STATISTICAL MECHANICS

HARPER'S CHEMISTRY SERIES
under the editorship
of Stuart A. Rice

STATISTICAL MECHANICS

Donald A. McQuarrie
Indiana University

HARPER & ROW
Publishers
New York Evanston San Francisco London

Sponsoring Editor: John A. Woods
Project Editor: Ralph Cato
Designer: T. R. Funderburk
Production Supervisor: Will C. Jomarrón
Compositor: Santype International Ltd.
Printer/Binder: Halliday Lithograph Corporation
Art Studio: Vantage Art, Inc.

Statistical Mechanics

Library of Congress Cataloging in Publication Data
McQuarrie, Donald Allan.
Statistical mechanics.

(Harper's chemistry series)
"Portions of this work were originally published under the title: Statistical thermodynamics."
Bibliography: p.
Includes index.
1. Statistical mechanics. 2. Statistical thermodynamics. I. Title.
QC174.8.M3 530.1'32 75-20499
ISBN 06-044366-9

CONTENTS

Preface xi

Preface to *Statistical Thermodynamics* xiii

Chapter 1 INTRODUCTION AND REVIEW 1

1–1 Introduction 1
1–2 Classical Mechanics 3
1–3 Quantum Mechanics 8
1–4 Thermodynamics 13
1–5 Mathematics 20

Chapter 2 THE CANONICAL ENSEMBLE 35

2–1 Ensemble Averages 35
2–2 Method of the Most Probable Distribution 37
2–3 The Evaluation of the Undetermined Multipliers, α and β 40
2–4 Thermodynamic Connection 44

Chapter 3 OTHER ENSEMBLES AND FLUCTUATIONS 51

3–1 Grand Canonical Ensemble 51
3–2 Other Ensembles 55
3–3 Fluctuations 57

Chapter 4 BOLTZMANN STATISTICS, FERMI-DIRAC STATISTICS, AND BOSE-EINSTEIN STATISTICS 68

4–1 The Special Case of Boltzmann Statistics 68
4–2 Fermi-Dirac and Bose-Einstein Statistics 73

Chapter 5 IDEAL MONATOMIC GAS 81

5–1 The Translational Partition Function 81
5–2 The Electronic and Nuclear Partition Functions 83
5–3 Thermodynamic Functions 85
5–4 A Digression on Atomic Term Symbols 87

Chapter 6 IDEAL DIATOMIC GAS 91

6–1 The Rigid Rotor–Harmonic Oscillator Approximation 91
6–2 The Vibrational Partition Function 96
6–3 The Rotational Partition Function of a Heteronuclear Diatomic Molecule 98
6–4 The Symmetry Requirement of the Total Wave Function of a Homonuclear Diatomic Molecule 101
6–5 The Rotational Partition Function of a Homonuclear Diatomic Molecule 104
6–6 Thermodynamic Functions 108

Chapter 7 CLASSICAL STATISTICAL MECHANICS 113

7–1 The Classical Partition Function 113
7–2 Phase Space and the Liouville Equation 117
7–3 Equipartition of Energy 121

Chapter 8 IDEAL POLYATOMIC GAS 129

8–1 The Vibrational Partition Function 130
8–2 The Rotational Partition Function 133
8–3 Thermodynamic Functions 136
8–4 Hindered Rotation 138

Chapter 9 CHEMICAL EQUILIBRIUM 142

9–1 The Equilibrium Constant in Terms of Partition Functions 142
9–2 Examples of the Calculation of Equilibrium Constants 144
9–3 Thermodynamic Tables 151

Chapter 10 QUANTUM STATISTICS 160

10–1 A Weakly Degenerate Ideal Fermi-Dirac Gas 162
10–2 A Strongly Degenerate Ideal Fermi-Dirac Gas 164
10–3 A Weakly Degenerate Ideal Bose-Einstein Gas 169
10–4 A Strongly Degenerate Ideal Bose-Einstein Gas 171
10–5 An Ideal Gas of Photons (Blackbody Radiation) 177
10–6 The Density Matrix 182
10–7 The Classical Limit from the Quantum Mechanical Expression for Q 185

