1928).[4]

By the same procedure, one can derive the following asymptotic result, which is clearly a generalization of (15):

$$\int_0^\infty \frac{\phi(x)\,dx}{e^{x-\xi} + 1} = \int_0^\xi \phi(x)\,dx + \frac{\pi^2}{6}\left(\frac{d\phi}{dx}\right)_{x=\xi} + \frac{7\pi^4}{360}\left(\frac{d^3\phi}{dx^3}\right)_{x=\xi}$$

$$+ \frac{31\pi^6}{15120}\left(\frac{d^5\phi}{dx^5}\right)_{x=\xi} + \cdots, \tag{16}$$

where $\phi(x)$ is any well-behaved function of x. It may be noted that the numerical coefficients in this expansion approach the limiting value 2.

Blakemore (1962) has tabulated numerical values of the function $f_\nu(e^\xi)$ in the range $-4 \leq \xi \leq +10$; his tables cover all integral orders from 0 to +5 and all half-odd integral orders from $-\frac{1}{2}$ to $+\frac{9}{2}$.

F. On Watson functions

In this appendix we examine the asymptotic behavior of the functions

$$W_d(\phi) = \int_0^\infty e^{-\phi x} \left[e^{-x} I_0(x)\right]^d \, dx \tag{1}$$

for $0 \leq \phi \ll 1$. First of all we note that if we set $\phi = 0$ the resulting integral converges only if $d > 2$. To see this, we observe that, with $\phi = 0$, convergence problems may arise in the limit of large x where the integrand

$$\left[e^{-x} I_0(x)\right]^d \approx (2\pi x)^{-d/2} \quad (x \gg 1). \tag{2}$$

Clearly, the integral will converge if $d > 2$; otherwise, it will diverge. We therefore conclude that

$$W_d(0) = \int_0^\infty \left[e^{-x} I_0(x)\right]^d dx \tag{3}$$

exists for $d > 2$.

Next we look at the derivative

$$W_d'(\phi) = -\int_0^\infty e^{-\phi x} \left[e^{-x} I_0(x)\right]^d x\, dx. \tag{4}$$

By the same argument as above, we conclude that

$$W_d'(0) = -\int_0^\infty \left[e^{-x} I_0(x)\right]^d x\, dx \tag{5}$$

exists for $d > 4$. The manner in which $W_d'(\phi)$ diverges for $d < 4$, as $\phi \to 0$, can be seen as follows:

$$\begin{aligned} W_d'(\phi) &= -\int_0^\infty e^{-y} \left[e^{-y/\phi} I_0(y/\phi)\right]^d \frac{1}{\phi^2} y\, dy \\ &\approx -\frac{1}{(2\pi)^{d/2} \phi^{(4-d)/2}} \int_0^\infty e^{-y} y^{(2-d)/2}\, dy \quad (\phi \ll 1) \\ &= -\frac{\Gamma\{(4-d)/2\}}{(2\pi)^{d/2} \phi^{(4-d)/2}}. \end{aligned} \tag{6}$$

Integrating (6) with respect to ϕ, and remembering the comments made earlier about $W_d(0)$, we obtain the desired results:

$$W_d(\phi) \approx \begin{cases} (2\pi)^{-d/2} \Gamma\{(2-d)/2\} \phi^{-(2-d)/2} + \text{const.} & \text{for} \quad d < 2 \quad \text{(7a)} \\ (2\pi)^{-1} \ln(1/\phi) + \text{const.} & \text{for} \quad d = 2 \quad \text{(7b)} \\ W_d(0) - (2\pi)^{-d/2} |\Gamma\{(2-d)/2\}| \phi^{(d-2)/2} & \text{for} \quad 2 < d < 4. \quad \text{(7c)} \end{cases}$$

For $d > 4$, we have a simpler result,

$$W_d(\phi) \approx W_d(0) - |W_d'(0)| \phi, \tag{8}$$

for in this case both $W_d(0)$ and $W_d'(0)$ exist.

The borderline case $d = 4$ presents some problems which can be simplified by splitting integral (4) into two parts:

$$\int_0^\infty = \int_0^1 + \int_1^\infty . \tag{9}$$

The first part is clearly finite; the divergence of the function $W_4'(\phi)$, as $\phi \to 0$, arises from the second part which, for $\phi \ll 1$, can be written as

$$\approx \int_1^\infty e^{-\phi x}(2\pi x)^{-2} x\,dx = \frac{1}{4\pi^2}E_1(\phi), \tag{10}$$

where $E_1(\phi)$ is the exponential integral; see Abramowitz and Stegun (1964), Chap. 5. Since $E_1(\phi)$, for $\phi \ll 1$, $\approx -\ln\phi$, we conclude that

$$W_4'(\phi) \approx \frac{1}{4\pi^2}\ln\phi. \tag{11}$$

Integrating (11) with respect to ϕ, we obtain

$$W_4(\phi) \approx W_4(0) - \frac{1}{4\pi^2}\phi\ln(1/\phi). \tag{12}$$

Equations (7), (8) and (12) constitute the main results of this appendix.

For the record, we quote some numbers:

$$W_3(0) = 0.50546, \qquad W_4(0) = 0.30987. \tag{13}$$

Notes

[1] By the term "in the ellipsoid" we mean "not external to the ellipsoid", i.e. the lattice points "on the ellipsoid" are also included. Other such expressions in the sequel carry a similar meaning.

[2] The reader may check that the introduction into (38) of a factor $\exp(-\gamma|k|)$, rather than $\exp(-\gamma k^2)$, leads to the representation (36).

[3] The behavior of $G_\nu(z)$ for $z > 1$ has been discussed by Clunie (1954).

[4] A more careful analysis carried out by Rhodes (1950), and followed by Dingle (1956), shows that the passage from eqn. (11) to (12) omits a term which, for large ξ, is of order $e^{-\xi}$. This term turns out to be $\cos\{(\nu-1)\pi\}F_\nu(e^{-\xi}) \equiv \cos\{(\nu-1)\pi\}F_\nu(1/z)$. For large z, this would be very nearly equal to $\cos\{(\nu-1)\pi\}/z$ and hence negligible in comparison with any of the terms appearing in (15). Of course, for $\nu = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$, which are the values occurring in most of the important applications of Fermi–Dirac statistics, the missing term is identically zero.

For $\nu = 2$, the inclusion of the missing term leads to the identity

$$f_2(e^{\xi}) + f_2(e^{-\xi}) = \frac{1}{2}\xi^2 + \frac{\pi^2}{6},$$

which is relevant to the contents of Sec. 8.3B.

BIBLIOGRAPHY

Abe, R. (1972) *Prog. Theor. Phys. Japan* **48**, 1414.

Abe, R. (1973) *Prog. Theor. Phys. Japan* **49**, 113, 1851.

Abe, R. and Hikami, S. (1973) *Prog. Theor. Phys. Japan* **49**, 442.

Abraham, D. B. (1986) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and J. L. Lebowitz (Academic Press, London), Vol. 10, pp. 1–74.

Abramowitz, M. and Stegun, I. A. (eds.) (1964) *Handbook of Mathematical Functions* (National Bureau of Standards, Washington, D.C.).

Abrikosov, A. A. and Khalatnikov, I. M. (1957) *J. Exptl. Theoret. Phys. USSR* **33**, 1154; English transl. *Soviet Phys. JETP* **6**, 888 (1958).

Abrikosov, A. A., Gor'kov, L. P. and Dzyaloshinskii, I. Y. (1965) *Quantum Field Theoretical Methods in Statistical Physics* (Pergamon Press, Oxford)

Aharony, A. (1976) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and M. S. Green (Academic Press, London), Vol. 6, pp. 357–424.

