International Series of Monographs in
Natural Philosophy, Volume 45

Statistical Mechanics

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

Pergamon Press

INTERNATIONAL SERIES OF MONOGRAPHS IN

NATURAL PHILOSOPHY

GENERAL EDITOR: D. TER HAAR

VOLUME 45

STATISTICAL MECHANICS

OTHER TITLES IN THE SERIES IN NATURAL PHILOSOPHY

Vol. 1. DAVIDOV—Quantum Mechanics
Vol. 2. FOKKER—Time and Space, Weight and Inertia
Vol. 3. KAPLAN—Interstellar Gas Dynamics
Vol. 4. ABRIKOSOV, GOR'KOV and DZYALOSHINSKII—Quantum Field Theoretical Methods in Statistical Physics
Vol. 5. OKUN'—Weak Interaction of Elementary Particles
Vol. 6. SHKLOVSKII—Physics of the Solar Corona
Vol. 7. AKHIEZER *et al.*—Collective Oscillations in a Plasma
Vol. 8. KIRZHNITS—Field Theoretical Methods in Many-body Systems
Vol. 9. KLIMONTOVICH—The Statistical Theory of Non-equilibrium Processes in a Plasma
Vol. 10. KURTH—Introduction to Stellar Statistics
Vol. 11. CHALMERS—Atmospheric Electricity (2nd Edition)
Vol. 12. RENNER—Current Algebras and their Applications
Vol. 13. FAIN and KHANIN—Quantum Electronics, Volume 1 – Basic Theory
Vol. 14. FAIN and KHANIN—Quantum Electronics, Volume 2 – Maser Amplifiers and Oscillators
Vol. 15. MARCH—Liquid Metals
Vol. 16. HORI—Spectral Properties of Disordered Chains and Lattices
Vol. 17. SAINT JAMES, THOMAS and SARMA—Type II Superconductivity
Vol. 18. MARGENAU and KESTNER—Theory of Intermolecular Forces
Vol. 19. JANCEL—Foundations of Classical and Quantum Statistical Mechanics
Vol. 20. TAKAHASHI—An Introduction to Field Quantization
Vol. 21. YVON—Correlations and Entropy in Classical Statistical Mechanics
Vol. 22. PENROSE—Foundations of Statistical Mechanics
Vol. 23. VISCONTI—Quantum Field Theory, Volume 1
Vol. 24. FURTH—Fundamental Principles of Theoretical Physics
Vol. 25. ZHELEZNYAKOV—Radioemission of the Sun and Planets
Vol. 26. GRINDLAY—An Introduction to the Phenomenological Theory of Ferroelectricity
Vol. 27. UNGER—Introduction to Quantum Electronics
Vol. 28. KOGA—Introduction to Kinetic Theory Stochastic Processes in Gaseous Systems
Vol. 29. GALASIEWICZ—Superconductivity and Quantum Fluids
Vol. 30. CONSTANTINESCU and MAGYARI—Problems in Quantum Mechanics
Vol. 31. KOTKIN and SERBO—Collection of Problems in Classical Mechanics
Vol. 32. PANCHEV—Random Functions and Turbulence
Vol. 33. TALPE—Theory of Experiments in Paramagnetic Resonance
Vol. 34. TER HAAR—Elements of Hamiltonian Mechanics (2nd Edition)
Vol. 35. CLARK and GRAINGER—Polarized Light and Optical Measurement
Vol. 36. HAUG—Theoretical Solid State Physics, Volume 1
Vol. 37. JORDAN and BEER—The Expanding Earth
Vol. 38. TODOROV—Analytic Properties of Feynman Diagrams in Quantum Field Theory
Vol. 39. SITENKO—Lectures in Scattering Theory
Vol. 40. SOBEL'MAN—Introduction to the Theory of Atomic Spectra
Vol. 41. ARMSTRONG and NICHOLLS—Emission, Absorption and Transfer of Radiation in Heated Atmospheres
Vol. 42. BRUSH—Kinetic Theory, Volume 3
Vol. 43. BOGOLYUBOV—A Method for Studying Model Hamiltonians
Vol. 44. TSYTOVICH—An Introduction to the Theory of Plasma Turbulence

STATISTICAL MECHANICS

by

R. K. PATHRIA

Department of Physics,
University of Waterloo, Ontario, Canada

PERGAMON PRESS

Oxford · New York · Toronto
Sydney · Braunschweig

Pergamon Press Ltd., Headington Hill Hall, Oxford
Pergamon Press Inc., Maxwell House, Fairview Park, Elmsford, New York 10523
Pergamon of Canada Ltd., 207 Queen's Quay West, Toronto 1
Pergamon Press (Aust.) Pty. Ltd., 19a Boundary Street, Rushcutters Bay, N.S.W. 2011, Australia
Vieweg & Sohn GmbH, Burgplatz 1, Braunschweig

First edition 1972
Library of Congress Catalog Card No. 73-181690

Printed in Hungary

08 016747 0

TO

PROFESSORS F. C. AULUCK AND D. S. KOTHARI

who initiated me into the study of this subject

This page intentionally left blank

CONTENTS

PREFACE

THIS book has arisen out of the notes of lectures that I have given 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 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.