Chapter 11 CRYSTALS 194

11–1 The Vibrational Spectrum of a Monatomic Crystal 194
11–2 The Einstein Theory of the Specific Heat of Crystals 197
11–3 The Debye Theory of the Heat Capacity of Crystals 200
11–4 Introduction to Lattice Dynamics 206
11–5 Phonons 212
11–6 Point Defects in Solids 214

Chapter 12 IMPERFECT GASES 222

12–1 The Virial Equation of State from the Grand Partition Function 224
12–2 Virial Coefficients in the Classical Limit 226
12–3 Second Virial Coefficient 233
12–4 Third Virial Coefficient 237
12–5 Higher Virial Coefficients for the Hard-Sphere Potential 239
12–6 Quantum Corrections to $B_2(T)$ 241
12–7 The Law of Corresponding States 243
12–8 Conclusion 245

Chapter 13 DISTRIBUTION FUNCTIONS IN CLASSICAL MONATOMIC LIQUIDS 254

13–1 Introduction 255
13–2 Distribution Functions 257
13–3 Relation of Thermodynamic Functions to $g(r)$ 261
13–4 The Kirkwood Integral Equation for $g(r)$ 264
13–5 The Direct Correlation Function 268
13–6 Density Expansions of the Various Distribution Functions 270
13–7 Derivation of Two Additional Integral Equations 274
13–8 Density Expansions of the Various Integral Equations 277
13–9 Comparisons of the Integral Equations to Experimental Data 279

Chapter 14 PERTURBATION THEORIES OF LIQUIDS 300

14–1 Statistical Mechanical Perturbation Theory 302
14–2 The van der Waals Equation 304
14–3 Several Perturbation Theories of Liquids 306

Chapter 15 SOLUTIONS OF STRONG ELECTROLYTES 326

15–1 The Debye–Hückel Theory 328
15–2 Some Statistical Mechanical Theories of Ionic Solutions 340

Chapter 16 KINETIC THEORY OF GASES AND MOLECULAR COLLISIONS 357

16–1 Elementary Kinetic Theory of Transport in Gases 358
16–2 Classical Mechanics and Molecular Collisions 365
16–3 Mean-Square Momentum Change During a Collision 370

Chapter 17 CONTINUUM MECHANICS 379

17–1 Derivation of the Continuity Equations 380
17–2 Some Applications of the Fundamental Equations of Continuum Mechanics 386
17–3 The Navier–Stokes Equation and Its Solution 391

Chapter 18 KINETIC THEORY OF GASES AND THE BOLTZMANN EQUATION 402

18–1 Phase Space and the Liouville Equation 402
18–2 Reduced Distribution Functions 405
18–3 Fluxes in Dilute Gases 406
18–4 The Boltzmann Equation 409
18–5 Some General Consequences of the Boltzmann Equation 411

Chapter 19 TRANSPORT PROCESSES IN DILUTE GASES 426

19–1 Outline of the Chapman–Enskog Method 426
19–2 Summary of Formulas 430
19–3 Transport Coefficients for Various Intermolecular Potentials 433
19–4 Extensions of the Boltzmann Equation 440

Chapter 20 THEORY OF BROWNIAN MOTION 452

20–1 The Langevin Equation 452
20–2 The Fokker–Planck Equation and the Chandrasekhar Equation 456

Chapter 21 THE TIME-CORRELATION FUNCTION FORMALISM, I 467

21–1 Absorption of Radiation 470
21–2 Classical Theory of Light Scattering 476
21–3 Raman Light Scattering 484
21–4 An Elementary Derivation of the Basic Formulas 489
21–5 Dielectric Relaxation 495
21–6 Time-Correlation Function Formalism of Molecular Spectroscopy 499
21–7 Derivation of the Basic Formulas from the Liouville Equation 507
21–8 Time-Correlation Function Expressions for the Thermal Transport Coefficients 512
21–9 Applications of the Time-Correlation Function Formulas for the Thermal Transport Coefficients 522

Chapter 22 THE TIME-CORRELATION FUNCTION FORMALISM, II 543

22–1 Inelastic Neutron Scattering 544
22–2 The Weiner–Khintchine Theorem 553
22–3 Laser Light Scattering 561
22–4 The Memory Function 572
22–5 Derivation of Thermal Transport Coefficients 579