Ahlers, G. (1980) *Rev. Mod. Phys.* **52**, 489.

Allen, S. and Pathria, R. K. (1989) *Can. J. Phys.* **67**, 952.

Allen, S. and Pathria, R. K. (1991) *Can. J. Phys.* **69**, 753.

Als-Nielsen, J. (1976) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and M. S. Green (Academic Press, London), Vol. 5a, pp. 87–164.

Als-Nielsen, J. and Birgeneau, R. J. (1977) *Am. J. Phys.* **45**, 554.

Anderson, P. W. (1984) *Basic Notions of Condensed Matter Physics* (Benjamin/Cummings Publishing Company, Menlo Park, California).

Auluck, F. C. and Kothari, D. S. (1946) *Proc. Camb. Phil. Soc.* **42**, 272.

Auluck, F. C. and Kothari, D. S. (1946–47) *Proc. Roy. Irish Academy*, p. 13.

Au-Yang, H. and Perk, J. H. H. (1984) *Phys. Lett. A* **104**, 131.

Bak, P., Krinsky, S. and Mukamel, D. (1976) *Phys. Rev. Lett.* **36**, 52.

Baker, G. A., Jr. (1990) *Quantitative Theory of Critical Phenomena* (Academic Press, New York).

Baltes, H. P. and Hilf, E. R. (1976) *Spectra of Finite Systems* (Bibliographisches Institut AG, Zürich).

Band, W. (1955) *An Introduction to Quantum Statistics* (Van Nostrand, Princeton).

Barber, M. N. (1983) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and J. L. Lebowitz (Academic Press, London), Vol. 8, pp. 145–266.

Barber, M. N. and Fisher, M. E. (1973) *Annals of Physics (N.Y.)* **77**, 1.

Bardeen, J., Cooper, L. N. and Schrieffer, J. R. (1957) *Phys. Rev.* **108**, 1175.

Barnett, S. J. (1944) *Proc. Am. Acad. Arts & Sci.* **75**, 109.

Barouch, E. *et al.* (1973) *Phys. Rev. Lett.* **31**, 1409.

Baxter, R. J. (1982) *Exactly Solved Models in Statistical Mechanics* (Academic Press, London).

Belinfante, F. J. (1939) *Physica* **6**, 849, 870.

Berlin, T. H. and Kac, M. (1952) *Phys. Rev.* **86**, 821.

Bernoulli, D. (1738) *Hydrodynamica* (Argentorati).

Beth, E. and Uhlenbeck, G. E. (1937) *Physica* **4**, 915; see also (1936) *Physica* **3**, 729.

Bethe, H. A. (1935) *Proc. Roy. Soc. London A* **150**, 552.

Betts, D. D. (1974) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and M. S. Green (Academic Press, London), Vol. 3, pp. 569–652.

Betts, D. D. and Lee, M. H. (1968) *Phys. Rev. Lett.* **20**, 1507.

Betts, D. D. *et al.* (1969) *Phys. Lett.* **29A**, 150; (1970) *ibid.* **32A**, 152.

Betts, D. D. *et al.* (1970) *Can. J. Phys.* **48**, 1566; (1971) *ibid.* **49**, 1327; (1973) *ibid.* **51**, 2249.

Betts, D. S. (1969) *Contemp. Phys.* **10**, 241.

Bijl, A. (1940) *Physica* **7**, 869.

Binder, K. (1983) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and J. L. Lebowitz (Academic Press, London), Vol. 8, pp. 1–144.

Binder, K. (ed.) (1986) *Monte Carlo Methods in Statistical Physics*, 2nd edn., Topics in Current Physics (Springer-Verlag, Berlin), Vol. 7.

Binder, K. (ed.) (1987) *Applications of the Monte Carlo Method in Statistical Physics*, 2nd edn., Topics in Current Physics (Springer-Verlag, Berlin), Vol. 36.

Binder, K. (ed.) (1992) *The Monte Carlo Method in Condensed Matter Physics*, Topics in Applied Physics (Springer-Verlag, Berlin), Vol. 71.

Blakemore, J. S. (1962) *Semiconductor Statistics* (Pergamon Press, Oxford).

Bogoliubov, N. N. (1947) *J. Phys. USSR* **11**, 23.

Bogoliubov, N. N. (1962) *Phys. Abhandl. S.U.* **6**, 113.

Bogoliubov, N. N. (1963) *Lectures on Quantum Statistics* (Gordon & Breach, New York).

Boltzmann, L. (1868) *Wien. Ber.* **58**, 517.

Boltzmann, L. (1871) *Wien. Ber.* **63**, 397, 679, 712.

Boltzmann, L. (1872) *Wien. Ber.* **66**, 275.

Boltzmann, L. (1875) *Wien. Ber.* **72**, 427.

Boltzmann, L. (1876) *Wien. Ber.* **74**, 503.

Boltzmann, L. (1877) *Wien. Ber.* **76**, 373.

Boltzmann, L. (1879) *Wien. Ber.* **78**, 7.

Boltzmann, L. (1884) *Ann. Phys.* **22**, 31, 291, 616.

Boltzmann, L. (1896, 1898) *Vorlesungen über Gastheorie*, 2 vols. (J. A. Barth, Leipzig). English transl. (1964) *Lectures on Gas Theory* (Berkeley, California).

Boltzmann, L. (1899) *Amsterdam Ber.* 477; see also (1909) *Wiss. Abhandl. von L. Boltzmann III*, 658.

Bose, S. N. (1924) *Z. Physik* **26**, 178.

Bragg, W. L. and Williams, E. J. (1934) *Proc. Roy. Soc. London A* **145**, 699.

Bragg, W. L. and Williams, E. J. (1935) *Proc. Roy. Soc. London A* **151**, 540; **152**, 231.

Brézin, E., le Guillou, J. C. and Zinn-Justin, J. (1976) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and M. S. Green (Academic Press, London), Vol. 6, pp. 125–247.

Brueckner, K. A. and Gammel, J. L. (1958) *Phys. Rev.* **109**, 1040.

Brueckner, K. A. and Sawada, K. (1957) *Phys. Rev.* **106**, 1117, 1128.

Brush, S. G. (1957) *Annals of Sci.* **13**, 188, 273.

Brush, S. G. (1958) *Annals of Sci.* **14**, 185, 243.

Brush, S. G. (1961a) *Am. J. Phys.* **29**, 593.

Brush, S. G. (1961b) *Am. Scientist* **49**, 202.

Brush, S. G. (1965–66) *Kinetic Theory*, 3 vols. (Pergamon Press, Oxford).

Brush, S. G. (1967) *Rev. Mod. Phys.* **39**, 883.

Buckingham, M. J. and Gunton, J. D. (1969) *Phys. Rev.* **178**, 848.

Bush, V. and Caldwell, S. H. (1931) *Phys. Rev.* **38**, 1898.

Cardy, J. L. (1987) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and J. L. Lebowitz (Academic Press, London), Vol. 11, pp. 55–126.

Cardy, J. L. (ed.) (1988) *Finite-size Scaling* (North-Holland Publishing Company, Amsterdam).

Chaba, A. N. and Pathria, R. K. (1973) *Phys. Rev. A* **8**, 3264.

Chandrasekhar, S. (1939) *Introduction to the Study of Stellar Structure* (University of Chicago Press, Chicago): now available in a Dover edition.

Chandrasekhar, S. (1943) *Rev. Mod. Phys.* **15**, 1.

Chandrasekhar, S. (1949) *Rev. Mod. Phys.* **21**, 383.