The 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. Hopefully, this will make the book useful for a wider readership.

Waterloo, Ontario, Canada R. K. P.

ACKNOWLEDGMENTS

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 which 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.

Waterloo, Ontario, Canada R. K. P.

This page intentionally left blank

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 *nonequilibrium* 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 ways, attempted to lay 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 the molecular bombardment on the walls of the container. Thus, Bernoulli and Herapath could show that, if the 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*, 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.* 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 the molecules. Krönig (1856) went a step further. Introducing the "quasi-statistical" assumption that, *at any time t*, one-sixth of the gas molecules could be assumed to be flying in

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

each of the six "independent" directions, namely $+x$, $-x$, $+y$, $-y$, $+z$ and $-z$, he derived the equation

$$P = \tfrac{1}{3}nmc^2, \tag{1}$$

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 the molecules; of course, from (1), he concluded 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 diffusion, viscosity or heat conduction, this is no longer the case.

It was at this stage that Clausius entered into the field. First of all, in 1857, he derived the ideal-gas law under assumptions far less stringent than Krönig's. He discarded both of the 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 the transport phenomena. It was in these studies that he introduced the famous "Stosszahlansatz"—the hypothesis on the number of collisions (among the molecules)—which had to play, later on, a prominent role in the monumental work of Boltzmann.* With Clausius, the introduction of the microscopic and statistical points of view into physical theory was definitive, rather than speculative. Accordingly, Maxwell, in a popular article entitled "Molecules", written for the *Encyclopedia Britannica*, has referred to him as the "principal founder of the kinetic theory of gases", while Gibbs, in his Clausius obituary notice, has called him the "father of statistical mechanics".†

The work of Clausius attracted Maxwell to the field. He made his first appearance with the great memoir "Illustrations in the dynamical theory of gases" (1860), in which he went much ahead of his predecessors by deriving his famous law of "distribution of molecular speeds". This derivation was based on the 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 skilfully used, leads to the same results as the ones following from the more fundamental equation due to Boltzmann.‡

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

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

‡ This has been demonstrated in Guggenheim (1960) where the coefficients of viscosity, thermal conductivity and diffusion of a gas of hard spheres have been calculated on the basis of Maxwell's transport equation.

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 the fundamental 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 of computing the entropy of a given physical system from a purely microscopic standpoint. 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 into 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 the macroscopic properties of systems in nonsteady states.

Things, however, proved quite harsh for Boltzmann. His H-theorem, and the consequent *irreversible* character of physical systems, came under heavy attack, mainly from Loschmidt (1876–7) 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,* while Lord Kelvin was emphasizing the "nineteenth-century clouds hovering over the dynamical theory of light and heat".†

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

> 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 the contemporary opposition were to cause the kinetic theory to sink into the oblivion which was the fate suffered by

---

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

† 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.

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

§ Quotation from Montroll (1963).

the wave theory of light through the authority of Newton. I am aware of the weakness of an 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 the 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 consider the development of the more sophisticated approach known as the *ensemble theory*, which may in fact be regarded as the statistical mechanics proper.* In this approach, the dynamical state of a given system, as characterized by the generalized coordinates q_i and the generalized momenta p_i of the system, 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, the "geometry" of the trajectory being governed by the equations of motion of the system and by the nature of the physical constraints acting on the system. To evolve an appropriate formalism, one considers the given system along with an infinitely large number of "mental copies" thereof, i.e. an *ensemble* of 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 in this space 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). However, eight years earlier, Boltzmann (1871) had already worked with essentially the same kind of ensembles.