Appendix A VALUES OF SOME PHYSICAL CONSTANTS AND ENERGY CONVERSION FACTORS 593

Appendix B FOURIER INTEGRALS AND THE DIRAC DELTA FUNCTION 595

Appendix C DEBYE HEAT CAPACITY FUNCTION 599

Appendix D HARD-SPHERE RADIAL DISTRIBUTION FUNCTION 600

Appendix E TABLES FOR THE m–6–8 POTENTIAL 604

Appendix F DERIVATION OF THE GOLDEN RULE OF PERTURBATION THEORY 608

Appendix G THE DIRAC BRA AND KET NOTATION 612

Appendix H THE HEISENBERG TIME-DEPENDENT REPRESENTATION 615

Appendix I THE POYNTING FLUX VECTOR 618

Appendix J THE RADIATION EMITTED BY AN OSCILLATING DIPOLE 622

Appendix K DIELECTRIC CONSTANT AND ABSORPTION 626

Index 631

PREFACE

Statistical Mechanics is the extended version of my earlier text, *Statistical Thermodynamics*, whose Preface is reprinted following these pages. The present volume is intended primarily for a two-semester course or for a second one-semester course in statistical mechanics. Whereas *Statistical Thermodynamics* deals principally with equilibrium systems whose particles are either independent or effectively independent, *Statistical Mechanics* treats equilibrium systems whose particles are strongly interacting as well as nonequilibrium systems. The first eleven chapters of *Statistical Thermodynamics* form also the first eleven chapters of *Statistical Mechanics.* Chapter 15, Imperfect Gases, becomes Chapter 12 of the present volume, and the next ten chapters, 13–22, are completely new. Chapter 13 deals with the radial distribution function approach to liquids, and Chapter 14 is a fairly detailed discussion of statistical mechanical perturbation theories of liquids. These theories were developed in the late 1960s and early 1970s and have brought the numerical calculation of the thermodynamic properties of simple dense fluids to a practical level. A number of the problems at the end of Chapter 14 require the student to calculate such properties and compare the results to experimental data. Chapter 15, on ionic solutions, is the last chapter on equilibrium systems. Section 15–2 is an introduction to the recent (1970s) advances in ionic solution theory that now allow one to calculate the thermodynamic properties of simple ionic solutions up to concentrations of 2 molar.

Chapters 16–22 treat systems that are not in equilibrium. Chapters 16 and 17 are meant to be somewhat of a review, although admittedly much of the material, particularly in Chapter 17, will be new. Nevertheless, these two chapters do serve as a background for the rest. Chapter 18 presents the rigorous kinetic theory of gases as formulated through the Boltzmann equation, the famous integro-differential equation whose solution gives the nonequilibrium distribution of a molecule in velocity space. The long-time or equilibrium solution of the Boltzmann equation is the well-known Maxwell–Boltzmann distribution (Chapter 7). Being an integro-differential equation, it is not surprising that its solution is fairly involved. We only outline the standard method of solution, called the Chapman–Enskog method, in Section 19–1, and the next two sections are a practical calculation of the transport properties of gases. In the last section of Chapter 19 we discuss Enskog's ad hoc extension of the Boltzmann equation to dense hard-sphere fluids. Chapter 20, which presents the Langevin equation and the Fokker–Planck equation, again is somewhat of a digression but does serve as a background to Chapters 21 and 22.

The 1950s saw the beginning of the development of a new approach to transport processes that has grown into one of the most active and fruitful areas of nonequilibrium statistical mechanics. This work was initiated by Green and Kubo, who showed that the phenomenological coefficients describing many transport processes and time-dependent phenomena in general could be written as integrals over a certain type of function called a time-correlation function. The time-correlation function associated with some particular process is in a sense the analog of the partition function for equilibrium systems. Although both are difficult to evaluate exactly, the appropriate properties of the system of interest can be formally expressed in terms of these functions, and they serve as basic starting points for computationally convenient approximations. Before the development of the time-correlation function formalism, there was no single unifying approach to nonequilibrium statistical mechanics such as Gibbs had given to equilibrium statistical mechanics.