Chang, C. H. (1952) *Phys. Rev.* **88**, 1422.

Chang, T. S. (1939) *Proc. Camb. Phil. Soc.* **35**, 265.

Chapman, S. (1916) *Trans. Roy. Soc. London A* **216**, 279.

Chapman, S. (1917) *Trans. Roy. Soc. London A* **217**, 115.

Chapman, S. and Cowling, T. G. (1939) *The Mathematical Theory of Non-uniform Gases* (Cambridge University Press).

Chisholm, J. S. R. and Borde, A. H. (1958) *An Introduction to Statistical Mechanics* (Pergamon Press, New York).

Chrétien, M., Gross, E. P. and Deser, S. (eds.) (1968) *Statistical Physics, Phase Transitions and Superfluidity*, 2 vols. (Gordon & Breach, New York).

Clausius, R. (1857) *Ann. Phys.* **100**, 353 (= (1857) *Phil. Mag.* **14**, 108).

Clausius, R. (1859) *Ann. Phys.* **105**, 239 (= (1859) *Phil. Mag.* **17**, 81).

Clausius, R. (1870) *Ann. Phys.* **141**, 124 (= (1870) *Phil. Mag.* **40**, 122).

des Cloizeaux, J. (1974) *Phys. Rev. A* **10**, 1665.

des Cloizeaux, J. (1982) in *Phase Transitions: Cargèse 1980*, eds. M. Lévy *et al.* (Plenum Press, New York), pp. 371–394.

Clunie, J. (1954) *Proc. Phys. Soc. A* **67**, 632.

Cohen, E. G. D. and van Leeuwen, J. M. J. (1960) *Physica* **26**, 1171.

Cohen, E. G. D. and van Leeuwen, J. M. J. (1961) *Physica* **27**, 1157.

Cohen, E. G. D. (ed.) (1962) *Fundamental Problems in Statistical Mechanics* (John Wiley, New York).

Cohen, M. and Feynman R. P. (1957) *Phys. Rev.* **107**, 13.

Compton, A. H. (1923a) *Phys. Rev.* **21**, 207, 483.

Compton, A. H. (1923b) *Phil. Mag.* **46**, 897.

Cooper, L. N. (1956) *Phys. Rev.* **104**, 1189.

Cooper, L. N. (1960) *Am. J. Phys.* **28**, 91.

Corak, Garfunkel, Satterthwaite and Wexler (1955) *Phys. Rev.* **98**, 1699.

Darwin, C. G. and Fowler, R. H. (1922a) *Phil. Mag.* **44**, 450, 823.

Darwin, C. G. and Fowler, R. H. (1922b) *Proc. Camb. Phil. Soc.* **21**, 262.

Darwin, C. G. and Fowler, R. H. (1923) *Proc. Camb. Phil. Soc.* **21**, 391, 730; see also Fowler, R. H. (1923) *Phil. Mag.* **45**, 1, 497; (1925) *Proc. Camb. Phil. Soc.* **22**, 861; (1926) *Phil. Mag.* **1**, 845; (1926) *Proc. Roy. Soc. London A* **113**, 432.

Debye, P. (1912) *Ann. Phys.* **39**, 789.

Dennison, D. M. (1927) *Proc. Roy. Soc. London A* **115**, 483.

Diehl, H. W. (1986) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and J. L. Lebowitz (Academic Press, London), Vol. 10, pp. 75–267.

Dingle, R. B. (1956) *J. App. Res. B* **6**, 225.

Dirac, P. A. M. (1926) *Proc. Roy. Soc. London A* **112**, 661, 671.

Dirac, P. A. M. (1929) *Proc. Camb. Phil. Soc.* **25**, 62.

Dirac, P. A. M. (1930) *Proc. Camb. Phil. Soc.* **26**, 361, 376.

Dirac, P. A. M. (1931) *Proc. Camb. Phil. Soc.* **27**, 240.

Domb, C. (1960) *Advances in Physics* **9**, 150.

Domb, C. (1974) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and M. S. Green (Academic Press, London), Vol. 3, pp. 357–484.

Domb, C. (1985) *Contemp. Phys.* **26**, 49.

Domb, C. and Green, M. S. (eds.) (1972–1976) *Phase Transitions and Critical Phenomena* (Academic Press, London), Vols. 1–6.

Domb, C. and Hunter, D. L. (1965) *Proc. Phys. Soc.* **86**, 1147.

Domb, C. and Lebowitz, J. L. (eds.) (1983–1992) *Phase Transitions and Critical Phenomena* (Academic Press, London), Vols. 7–15.

Drude, P. (1900) *Ann. Phys.* **1**, 566; **3**, 369.

Ehrenfest, P. (1905) *Wiener Ber.* **114**, 1301.

Ehrenfest, P. (1906) *Phys. Zeits.* **7**, 528.

Ehrenfest, P. and T. (1912) *Enzyklopädie der mathematischen Wissenschaften*, vol. IV (Teubner, Leipzig). English transl. (1959) *The Conceptual Foundations of the Statistical Approach in Mechanics* (Cornell University Press, Ithaca).

Einstein, A. (1902) *Ann. Phys.* **9**, 417.

Einstein, A. (1903) *Ann. Phys.* **11**, 170.

Einstein, A. (1905a) *Ann. Phys.* **17**, 132; see also (1906b) *ibid.* **20**, 199.

Einstein, A. (1905b) *Ann. Phys.* **17**, 549; see also (1906a) *ibid.* **19**, 289, 371.

Einstein, A. (1905c) *Ann. Phys.* **17**, 891; see also (1905d) *ibid.* **18**, 639 and (1906c) *ibid.* **20**, 627.

Einstein, A. (1907) *Jb. Radioakt.* **4**, 411.

Einstein, A. (1909) *Physik Z.* **10**, 185.

Einstein, A. (1910) *Ann. Phys.* **33**, 1275.

Einstein, A. (1924) *Berliner Ber.* **22**, 261.

Einstein, A. (1925) *Berliner Ber.* **1**, 3.

Eisenschitz, R. (1958) *Statistical Theory of Irreversible Processes* (Oxford University Press).

Elcock, E. W. and Landsberg, P. T. (1957) *Proc. Phys. Soc.* **70**, 161.

Enskog, D. (1917) *Kinetische Theorie der Vorgänge in mässig verdünnten Gasen*, dissertation (Almquist & Wiksell, Uppsala).

Eyring, H., Henderson, D., Stover, B. J. and Eyring, E. M. (1963) *Statistical Mechanics and Dynamics* (John Wiley, New York).

Farquhar, I. E. (1964) *Ergodic Theory in Statistical Mechanics* (Interscience Publishers, New York).

Fay, J. A. (1965) *Molecular Thermodynamics* (Addison-Wesley, Reading, Mass.).

Fedosov, B. V. (1963) *Dokl. Akad. Nauk SSSR* **151**, 786; English transl. *Sov. Math.* (Doklady) **4**, 1092.

Fedosov, B. V. (1964) *Dokl. Akad. Nauk SSSR* **157**, 536; English transl. *Sov. Math.* (Doklady) **5**, 988.

Fermi, E. (1926) *Z. Physik* **36**, 902.

Fermi, E. (1928) *Zeit. für Phys.* **48**, 73; see also (1927) *Acc. Lencei* **6**, 602.

Fetter, A. L. (1963) *Phys. Rev. Lett.* **10**, 507.

Fetter, A. L. (1965) *Phys. Rev.* **138**, A 429.

Feynman, R. P. (1953) *Phys. Rev.* **91**, 1291, 1301.

Feynman, R. P. (1954) *Phys. Rev.* **94**, 262.