The most important quantity in the ensemble theory is the *density function*, $\varrho(q, p; t)$, of the G-points in the phase space; a stationary distribution ($\partial\varrho/\partial t = 0$) characterizes a *stationary ensemble*, which in turn represents a system *in equilibrium*. Maxwell and Boltzmann confined their study to those ensembles for which the function ϱ depended solely on the energy E of the system. This included the special case of the *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*, $\bar{f}$, pertaining to *any* member of the ensemble. Now, $\bar{f}$ 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 which enables us to bring about a direct contact between the theory and the 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, by his *Elementary Principles of Statistical Mechanics* (1902), turned the ensemble theory into a most efficient tool for the theorist. He emphasized the use of "generalized" ensembles and developed

* For a review of the historical development of kinetic theory leading to statistical mechanics, see Brush (1957, 1958, 1961, 1965–6).

schemes which, in principle, enabled one to compute the complete set of thermodynamic quantities of a given system from purely mechanical properties of its microscopic constituents.* 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, namely mechanics, electrodynamics and thermodynamics, were all focused. At the same time, it disclosed all the strengths and weaknesses of these disciplines. It would have been strange 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 (1923) 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.† It was then natural that someone tried to derive Planck's radiation formula by treating the 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 was provided by the manner in which Planck's formula was derived by Bose. In his historic paper of 1924, Bose treated the 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 this paper into German from an English manuscript sent to him by Bose, 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, the photons

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

† 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).

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).* 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.† In the same paper Einstein discovered the famous phenomenon of *Bose–Einstein condensation* which, thirteen years later, was adopted by London (1938) as the basis for a microscopic understanding of the curious properties of liquid helium 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 leads to the Fermi–Dirac distribution as well, provided that one limits the occupancy of an energy state to *at most* one particle!‡

Soon after its appearance, the Fermi–Dirac statistics were applied, by Fowler (1926), to discuss the equilibrium states of the white dwarf stars and, by Pauli (1927), to explain the weak, temperature-independent paramagnetism of the alkali metals; in each case, one had to deal with a "highly degenerate" gas of electrons. 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 (1904–5). 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 has been considerably extended so as to be applicable to molecules, solids and nuclei as well.§

Thus, the whole structure of statistical mechanics was overhauled by the introduction of the concept of indistinguishability of (identical) particles.‖ The statistical aspect of the problem, which was already there in view of the large number of particles present, was

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

† 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. $n_i = 0, 1, 2, \ldots$.

‡ 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.

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

‖ Of course, in many a situation, especially at high temperatures and low densities, where the wave nature of the particles is not very important the classical statistics are still applicable.

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. This was carried out step by step. First of all, Landau (1927) and von Neumann (1927) introduced the so-called *density matrix*, which was meant to be 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).*

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.† 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 the application of the basic formalism to actual physical problems more fruitful. A review of these developments is clearly out of place here; we better discuss them at their appropriate place in the text.

* A detailed treatment of this development has been given by Kramers (1938).
† See also Lüders and Zumino (1958).

This page intentionally left blank

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.* 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 developed more on speculation than on calculation, began to emerge as a real, mathematical theory. Its initial successes were indeed 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 the one hand and the dynamics of the molecules 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 should, in some measure, 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.

* 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.

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 the order of 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 is generally denoted by the symbol n and is referred to as the *particle density*, 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 of the system. If the particles comprising the system could be regarded as noninteracting, the total energy E of the system 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 with energy ε_i. Clearly,

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

According to quantum mechanics, the possible values of the single-particle energies ε_i are discrete and their magnitude depends 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 spacings of the different energy values are so small in comparison with the total energy of the system that the parameter E might be regarded as almost a *continuous* variable. This would be true even if the particles were mutually interacting; however, 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 particular *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 noninteracting 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's 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 particular *microstate*, or *complexion*, of the given system. In general, the various microstates, or complexions, of a given system could be identified with the independent solutions $\psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ 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 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 a function of this parameter.* Curiously enough, it is from the magnitude of the number Ω, and from the nature of its dependence on the parameters N, V and E, that the 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 when 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 here the problem of the "thermal contact" between two 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 $\Omega(N, V, E)$

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 complexions, and the macrostate of A_2 be represented

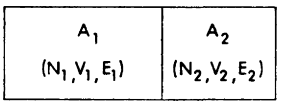

FIG. 1.1. Two physical systems, being brought into thermal contact.

by the parameters N_2, V_2 and E_2, so that it has $\Omega_2(N_2, V_2, E_2)$ possible complexions. The mathematical form of the function Ω_1 need not be the same as that of the function Ω_2, because that ultimately depends upon the nature of the system. We, of course, believe that all thermodynamic properties of the systems A_1 and A_2 can be completely derived from the respective functions $\Omega_1(N_1, V_1, E_1)$ and $\Omega_2(N_2, V_2, E_2)$.