Chapters 21 and 22, two long chapters, introduce the time-correlation function approach. We have chosen to introduce the time-correlation function formalism through the absorption of electromagnetic radiation by a system of molecules since the application is of general interest and the derivation of the key formulas is quite pedagogical and requires no special techniques. After presenting a similar application to light scattering, we then develop the formalism in a more general way and apply the general formalism to dielectric relaxation, thermal transport, neutron scattering, light scattering, and several others.

There are also eleven Appendixes. The first three are from *Statistical Thermodynamics*, but the rest are new and are meant to supplement the textual material. Of particular interest is Appendix D, which lists a FORTRAN program for the hard-sphere radial distribution. This program allows one to apply the equations of Chapter 14 (perturbation theories of liquids) with ease and is central to a new, practical approach to the theory of liquids.

As in *Statistical Thermodynamics*, the intention here is to present a readable introduction to the topics covered rather than a rigorous, formal development. In addition, a great number of problems is included at the end of each chapter in order either to increase the student's understanding of the material or to introduce him or her to selected extensions. Generally, this volume is an extension of *Statistical Thermodynamics* to more complex and more interesting systems, but the spirit and level of the two are comparable.

Donald A. McQuarrie

Preface to *Statistical Thermodynamics*

Statistical Thermodynamics represents the expansion of notes for a one-semester, first-year graduate course in statistical mechanics. The present volume comprises the first 15 chapters of a larger work entitled *Statistical Mechanics* that is to be published in 1976. There is intentionally more material in the present volume than should be necessary for a single semester; additional topics have been included in an attempt to recognize each instructor's personal prejudices. It is assumed that students have had a standard course in physical chemistry but have not previously had a serious course in quantum mechanics. The mathematical prerequisite of the book is, at most, a knowledge of linear differential equations with constant coefficients. All of the necessary mathematical techniques have been developed within the book itself, either in the main text, in problems, or in Appendix B.

There is no doubt that the term "statistical thermodynamics" or "statistical mechanics" either frightens or brings forth scorn from many people. To eradicate this attitude, a serious attempt has been made to write this book with the experimentalist in mind, whether he be instructor or student. There are extensive applications of the formal equations to experimental data both in the text itself and in the numerous applied problems. For example, a large number of figures and tables in the book compare theoretical predictions to experimental results for a variety of real systems. This has, I hope, not been done at the expense of a sensible degree of rigor or logical clarity; I have tried to achieve a balance in this respect. What I ask of the student is to make a serious effort to conceptualize and abstract for the first few chapters, because from then on he will find that practical results and rewards come quite rapidly and easily.

The first chapter is called a review, but no doubt much of the material will be new to many students. The topics included here are particularly those points of classical mechanics, quantum mechanics, and mathematics that play a central role in the development of statistical mechanics. Chapter 2 introduces the fundamental postulates and concepts of statistical mechanics with particular reference to the so-called canonical ensemble, in other words, to thermodynamic (macroscopic) systems with fixed volume, temperature, and number of particles. This chapter is fairly formal, but it concludes with problems that anticipate the utility of the formal equations derived in the chapter.

The approach used to develop satistical mechanics can be credited (at least in the textbook literature) to T. L. Hill's *Introduction to Statistical Thermodynamics* (1960). There are basically two reasons for this choice: It is the most appealing approach from both a logical and pedagogical point of view, but probably more to the point, I was a student of Hill shortly after the publication of his book and grew up statistical-mechanically with his approach.

Chapter 3 extends the formal development of Chapter 2, and Chapter 4 illustrates how the equations of Chapters 2 and 3 may be reduced to very practical formulas that are applicable to an amazing variety of systems. These applications, in fact, largely constitute the balance of this volume. The common feature of the systems discussed in Chapters 5 through 14 is this: Atoms or molecules making up these systems are either independent or "effectively" independent. This results in equations that are not only easy to use but lead to comparisons with experiment that are usually quite decent and rewarding.

The final chapter investigates imperfect gases, the simplest system in which atoms or molecules cannot be treated as independent or even effectively independent. Here the intermolecular potential begins to emerge as a major complicating factor. A principal goal of this chapter is the introduction of one of the more advanced techniques that can be used to treat complicated systems. Chapters 13 through 15 of the larger volume *Statistical Mechanics* (which might be covered in the first weeks of a second semester) treat other systems in which the intermolecular potential plays a central role. The inclusion of Chapter 15 on imperfect gases in the present volume is to assure the skeptical student that statistical mechanics is quite capable of treating general nonindependent systems—and possibly to entice him into a more challenging and much more exciting second semester when he can, in fifteen weeks, be brought close to a "literature level" in statistical mechanics.