Feynman, R. P. (1955) *Progress In Low Temperature Physics*, ed. C. J. Gorter (North-Holland, Amsterdam) 1.

Feynman, R. P. and Cohen, M. (1956) *Phys. Rev.* **102**, 1189.

Fisher, M. E. (1959) *Physica* **25**, 521.
Fisher, M. E. (1964) *J. Math. Phys.* **5**, 944.
Fisher, M. E. (1967) *Rep. Prog. Phys.* **30**, 615.
Fisher, M. E. (1969) *Phys. Rev.* **180**, 594.
Fisher, M. E. (1972) in *Essays in Physics*, eds. G. K. T. Conn and G. N. Fowler (Academic Press, London), Vol. 4, pp. 43–89.
Fisher, M. E. (1983) in *Critical Phenomena*, ed. F. J. W. Hahne (Springer-Verlag, Berlin), pp. 1–139.
Fisher, M. E., Barber, M. N. and Jasnow, D. (1973) *Phys. Rev. A* **8**, 1111.
Fisher, M. E. and Privman, V. (1985) *Phys. Rev. B* **32**, 447.
Fokker, A. D. (1914) *Ann. Phys.* **43**, 812.
Fowler, R. H. (1926) *Monthly Not. R.A.S.* **87**, 114.
Fowler, R. H. (1955) *Statistical Mechanics*, 2nd ed. (Cambridge University Press).
Fowler, R. H. and Guggenheim, E. A. (1940) *Proc. Roy. Soc. London A* **174**, 189.
Fowler, R. H. and Guggenheim, E. A. (1960) *Statistical Thermodynamics* (Cambridge University Press).
Fowler, R. H. and Nordheim, L. (1928) *Proc. Roy. Soc. London A* **119**, 173.
Fujiwara, ter Haar and Wergeland (1970) *J. Stat. Phys.* **2**, 329.
Galitskii, V. M. (1958) *J. Exptl. Theor. Phys. USSR* **34**, 151; English transl. *Soviet Physics JETP* **7**, 104.
Gaunt, D. S. and Guttmann, A. J. (1974) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and M. S. Green (Academic Press, London), Vol. 3, pp. 181–243.
de Gennes, P. -G. (1972) *Phys. Lett.* **38A**, 339.
de Gennes, P. -G. (1979) *Scaling Concepts in Polymer Physics* (Cornell University Press).
Gibbs, J. W. (1902) *Elementary Principles in Statistical Mechanics* (Yale University Press, New Haven); reprinted by Dover Publications, New York (1960). See also *A Commentary on the Scientific Writings of J. Willard Gibbs*, ed. A. Haas, Yale University Press, New Haven (1936), Vol. II.
Ginzburg, V. L. (1960) *Sov. Phys. Solid State* **2**, 1824.
Glasser, M. L. (1970) *Am. J. Phys.* **38**, 1033.
Gol'dman, Krivchenkov, Kogan and Galitskii (1960) *Problems in Quantum Mechanics*, translated by D. ter Haar (Academic Press, New York).
Gombás, P. (1949) *Die statistische Theorie des Atoms und ihre Anwendungen* (Springer, Vienna).
Gombás, P. (1952) *Ann. Phys.* **10**, 253.
Gradshteyn, I. S. and Ryzhik, I. M. (1965) *Table of Integrals, Series and Products* (Academic Press, New York).
Green, H. S. and Hurst, C. A. (1964) *Order–Disorder Phenomena* (Interscience Publishers, New York).
Greenspoon, S. and Pathria, R. K. (1974) *Phys. Rev. A* **9**, 2103.
Griffiths, R. B. (1965a) *Phys. Rev. Lett.* **14**, 623.
Griffiths, R. B. (1965b) *J. Chem. Phys.* **43**, 1958.
Griffiths, R. B. (1970) *Phys. Rev. Lett.* **24**, 715.
Griffiths, R. B. (1972) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and M. S. Green (Academic Press, London), Vol. 1, pp. 7–109.
de Groot, S. R. (1951) *Thermodynamics of Irreversible Processes* (North-Holland, Amsterdam).
de Groot, S. R. and Mazur, P. (1962) *Non-equilibrium Thermodynamics* (North-Holland, Amsterdam).
Gross, E. P. (1961) *Nuovo Cim.* **20**, 454.
Gross, E. P. (1963) *J. Math. Phys.* **4**, 195.
Guggenheim, E. A. (1935) *Proc. Roy. Soc. London A* **148**, 304.
Guggenheim, E. A. (1938) *Proc. Roy. Soc. London A* **169**, 134.
Guggenheim, E. A. (1959) *Boltzmann's Distribution Law* (North-Holland, Amsterdam).
Guggenheim, E. A. (1960) *Elements of the Kinetic Theory of Gases* (Pergamon Press, Oxford).
Gunton, J. D. and Buckingham, M. J. (1968) *Phys. Rev.* **166**, 152.
Gupta, H. (1947) *Proc. Nat. Ins. Sci. India* **13**, 35.
Guttmann, A. J. (1989) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and J. L. Lebowitz (Academic Press, London), Vol. 13, pp. 1–234.
ter Haar, D. (1954) *Elements of Statistical Mechanics* (Rinehart, New York).
ter Haar, D. (1955) *Rev. Mod. Phys.* **27**, 289.
ter Haar, D. (1961) *Rep. Prog. Phys.* **24**, 304.
ter Haar, D. (1966) *Elements of Thermostatistics* (Holt, Rinehart & Winston, New York).
ter Haar, D. (1967) *The Old Quantum Theory* (Pergamon Press, Oxford).
ter Haar, D. (1968) *On the History of Photon Statistics*, Course 42 of the International School of Physics "Enrico Fermi" (Academic Press, New York).

ter Haar, D. and Wergeland, H. (1966) *Elements of Thermodynamics* (Addison-Wesley, Reading, Ma.

Haber, H. E. and Weldon, H. A. (1981) *Phys. Rev. Lett.* **46**, 1497.

Haber, H. E. and Weldon, H. A. (1982) *Phys. Rev. D* **25**, 502.

Hanbury Brown, R. and Twiss, R. Q. (1956) *Nature* **177**, 27; **178**, 1447.

Hanbury Brown, R. and Twiss, R. Q. (1957) *Proc. Roy. Soc. London A* **242**, 300; **243**, 291.

Hanbury Brown, R. and Twiss, R. Q. (1958) *Proc. Roy. Soc. London A* **248**, 199, 222.

Heisenberg, W. (1928) *Z. Physik* **49**, 619.

Heller, P. (1967) *Rep. Prog. Phys.* **30**, 731.

Henry, W. E. (1952) *Phys. Rev.* **88**, 561.

Henshaw, D. G. (1960) *Phys. Rev.* **119**, 9.

Henshaw, D. G. and Woods, A. D. B. (1961) *Phys. Rev.* **121**, 1266.

Herapath, J. (1821) *Ann. Philos.* **1**, 273, 340, 401.

Hill, T. L. (1953) *J. Phys. Chem.* **57**, 324.

Hill, T. L. (1956) Statistical Mechanics (McGraw-Hill, New York).

Hill, T. L. (1960) *Introduction to Statistical Thermodynamics* (Addison-Wesley, Reading, Mass.).

Hirschfelder, J. O., Curtiss, C. F. and Bird, R. B. (1954) *Molecular Theory of Gases and Liquids* (John Wiley, New York).

Holliday, D. and Sage, M. L. (1964) *Annals of Physics* **29**, 125.

van Hove, L. (1949) *Physica*, **15**, 951.

Huang, K. (1959) *Phys. Rev.* **115**, 765.