We now bring the two systems into thermal contact with each other, thus allowing the possibility of an 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 unchanged. The energies E_1 and E_2, however, become variable and the only condition that restricts their variation is

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

* It may be noted that the manner in which the ε_i's 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 1.8. 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 enter into the boundary conditions imposed on the wave functions of the system.

Here, $E^{(0)}$ denoted 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

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

Clearly, the number $\Omega^{(0)}$ itself varies with E_1. The question then arises: at what value of the variable E_1 will the composite system be in equilibrium? Or, in other words, at what stage of the energy exchange will A_1 and A_2 be in 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, naturally proceeds in a direction that enables it to assume an ever-increasing number of complexions until it finally settles down in a state that affords the *largest possible* number of complexions. Statistically speaking, we regard a state with a larger number of complexions as a more probable state, and the one with the largest number of complexions as the most probable. Detailed studies show that for a typical system the number of complexions pertaining to any state which departs even slightly from the most probable state is "orders of magnitude" smaller than the number pertaining to the most probable state. 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 $\bar{E}_1$ (and that of E_2 by $\bar{E}_2$), we obtain, on maximizing $\Omega^{(0)}$,

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

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

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

Thus, our condition for equilibrium reduces to the equality of the respective values of the parameter $\{\partial \ln \Omega(N, V, E)/\partial E\}_{N,V}$ for the sub-systems A_1 and A_2. We denote this parameter by the symbol β:

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

whence the condition for equilibrium becomes

$$\beta_1 = \beta_2. \quad (4)$$

We thus find that when two physical systems are brought into thermal contact, which enables an exchange of energy to take place between them, this exchange continues *until*

* It should be 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).

the equilibrium values $\bar{E}_1$ and $\bar{E}_2$ of the variables E_1 and E_2 are reached. Once these values are reached, there is no more 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.* It is then natural that the parameter β be regarded as the statistical analogue of the *thermodynamic temperature* T. To determine the precise relationship between β and T, we consider the same problem from the conventional, thermodynamic point of view.

When the sub-systems A_1 and A_2 were isolated, their individual entropies $S_1(N_1, V_1, E_1)$ and $S_2(N_2, V_2, E_2)$ were determined by their respective macrostates. When they came into thermal contact, an exchange of energy set in. Let us consider, during this period of exchange, the transfer of an energy ΔE from the sub-system A_1 to the sub-system A_2. The corresponding change in the entropy of the composite system $A^{(0)}$ would be

$$\Delta S^{(0)} = (\Delta S)_{A_1} + (\Delta S)_{A_2} = \left(\frac{\partial S_1}{\partial E_1}\right)_{N_1, V_1} (-\Delta E) + \left(\frac{\partial S_2}{\partial E_2}\right)_{N_2, V_2} (+\Delta E)$$

$$= \Delta E \left[\left(\frac{\partial S_2}{\partial E_2}\right)_{N_2, V_2} - \left(\frac{\partial S_1}{\partial E_1}\right)_{N_1, V_1} \right]. \tag{5}$$

According to the second law of thermodynamics, the total entropy $S^{(0)}$ must increase, or at best stay constant, in the transfer process. Hence, for the "natural" direction of energy flow to be from A_1 to A_2, i.e. for ΔE to be positive, we must have

$$\left(\frac{\partial S_2}{\partial E_2}\right)_{N_2, V_2} > \left(\frac{\partial S_1}{\partial E_1}\right)_{N_1, V_1}. \tag{6}$$

The foregoing condition is essentially the same as the empirical one, namely

$$T_1 > T_2, \tag{7}$$

provided we recall the thermodynamic formula

$$\left(\frac{\partial S}{\partial E}\right)_{N, V} = \frac{1}{T}. \tag{8}$$

It is clear that the state of equilibrium for the composite system $A^{(0)}$ will obtain when $\Delta S^{(0)} = 0$, i.e. when

$$T_1 = T_2. \tag{9}$$†

Comparing the respective conditions of equilibrium, viz. (4) and (9), we conclude that the results of the statistical approach are essentially equivalent to those of the conventional thermodynamic approach. A further comparison of the formulae (3) and (8) prompts us to draw the desired correspondence between the statistical and the thermodynamic para-

* 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 equilibrium.

† It may be pointed out here that the onset of equilibrium does not necessarily require: $\Delta E = 0$. In fact, when $T_1 = T_2$, the process of energy transfer becomes *reversible* and ΔE is then as likely to be positive as negative; of course, we must have: $\overline{\Delta E} = 0$.