Although this book has evolved from teaching a course in statistical mechanics for several years, much effort has been made to produce a book that is equally suitable for self-study. (I have also kept in mind the apprehensive instructor who is not a professional theoretical chemist.) To this purpose, there is a fairly extensive and specific bibliography available at the end of each chapter to direct either student or instructor to other treatments of the material covered. Obviously no book of this type is written in a vacuum, and the bibliography represents those sources that were particularly useful to me.

One of the most important and useful features of this book is the extensive collection of problems presented after each chapter. The only way to learn statistical mechanics, or any other physical chemical subject, is to work a great many problems. Consequently, there are over 450 problems in the book, ranging in difficulty from "plug in the number and get an answer" to those that require little writing but much ingenuity. A few of the problems fill in steps leading from one equation to another, while others extend material to research papers in the literature. Most of the problems, however, are intended as applications of the equations and concepts presented in chapters. It is primarily the problems that dictate the level of a course such as statistical mechanics, and the variety and number included here allow an instructor to present a straightforward application-oriented course that should be perfectly reasonable for seniors or a fairly demanding introductory course for theoretical graduate students.

D. A. M.

ACKNOWLEDGMENTS

I should like to acknowledge the people who have helped me in putting this book together: Donald DuPré and Stuart Rice for reviewing most of the manuscript; all my students, particularly Steve Brenner, Jim Ely, Dennis Isbister, Allen Nelson, and Bernard Pailthorpe, who had to suffer through several earlier versions of the final manuscript; Wilmer Olivares for reading all of the page proofs; and Lynn Johnson, who typed the final manuscript and is more than happy to be finished. After reading many prefaces recently, it would seem appropriate at this point to thank my wife, Carole, for "encouragement and understanding during the writing of this book." It so happens, however, that she never did encourage me nor was she especially understanding, but she actually did something far better. As a biochemistry graduate student, she took both semesters of the course here (*not* from me) and found out for herself that statistical mechanics can be a rich and exciting field that has important applications not only in chemistry and physics but in biology as well. I hope that this book starts its readers along the same route.

STATISTICAL MECHANICS

CHAPTER 1

INTRODUCTION AND REVIEW

1–1 INTRODUCTION

Statistical mechanics is that branch of physics which studies macroscopic systems from a microscopic or molecular point of view. The goal of statistical mechanics is the understanding and prediction of macroscopic phenomena and the calculation of macroscopic properties from the properties of the individual molecules making up the system.

Present-day research in statistical mechanics varies from mathematically sophisticated discussions of general theorems to almost empirical calculations based upon simple, but nevertheless useful, molecular models. An example of the first type of research is the investigation of the question of whether statistical mechanics, as it is formulated today, is even capable of predicting the existence of a first-order phase transition. General questions of this kind are by their nature mathematically involved and are generally beyond the level of this book. We shall, however, discuss such questions to some extent later on. On the other hand, for many scientists statistical mechanics merely provides a recipe or prescription which allows them to calculate the properties of the physical systems which they are studying.

The techniques of statistical mechanics have been used in attacking a wide variety of physical problems. A quick glance through this text will show that statistical mechanics has been applied to gases, liquids, solutions, electrolytic solutions, polymers, adsorption, metals, spectroscopy, transport theory, the helix-coil transition of DNA, the electrical properties of matter, and cell membranes, among others.

Statistical mechanics may be broadly classified into two parts, one dealing with systems in equilibrium and the other with systems not in equilibrium. The treatment of systems in equilibrium is usually referred to as *statistical thermodynamics*, since it forms a bridge between thermodynamics (often called classical thermodynamics) and molecular physics.