Huang, K. (1960) *Phys. Rev.* **119**, 1129.

Huang, K. (1963) *Statistical Mechanics*, 1st ed. (John Wiley, New York).

Huang, K. (1987) *Statistical Mechanics*, 2nd ed. (John Wiley, New York).

Huang, K. and Klein, A. (1964) *Annals of Physics* **30**, 203.

Huang, K. and Yang, C. N. (1957) *Phys. Rev.* **105**, 767.

Huang, K., Yang, C. N. and Luttinger, J. M. (1957) *Phys. Rev.* **105**, 776.

Hupse, J. C. (1942) *Physica* **9**, 633.

Hurst, C. A. and Green, H. S. (1960) *J. Chem. Phys.* **33**, 1059.

Ising, E. (1925) *Z. Physik* **31**, 253.

Jackson, H. W. and Feenberg, E. (1962) *Rev. Mod. Phys.* **34**, 686.

Jasnow, D. (1986) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and J. L. Lebowitz (Academic Press, London), Vol. 10, pp. 269–363.

Jeans, J. H. (1905) *Phil. Mag.* **10**, 91.

Johnson, J. B. (1927a) *Nature* **119**, 50.

Johnson, J. B. (1927b) *Phys. Rev.* **29**, 367.

Johnson, J. B. (1928) *Phys. Rev.* **32**, 97.

Josephson, B. D. (1967) *Proc. Phys. Soc.* **92**, 269, 276.

Joule, J. P. (1851) *Mem. and Proc. Manchester Lit. and Phil. Soc.* **9**, 107 (= (1857) *Phil. Mag.* **14**, 211).

Joyce, G. S. (1972) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and M. S. Green (Academic Press, London), Vol. 2, pp. 375–442.

Kac, M. and Thompson, C. J. (1971) *Phys. Norveg.* **5**, 163.

Kac, M., Uhlenbeck. G. E. and Hemmer, P. C. (1963) *J. Math. Phys.* **4**, 216, 229; see also (1964) *ibid.* **5**, 60.

Kac, M. and Ward, J. C. (1952) *Phys. Rev.* **88**, 1332.

Kadanoff, L. P. (1966a) *Nuovo Cim.* **44**, 276.

Kadanoff, L. P. (1966b) *Physics* **2**, 263.

Kadanoff, L. P. (1976a) *Annals of Physics* **100**, 359.

Kadanoff, L. P. (1976b) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and M. S. Green (Academic Press, London), Vol. 5a, pp. 1–34.

Kadanoff, L. P. *et al.* (1967) *Rev. Mod. Phys.* **39**, 395.

Kahn, B. (1938) *On the Theory of the Equation of State*, Thesis, Utrecht; English translation appears in de Boer and Uhlenbeck, eds. (1965) *Studies in Statistical Mechanics*, Vol. III (North-Holland, Amsterdam).

Kahn, B. and Uhlenbeck, G. E. (1938) *Physica* **5**, 399.

Kappler, E. (1931) *Ann. Phys.* **11**, 233.

Kappler, E. (1938) *Ann. Phys.* **31**, 377, 619.

Kasteleyn, P. W. (1963) *J. Math. Phys.* **4**, 287.

Katsura, S. (1959) *Phys. Rev.* **115**, 1417.

Kaufman, B. (1949) *Phys. Rev.* **76**, 1232.

Kaufman, B. and Onsager, L. (1949) *Phys. Rev.* **76**, 1244.

Kawatra, M. P. and Pathria, R. K. (1966) *Phys. Rev.* **151**, 132.

Keesom, P. H. and Pearlman, N. (1953) *Phys. Rev.* **91**, 1354.

Khalatnikov, I. M. (1965) *Introduction to the Theory of Superfluidity* (W. A. Benjamin, New York).

Khintchine, A. (1934) *Math. Ann.* **109**, 604.

Kiess, E. (1987) *Am. J. Phys.* **55**, 1006.

Kilpatrick, J. E. (1953) *J. Chem. Phys.* **21**, 274.

Kilpatrick, J. E. and Ford, D. I. (1969) *Am. J. Phys.* **37**, 881.

Kirkwood, J. G. (1965) *Quantum Statistics and Cooperative Phenomena* (Gordon & Breach, New York).

Kittel, C. (1958) *Elementary Statistical Physics* (John Wiley, New York).

Kittel, C. (1969) *Thermal Physics* (John Wiley, New York).

Klein, M. J. and Tisza, L. (1949) *Phys. Rev.* **76**, 1861.

de Klerk, D., Hudson R. P. and Pellam, J. R. (1953) *Phys. Rev.* **89**, 326, 662.

Knobler, C. M. and Scott, R. L. (1984) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and J. L. Lebowitz (Academic Press, London), Vol. 9, pp. 163–231.

Kosterlitz, J. M. (1974) *J. Phys. C* **7**, 1046.

Kosterlitz, J. M. and Thouless, D. J. (1972) *J. Phys. C* **5**, L124.

Kosterlitz, J. M. and Thouless, D. J. (1973) *J. Phys. C* **6**, 1181.

Kosterlitz, J. M. and Thouless, D. J. (1978) *Prog. Low Temp. Phys.* **78**, 371.

Kothari, D. S. and Auluck, F. C. (1957) *Curr. Sci.* **26**, 169.

Kothari, D. S. and Singh, B. N. (1941) *Proc. Roy. Soc. London A* **178**, 135.

Kothari, D. S. and Singh, B. N. (1942) *Proc. Roy. Soc. London A* **180**, 414.

Kramers, H. A. (1938) *Proc. Kon. Ned. Akad. Wet. (Amsterdam)* **41**, 10.

Kramers, H. A. and Wannier, G. H. (1941) *Phys. Rev.* **60**, 252, 263.

Krönig, A. (1856) *Ann. Phys.* **99**, 315.

Kubo, R. (1956) *Can. J. Phys.* **34**, 1274.

Kubo, R. (1957) *Proc. Phys. Soc. Japan* **12**, 570.

Kubo, R. (1959) *Some Aspects of the Statistical Mechanical Theory of Irreversible Processes*, University of Colorado *Lectures in Theoretical Physics* (Interscience Publishers, New York), Vol. 1, p. 120.

Kubo, R. (1965) *Statistical Mechanics* (Interscience Publishers, New York).

Kubo, R. (1966) *Rep. Prog. Phys.* **29**, 255.

Landau, L. D. (1927) *Z. Phys.* **45**, 430; reprinted in the *Collected Papers of L. D. Landau*, ed. D. ter Haar (Pergamon Press, Oxford), p. 8.

Landau, L. D. (1930) *Z. Phys.* **64**, 629.

Landau, L. D. (1937) *Phys. Z. Sowjetunion* **11**, 26; reprinted in the *Collected Papers of L. D. Landau*, ed. D. ter Haar (Pergamon Press, Oxford), p. 193.

Landau, L. D. (1941) *J. Phys. USSR* **5**, 71.

Landau, L. D. (1947) *J. Phys. USSR* **11**, 91.

Landau, L. D. (1956) *J. Exptl. Theoret. Phys. USSR* **30**, 1058; English transl. *Soviet Physics JETP* **3**, 920 (1957).

Landau, L. D. and Lifshitz, E. M. (1958) *Statistical Physics* (Pergamon Press, Oxford).

Landsberg, P. T. (1954a) *Proc. Nat. Acad. Sciences (U.S.A.)* **40**, 149.

Landsberg, P. T. (1954b) *Proc. Camb. Phil. Soc.* **50**, 65.

Landsberg, P. T. (1961) *Thermodynamics with Quantum Statistical Illustrations* (Interscience Publishers, New York).

Landsberg, P. T. and Dunning-Davies, J. (1965) *Phys. Rev.* **138**, A1049; see also their article in the *Statistical Mechanics of Equilibrium and Non-equilibrium* (1965), ed. J. Meixner (North-Holland Publishing Co., Amsterdam).

Langevin, P. (1905a) *J. Phys.* **4**, 678.

Langevin, P. (1905b) *Ann. Chim. et Phys.* **5**, 70.

Langevin, P. (1908) *Comptes Rend. Acad. Sci. Paris* **146**, 530.

Lawrie, I. D. and Sarbach, S. (1984) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and J. L. Lebowitz (Academic Press, London), Vol. 9, pp. 1–161.

Lee, J. F., Sears, F. W. and Turcotte, D. L. (1963) *Statistical Thermodynamics* (Addison-Wesley, Reading, Mass.).

Lee, T. D., Huang, K. and Yang, C. N. (1957) *Phys. Rev.* **106**, 1135.

Lee, T. D. and Yang, C. N. (1952) *Phys. Rev.* **87**, 410; see also *ibid.*, p. 404.

Lee, T. D. and Yang, C. N. (1957) *Phys. Rev.* **105**, 1119.

Lee, T. D. and Yang, C. N. (1958) *Phys. Rev.* **112**, 1419.

Lee, T. D. and Yang, C. N. (1959a,b) *Phys. Rev.* **113**, 1165; **116**, 25.

Lee, T. D. and Yang, C. N. (1960a,b,c) *Phys. Rev.* **117**, 12, 22, 897.

Lennard-Jones, J. E. (1924) *Proc. Roy. Soc. London A* **106**, 463.

Lenz, W. (1920) *Z. Physik* **21**, 613.

Lenz, W. (1929) *Z. Physik* **56**, 778.

Lewis, H. W. and Wannier, G. H. (1952) *Phys. Rev.* **88**, 682.

Lieb, E. H. and Mattis, D. C. (1966) *Mathematical Physics in One Dimension* (Academic Press, New York).

Lines, M. E. and Glass, A. M. (1977) *Principles and Applications of Ferroelectrics and Related Materials* (Clarendon Press, Oxford).

Liouville, J. (1838) *J. de Math.* **3**, 348.

London, F. (1938a) *Nature* **141**, 643.

London, F. (1938b) *Phys. Rev.* **54**, 947.

London, F. (1954) *Superfluids*, Vols. 1 and 2 (John Wiley, New York); reprinted by Dover Publications, New York (1964).

Lorentz, H. A. (1904-05) *Proc. Kon. Ned. Akad. Wet. Amsterdam* **7**, 438, 585, 684. See also (1909) *The Theory of Electrons* (Teubner, Leipzig), secs. 47–50; reprinted by Dover Publications, New York (1952).

Loschmidt, J. (1876) *Wien. Ber.* **73**, 139.

Loschmidt, J. (1877) *Wien. Ber.* **75**, 67.

Luck, J. M. (1985) *Phys. Rev. B* **31**, 3069.

Lüders, G. and Zumino, B. (1958) *Phys. Rev.* **110**, 1450.

Ma, S. -K. (1973) *Phys. Rev. A*7, 2712.

Ma, S. -K. (1976a) *Modern Theory of Critical Phenomena* (Benjamin/Cummings Publishing Co., Reading, Mass.).

Ma, S. -K. (1976b) *Phys. Rev. Lett.* **37**, 461.

Ma, S. -K. (1976c) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and M. S. Green (Academic Press, London), Vol. 6, pp. 249–292.

MacDonald, D. K. C. (1948–49) *Rep. Prog. Phys.* **12**, 56.

MacDonald, D. K. C. (1950) *Phil. Mag.* **41**, 814.

MacDonald, D. K. C. (1962) *Noise and Fluctuations: An Introduction* (John Wiley, New York).

MacDonald, D. K. C. (1963) *Introductory Statistical Mechanics for Physicists* (John Wiley, New York).

Majumdar, R. (1929) *Bull. Calcutta Math. Soc.* **21**, 107.

Mandel, L., Sudarshan, E. C. G. and Wolf, E. (1964) *Proc. Phys. Soc. London* **84**, 435.

Manoliu, A. and Kittel, C. (1979) *Am. J. Phys.* **47**, 678.

March, N. H. (1957) *Adv. in Phys.* **6**, 1.

Martin, P. C. and de Dominicis, C. (1957) *Phys. Rev.* **105**, 1417.

Matsubara, T. and Matsuda, H. (1956) *Prog. Theor. Phys. Japan* **16**, 416.

Maxwell, J. C. (1860) *Phil. Mag.* **19**, 19; **20**, 1.

Maxwell, J. C. (1867) *Trans. Roy. Soc. London* **157**, 49 (= (1868) *Phil. Mag.* **35**, 129, 185).

Maxwell, J. C. (1879) *Camb. Phil. Soc. Trans.* **12**, 547.

Mayer, J. E. *et al.* (1937) *J. Chem. Phys.* **5**, 67, 74.

Mayer, J. E. *et al.* (1938) *J. Chem. Phys.* **6**, 87, 101.

Mayer, J. E. *et al.* (1942) *J. Chem. Phys.* **10**, 629.

Mayer, J. E. *et al.* (1951) *J. Chem. Phys.* **19**, 1024.

Mayer, J. E. and Mayer, M. G. (1940) *Statistical Mechanics* (John Wiley, New York).

McCoy, B. M. and Wu, T. T. (1973) *The Two-dimensional Ising Model* (Harvard University Press).

Mehra, J. and Pathria, R. K. (1994) in *The Beat of a Different Drum: The Life and Science of Richard Feynman* by J. Mehra (Clarendon Press, Oxford), Chap. 17, pp. 348–391.

Mermin, N. D. and Wagner, H. (1966) *Phys. Rev. Lett.* **17**, 1133, 1307.

Milne, E. (1927) *Proc. Camb. Phil. Soc.* **23**, 794.

Mohling, F. (1961) *Phys. Rev.* **122**, 1062.

Montroll, E. (1963) *The Many-body Problem*, ed. J. K. Percus (Interscience Publishers, New York).

Montroll, E. (1964) *Applied Combinatorial Mathematics*, ed. E. F. Beckenbach (John Wiley, New York).

Morse, P. M. (1969) *Thermal Physics*, 2nd ed. (W. A. Benjamin, New York).

Mukamel, D. (1975) *Phys. Rev. Lett.* **34**, 481.

Nelson, D. R. (1983) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and J. L. Lebowitz (Academic Press, London), Vol. 7, pp. 1–99.

von Neumann, J. (1927) *Göttinger Nachr.* 1, 24, 273.

Niemeijer, Th. and van Leeuwen, J. M. J. (1976) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and M. S. Green (Academic Press, London), Vol. 6, pp. 425–505.

Nienhuis, B. (1987) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and J. L. Lebowitz (Academic Press, London), Vol. 11, pp. 1–53.

Nozières, P. (1964) *Theory of Interacting Fermi Systems* (W. A. Benjamin, New York).

Nyquist, H. (1927) *Phys. Rev.* **29**, 614.

Nyquist, H. (1928) *Phys. Rev.* **32**, 110.

Ono, S. (1951) *J. [illegible]. Phys.* **19**, 504.

Onsager, L. (1931) *Phys. Rev.* **37**, 405; **38**, 2265.

ONSAGER, L. (1944) *Phys. Rev.* **65**, 117.

ONSAGER, L. (1949) *Nuovo Cim.* **6** (Suppl. 2), 249, 261.