Thermodynamics provides us with mathematical relations between the various experimental properties of macroscopic systems in equilibrium. An example of such a thermodynamic relation is that between the molar heat capacities at constant pressure and at constant volume,

$$C_p - C_v = \left[p + \left(\frac{\partial E}{\partial V}\right)_T\right]\left(\frac{\partial V}{\partial T}\right)_p \tag{1-1}$$

Another, and one that we shall use in the next chapter, is

$$\left(\frac{\partial E}{\partial V}\right)_{N,T} - T\left(\frac{\partial p}{\partial T}\right)_{N,V} = -p \tag{1-2}$$

Note that thermodynamics provides connections between many properties, but does not supply information concerning the magnitude of any one. Neither does it attempt to base any relation on molecular models or interpretations. This, in fact, is both the power and weakness of thermodynamics. It is a general discipline which does not need to recognize or rely upon the existence of atoms and molecules. Its many relations would remain valid even if matter were continuous. In addition, there are many systems (such as biological systems) which are too complicated to be described by an acceptable molecular theory, but here again the relations given by thermodynamics are exact. This great generality, however, is paid for by its inability to calculate physical properties separately or to supply physical interpretations of its equations. When one seeks a molecular theory which can do just this, one then enters the field of statistical thermodynamics. Thus thermodynamics and statistical thermodynamics treat the same systems. Thermodynamics provides general relations without the need of ever considering the ultimate constitution of matter, while statistical thermodynamics, on the other hand, assumes the existence of atoms and molecules to calculate and interpret thermodynamic quantities from a molecular point of view.

Statistical thermodynamics itself may be further divided into two areas: first, the study of systems of molecules in which molecular interactions may be neglected (such as dilute gases), and second, the study of systems in which the molecular interactions are of prime importance (such as liquids). We shall see that the neglect of intermolecular interactions enormously simplifies our problem. Chapters 4 through 11 of the book are devoted to the treatment of systems in which these interactions either may be ignored or highly simplified. This is the kind of statistical thermodynamics to which most undergraduates have been exposed and, to some extent, represents typical statistical thermodynamical research done in the 1930s. The more interesting and challenging problems, however, concern systems in which these molecular interactions cannot be neglected; Chapters 12 through 15 of the present volume are devoted to the study of such systems. It is in this latter area that a great deal of the research of the 1940s, 1950s, and 1960s was carried out. There are, of course, many important problems of this sort still awaiting attack. The theory of concentrated electrolyte solutions and the proof for the existence of first-order phase transitions are just two examples.

The most difficult branch of statistical mechanics, both mathematically and conceptually, is the study of systems not in equilibrium. This field is often referred to as *nonequilibrium statistical mechanics*. This is presently a very active area of research.

There are still some important unsolved conceptual problems in nonequilibrium statistical mechanics. Nevertheless, in the 1950s great strides were made toward the establishment of a firm basis for nonequilibrium statistical mechanics, commensurate with that of equilibrium statistical mechanics, or what we have called statistical thermodynamics. Chapters 16 through 22 of *Statistical Mechanics* present an introduction to some of the more elementary of these fairly new and useful concepts and techniques.

In Chapter 2 we shall introduce and discuss the basic concepts and assumptions of statistical thermodynamics. We shall present these ideas in terms of quantum mechanical properties such as energy states, wave functions, and degeneracy. Although it may appear at this point that quantum mechanics is a prerequisite for statistical thermodynamics, it will turn out that a satisfactory version of statistical thermodynamics can be presented by using only a few quantum mechanical ideas and results. We assume that the student is familiar with only the amount of quantum mechanics taught in most present-day physical chemistry courses. About the only requirement of the first few chapters is an understanding that the Schrödinger equation determines the possible energy values E_j available to the system and that these may have a degeneracy associated with them which we denote by $\Omega(E_j)$.

Before discussing the principles, however, we shall present in this chapter a discussion of some of the terms or concepts that are particularly useful in statistical thermodynamics. In Section 1–2 we shall treat classical mechanics, including an introduction to the Lagrangian and Hamiltonian formalisms. In Section 1–3 we shall briefly review the main features of quantum mechanics and give the solutions of the Schrödinger equation for some important systems. The only new material in this section to most students probably will be the discussion of the eigenvalues or energy levels of a many-body system. Then in Section 1–4 we shall review thermodynamics briefly, since it is assumed that the reader is familiar with the three laws of thermodynamics and the tedious manipulations of partial derivatives. Two important topics that are not usually discussed in elementary physical chemistry texts are introduced, however. These two topics are the Legendre transformation and Euler's theorem, both of which are useful in studying statistical thermodynamics. Finally, in Section 1–5 we shall discuss some mathematical techniques and results that are particularly useful in statistical thermodynamics. Much of this section may be new material to the reader.