ORNSTEIN, L. S. and ZERNIKE, F. (1914) *Proc. Akad. Sci. Amsterdam* **17**, 793.

ORNSTEIN, L. S. and ZERNIKE, F. (1916) *Proc. Akad. Sci. Amsterdam* **18**, 1520; reproduced in *The Equilibrium Theory of Classical Fluids*, eds. A. L. FRISCH and J. L. LEBOWITZ (Benjamin, New York, 1964).

PAIS, A. and UHLENBECK, G. E. (1959) *Phys. Rev.* **116**, 250.

PATASHINSKII, A. Z. and POKROVSKII, V. L. (1966) *Soviet Physics JETP* **23**, 292.

PATHRIA, R. K. (1966) *Nuovo Cim. (Supp.), Ser. I*, **4**, 276.

PATHRIA, R. K. (1972) *Statistical Mechanics*, 1st ed. (Pergamon Press, Oxford).

PATHRIA, R. K. (1974) *The Theory of Relativity*, 2nd ed. (Pergamon Press, Oxford).

PATHRIA, R. K. and KAWATRA, M. P. (1962) *Prog. Theor. Phys. Japan* **27**, 638, 1085.

PAULI, W. (1925) *Z. Physik* **31**, 776.

PAULI, W. (1927) *Z. Physik* **41**, 81.

PAULI, W. (1940) *Phys. Rev.* **58**, 716.

PAULI, W. and BELINFANTE, F. J. (1940) *Physica* **7**, 177.

PEIERLS, R. E. (1936) *Proc. Camb. Phil. Soc.* **32**, 471, 477.

PENROSE, O. and ONSAGER, L. (1956) *Phys. Rev.* **104**, 576.

PERCUS, J. K. (1963) *The Many-body Problem* (Interscience Publishers, New York).

PINES, D. (1962) *The Many-body Problem* (W. A. BENJAMIN, New York).

PITAEVSKII, L. P. (1959) *Soviet Phys. JETP* **9**, 830.

PITAEVSKII, L. P. (1961) *Soviet Phys. JETP* **13**, 451.

PLANCK, M. (1900) *Verhandl. Deut. Physik Ges.* **2**, 202, 237.

PLANCK, M. (1908) *Ann. Phys.* **26**, 1.

PLANCK, M. (1917) *Sitz. der Preuss. Akad.*, p. 324.

PLISCHKE, M. and BERGERSEN, B. (1989) *Equilibrium Statistical Physics* (Prentice Hall, Englewood Cliffs, New Jersey, U.S.A.).

POSPIŠIL, W. (1927) *Ann. Phys.* **83**, 735.

POTTS, R. B. and WARD, J. C. (1955) *Prog. Theor. Phys. Japan* **13**, 38.

PRIGOGINE, I. (1967) *Introduction to Thermodynamics of Irreversible Processes*, 3rd ed. (John Wiley, New York).

PRIVMAN, V. (ed.) (1990) *Finite Size Scaling and Numerical Simulation of Statistical Systems* (World Scientific, Singapore).

PRIVMAN, V. and FISHER, M. E. (1984) *Phys. Rev. B* **30**, 322.

PRIVMAN, V. HOHENBERG, P. C. and AHARONY, A. (1991) in *Phase Transitions and Critical Phenomena*, eds. C. DOMB and J. L. LEBOWITZ (Academic Press, London), Vol. 14, pp. 1–134, 364–367.

PURCELL, E. M. (1956) *Nature* **178**, 1449.

PURCELL, E. M. and POUND, R. V. (1951) *Phys. Rev.* **81**, 279.

RAMSEY, N. F. (1956) *Phys. Rev.* **103**, 20.

RAYFIELD, G. W. and REIF, F. (1964) *Phys. Rev.* **136**, A1194; see also (1963) *Phys. Rev. Lett.* **11**, 305.

RAYLEIGH, LORD (1900) *Phil. Mag* **49**, 539.

REE, F. H. and HOOVER, W. G. (1964) *J. Chem. Phys.* **40**, 939.

REIF, F. (1965) *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, New York).

RHODES, P. (1950) *Proc. Roy. Soc. London A* **204**, 396.

RIECKE, E. (1898) *Ann. Phys.* **66**, 353, 545.

RIECKE, E. (1900) *Ann. Phys.* **2**, 835.

ROBERTS, T. R. and SYDORIAK, S. G. (1955) *Phys. Rev.* **98**, 1672.

ROBINSON, J. E. (1951) *Phys. Rev.* **83**, 678.

ROWLINSON, J. S. and SWINTON, F. L. (1982) *Liquids and Liquid Mixtures*, 3rd ed. (Butterworth Scientific, London).

RUSHBROOKE, G. S. (1938) *Proc. Roy. Soc. London A* **166**, 296.

RUSHBROOKE, G. S. (1955) *Introduction to Statistical Mechanics* (Clarendon Press, Oxford).

RUSHBROOKE, G. S. (1963) *J. Chem. Phys.* **39**, 842.

RUSHBROOKE, G. S., BAKER, G. A. JR., and WOOD, P. J. (1974) in *Phase Transitions and Critical Phenomena*, eds. C. DOMB and M. S. GREEN (Academic Press, London), Vol. 3, pp. 245–356.

SACKUR, O. (1911) *Ann. Phys.* **36**, 958.

SACKUR, O. (1912) *Ann. Phys.* **40**, 67.

SCHIFF, L. I. (1968) *Quantum Mechanics*, 3rd ed. (McGraw-Hill, New York).

SCHRÖDINGER, E. (1960) *Statistical Thermodynamics* (Cambridge University Press).

SCHULTZ, T. D., LIEB, E. H. and MATTIS, D. C. (1964) *Rev. Mod. Phys.* **36**, 856.

SCOTT, G. G. (1951) *Phys. Rev.* **82**, 542.

SCOTT, G. G. (1952) *Phys. Rev.* **87**, 697.

SHANNON, C. E. (1948) *Bell Syst. Tech. J.* **27**, 379, 623.