1–2 CLASSICAL MECHANICS

NEWTONIAN APPROACH

Everyone knows the equation $F = ma$. What this equation really says is that the rate of change of momentum is equal to the applied force. If we denote the momentum by $\mathbf{p}$, we have then a more general version of Newton's second law, namely,

$$\frac{d\mathbf{p}}{dt} \equiv \dot{\mathbf{p}} = \mathbf{F} \tag{1–3}$$

If the mass is independent of time, then $d\mathbf{p}/dt = m d\dot{\mathbf{r}}/dt = m\ddot{\mathbf{r}} = m\mathbf{a}$. If $\mathbf{F}$ is given as a function of position $\mathbf{F}(x, y, z)$, then Eq. (1–3) represents a set of second-order differential equations in x, y, and z whose solutions give x, y, and z as a function of time if some initial conditions are known. Thus Eq. (1–3) is called an equation of motion. We shall consider three applications of this equation.

Example 1. Solve the equation of motion of a body of mass m shot vertically upward with an initial velocity v_0 in a gravitational field.

If we choose the x-axis (positive in the upward direction) to be the height of the body, then we have

$$m\ddot{x} = -mg$$

where mg is the magnitude of the force. The negative sign indicates that the force is acting in a downward direction. The solution to this differential equation for x is then

$$x(t) = -\tfrac{1}{2}gt^2 + v_0 t + x_0 \tag{1–4}$$

with x_0 equal to $x(0)$, which in our case is 0. This then gives the position of the body at any time after it was projected. The extension of this problem to two dimensions (i.e., a shell shot out of a cannon) and the inclusion of viscous drag on the body are discussed in Problems 1–1 and 1–2.

Example 2. Set up and solve the equation of motion of a simple harmonic oscillator.

Let x_0 be the length of the unstrained spring. Hooke's law says that the force on the mass attached to the end of the spring is $F = -k(x - x_0)$. If we let $\xi = x - x_0$, we can write

$$\frac{d^2\xi}{dt^2} + \frac{k}{m}\xi = 0 \tag{1–5}$$

whose solution is

$$\xi(t) = A \sin \omega t + B \cos \omega t \tag{1–6}$$

The quantity

$$\omega = (k/m)^{1/2} \tag{1–7}$$

is the natural vibrational frequency of the system. Equation (1–6) can be written in an alternative form (see Problem 1–5)

$$\xi(t) = C \sin(\omega t + \phi) \tag{1–8}$$

This shows more clearly that the mass undergoes simple harmonic motion with frequency ω. Problems 1–3 through 1–5 illustrate some of the basic features of simple harmonic motion.

Example 3. Two-dimensional motion of a body under coulombic attraction to a fixed center.

In this case the force is $\mathbf{F} = -K\mathbf{r}/r^3$, that is, it is of magnitude $-K/r^2$ and directed radially. Newton's equations become

$$\begin{aligned} m\ddot{x} &= F_x = -\frac{Kx}{(x^2 + y^2)^{3/2}} \\ m\ddot{y} &= F_y = -\frac{Ky}{(x^2 + y^2)^{3/2}} \end{aligned} \tag{1–9}$$

Unlike our previous examples, these two equations are difficult to solve. Since the force depends, in a natural way, on the polar coordinates r and θ, it is more convenient for us to set the problem up in a polar coordinate system. Using then

$$\begin{aligned} x &= r \cos \theta \\ y &= r \sin \theta \end{aligned}$$

and some straightforward differentiation, we get

$$\left\{m(\ddot{r} - \dot{\theta}^2 r) + \frac{K}{r^2}\right\}\cos \theta - m(r\ddot{\theta} + 2\dot{\theta}\dot{r})\sin \theta = 0 \tag{1–10a}$$

$$\left\{m(\ddot{r} - \dot{\theta}^2 r) + \frac{K}{r^2}\right\}\sin \theta + m(r\ddot{\theta} + 2\dot{\theta}\dot{r})\cos \theta = 0 \tag{1–10b}$$