Shannon, C. E 49) *The Mathematical Theory of Communication* (University of Illinois Press, Urbana).
Simon, F. (1930) *Ergeb. Exakt. Naturwiss.* **9**, 222.
Singh, S. and Pandita, P. N. (1983) *Phys. Rev. A* **28**, 1752.
Singh, S. and Pathria, R. K. (1985a) *Phys. Rev. B* **31**, 4483.
Singh, S. and Pathria, R. K. (1985b) *Phys. Rev. Lett.* **55**, 347.
Singh, S. and Pathria, R. K. (1986a) *Phys. Rev. Lett.* **56**, 2226.
Singh, S. and Pathria, R. K. (1986b) *Phys. Rev. B* **34**, 2045.
Singh, S. and Pathria, R. K. (1987a) *Phys. Rev. B* **36**, 3769.
Singh, S. and Pathria, R. K. (1987b) *J. Phys. A* **20**, 6357.
Singh, S. and Pathria, R. K. (1988) *Phys. Rev. B* **38**, 2740.
Singh, S. and Pathria, R. K. (1992) *Phys. Rev. B* **45**, 9759.
Smith, B. L. (1969) *Contemp. Phys.* **10**, 305.
Smoluchowski, M. v. (1906) *Ann. Phys.* **21**, 756.
Smoluchowski, M. v. (1908) *Ann. Phys.* **25**, 205.
Sokal, A. D. (1981) *J. Stat. Phys.* **25**, 25, 51.
Sommerfeld, A. (1928) *Z. Physik* **47**, 1; see also Sommerfeld and Frank (1931) *Rev. Mod. Phys.* **3**, 1.
Sommerfeld, A. (1932) *Z. Physik* **78**, 283.
Sommerfeld, A. (1956) *Thermodynamics and Statistical Mechanics* (Academic Press, New York).
Stanley, H. E. (1968) *Phys. Rev.* **176**, 718.
Stanley, H. E. (1969a) *J. Phys. Soc. Japan* **26S**, 102.
Stanley, H. E. (1969b) *Phys. Rev.* **179**, 570.
Stanley, H. E. (1971) *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press).
Stanley, H. E. (1974) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and M. S. Green (Academic Press, London), Vol. 3, pp. 485–567.
Stefan, J. (1879) *Wien. Ber.* **79**, 391.
Stierstadt, K. *et al.* (1990) *Physics Data: Experimental Values of Critical Exponents and Amplitude Ratios of Magnetic Phase Transitions* (Fachinformationszentrum, Karlsruhe, Germany).
Tarko, H. B. and Fisher, M. E. (1975) *Phys. Rev. B* **11**, 1217.
Tetrode, H. (1912) *Ann. Phys.* **38**, 434; corrections to this paper appear in (1912) *ibid.* **39**, 255.
Tetrode, H. (1915) *Proc. Kon. Ned. Akad. Amsterdam* **17**, 1167.
Thomas, L. H. (1927) *Proc. Camb. Phil. Soc.* **23**, 542.
Thompson, C. J. (1972a) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and M. S. Green (Academic Press, London), Vol. 1, pp. 177–226.
Thompson, C. J. (1972b) *Mathematical Statistical Mechanics* (Princeton University Press).
Thompson, C. J. (1988) *Classical Equilibrium Statistical Mechanics* (Clarendon Press, Oxford).
Tilley, D. R. and Tilley, J. (1990) *Superfluidity and Superconductivity*, 3rd ed. (Adam Hilger, Bristol).
Tisza, L. (1938a) *Nature* **141**, 913.
Tisza, L. (1938b) *Compt. Rend. Paris* **207**, 1035, 1186.
Tolman, R. C. (1934) *Relativity, Thermodynamics and Cosmology* (Clarendon Press, Oxford).
Tolman, R. C. (1938) *The Principles of Statistical Mechanics* (Oxford University Press).
Tracy, C. A. and McCoy, B. M. (1973) *Phys. Rev. Lett.* **31**, 1500.
Tuttle, E. R. and Mohling, F. (1966) *Annals of Physics* **38**, 510.
Uhlenbeck, G. E. (1966) in *Critical Phenomena*, eds. M. S. Green and J. V. Sengers (National Bureau of Standards, Washington), pp. 3–6.
Uhlenbeck, G. E. and Beth, E. (1936) *Physica* **3**, 729; see also (1937) *Physica* **4**, 915.
Uhlenbeck, G. E. and Ford, G. W. (1963) *Lectures in Statistical Mechanics* (American Mathematical Society, Providence).
Uhlenbeck, G. E. and Gropper, L. (1932) *Phys. Rev.* **41**, 79.
Uhlenbeck, G. E. and Ornstein, L. S. (1930) *Phys. Rev.* **36**, 823.
Ursell, H. D. (1927) *Proc. Camb. Phil. Soc.* **23**, 685.
Vdovichenko, N. V. (1965) *Soviet Phys. JETP* **20**, 477; **21**, 350.
Vinen, W. F. (1961) *Proc. Roy. Soc. London A* **260**, 218; see also *Progress in Low Temperature Physics*, Vol. III (1961), ed. C. J. Gorter (North-Holland Publishing Co., Amsterdam).
Vinen, W. F. (1968) *Rep. Prog. Phys.* **31**, 61.
Vollhardt, D. and Wölfle, P. (1990) *The Superfluid Phases of Helium 3* (Taylor and Francis, London).
Voronel, A. V. (1976) in *Phase Transitions and Critical Phenomena*, eds. C. Domb and M. S. Green (Academic Press, London), Vol. 5b, pp. 343–394.
van der Waals, J. D. (1873) *Over de Continuiteit van den Gas-en Vloeistoftoestand*, Thesis, Leiden.
Walker, C. B. (1956) *Phys. Rev.* **103**, 547.

WALLACE, D. J. (1976) in *Phase Transitions and Critical Phenomena*, eds. C. DOMB and M. S. GREEN (Academic Press, London), Vol. 6, pp. 293–356.
WALTON, A. J. (1969) *Contemp. Phys.* **10**, 181.
WANNIER, G. H. (1945) *Rev. Mod. Phys.* **17**, 50.
WANNIER, G. H. (1966) *Statistical Physics* (John Wiley, New York).
WATERSTON, J. J. (1892) *Phil. Trans. Roy. Soc. London A* **183**, 5, 79; reprinted in his collected papers (1928), ed. J. S. HALDANE (Oliver & Boyd, Edinburgh), pp. 207, 318. An abstract of Waterston's work did appear earlier; see *Proc. Roy. Soc. London* **5**, 604 (1846).
WATSON, P. G. (1969) *J. Phys. C* **2**, 1883, 2158.
WAX, N. (ed.) (1954) *Selected Papers on Noise and Stochastic Processes* (Dover Publications, New York).
WEINSTOCK, R. (1969) *Am. J. Phys.* **37**, 1273.
WELLER, W. (1963) *Z. Naturforsch.* **18A**, 79.
WERGELAND, H. (1969) *Lettere al Nuovo Cimento* **1**, 49.
WHITMORE, S. C. and ZIMMERMANN, W. (1965) *Phys. Rev. Lett.* **15**, 389.
WIDOM, B. (1965) *J. Chem. Phys.* **43**, 3892, 3898.
WIEBES, NIELS-HAKKENBERG and KRAMERS (1957) *Physica* **23**, 625.
WIEN, W. (1896) *Ann. Phys.* **58**, 662.
WIENER, N. (1930) *Act. Math. Stockholm* **55**, 117.
WILKS, J. (1961) *The Third Law of Thermodynamics* (Oxford University Press).
WILSON, A. H. (1960) *Thermodynamics and Statistical Mechanics* (Cambridge University Press).
WILSON, K. G. (1971) *Phys. Rev. B* **4**, 3174, 3184.
WILSON, K. G. (1972) *Phys. Rev. Lett.* **28**, 548.
WILSON, K. G. (1975) *Rev. Mod. Phys.* **47**, 773.
WILSON, K. G. and FISHER, M. E. (1972) *Phys. Rev. Lett.* **28**, 240.
WOODS, A. D. B. (1966) *Quantum Fluids*, ed. D. F. BREWER (North-Holland Publishing Co., Amsterdam), p. 239.
WU, F. Y. (1982) *Rev. Mod. Phys.* **54**, 235.
WU, T. T. (1959) *Phys. Rev.* **115**, 1390.
WU, T. T. *et al.* (1976) *Phys. Rev. B* **13**, 316.
YANG, C. N. (1952) *Phys. Rev.* **85**, 808.
YANG, C. N. and LEE, T. D. (1952) *Phys. Rev.* **87**, 404; see also *ibid.*, p 410.
YANG, C. N. and YANG, C. P. (1964) *Phys. Rev. Lett.* **13**, 303
YANG, C. P. (1962) *J. Math. Phys.* **3**, 797.
YARNELL, J. L. *et al.* (1959) *Phys. Rev.* **113**, 1379, 1386.
YEOMANS, J. M. (1992) *Statistical Mechanics of Phase Transitions* (Clarendon Press, Oxford).
ZASADA, C. S. and PATHRIA, R. K. (1976) *Phys. Rev. A* **14**, 1269.
ZERMELO, E. (1896) *Ann. Phys.* **57**, 485; **59**, 793.

INDEX