By multiplying the first of these equations by $\cos\theta$ and the second by $\sin\theta$ and then adding the two, one gets

$$m(\ddot{r} - \dot{\theta}^2 r) + \frac{K}{r^2} = 0 \tag{1–11}$$

But this is just the term in braces in Eq. (1–10), which leads us to the result that

$$m(r\ddot{\theta} + 2\dot{\theta}\dot{r}) = 0 \tag{1–12}$$

as well. Equation (1–12) can be written in the form

$$\frac{1}{r}\frac{d}{dt}(mr^2\dot{\theta}) = 0 \tag{1–13}$$

which implies that

$$mr^2\dot{\theta} = \text{constant} \tag{1–14}$$

This quantity, $mr^2\dot{\theta}$, which maintains a fixed value during the motion of the particle, is called the angular momentum of the particle and is denoted by l. The angular momentum is always conserved if the force is central, that is, directed along $\mathbf{r}$ (see Problem 1–10).

Equation (1–14) can be used to eliminate $\dot{\theta}$ from Eq. (1–11) to give an equation in r alone, called the radial equation:

$$m\ddot{r} - \frac{l^2}{mr^3} + \frac{K}{r^2} = 0 \tag{1–15}$$

This equation can be solved (at least numerically) to give $r(t)$, which together with Eq. (1–14) gives $\theta(t)$.

Even though the solution in this example is somewhat involved using polar coordinates, it is nevertheless much easier than if we had used Cartesian coordinates. This is just one example of many possibilities, which show that it is advantageous to recognize the symmetry of the problem by using the appropriate coordinate system.

This example was introduced, however, to illustrate another important point. Notice that Eq. (1–15) for r can be written as a Newtonian equation (i.e., in the form $F = ma$)

$$m\ddot{r} = -\frac{K}{r^2} + \frac{l^2}{mr^3}$$

if we interpret the term l^2/mr^3 as a force. This force is the well-known centrifugal force and must be introduced into the equation for $m\ddot{r}$.

This constitutes the main disadvantage of the Newtonian approach. The form of the equation $m\ddot{\eta} = F_\eta$ (where η is some general coordinate) is useful only in Cartesian systems, unless we are prepared to define additional forces, such as the centrifugal force in the above example. At times these necessary additional forces are fairly obscure.

There exist more convenient formulations of classical mechanics which are not tied to any one coordinate system. The two formulations that we are about to introduce are, in fact, independent of the coordinate system employed. These are the Lagrangian and the Hamiltonian formulations.

LAGRANGIAN APPROACH

Let K be the kinetic energy of a particle. In Cartesian coordinates

$$K(\dot{x}, \dot{y}, \dot{z}) = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

Let the potential energy be U. In many problems U is a function of position only, and so we write $U(x, y, z)$. Newton's equations are

$$m\ddot{x} = -\frac{\partial U}{\partial x}$$

with similar equations for y and z. Now introduce a new function

$$L(x, y, z, \dot{x}, \dot{y}, \dot{z}) \equiv K(\dot{x}, \dot{y}, \dot{z}) - U(x, y, z)$$

This function is called the Lagrangian of the system. In terms of L, we have

$$\frac{\partial L}{\partial \dot{x}} = \frac{\partial K}{\partial \dot{x}} = m\dot{x}$$

$$\frac{\partial L}{\partial x} = -\frac{\partial U}{\partial x}$$

and we can write Newton's equations in the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = \frac{\partial L}{\partial x} \tag{1-16}$$

with similar equations for y and z. These are Lagrange's equations of motion in Cartesian coordinates. The remarkable and useful property of Lagrange's equations is that they have the same form in any coordinate system. If the x, y, z are transformed into any other system, say q_1, q_2, q_3, Lagrange's equations take the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) = \frac{\partial L}{\partial q_j} \qquad j = 1, 2, 3 \tag{1-17}$$

This can be proved by writing $x = x(q_1, q_2, q_3)$, $y = y(q_1, q_2, q_3)$, and $z = z(q_1, q_2, q_3)$ and then transforming Eq. (1–16) into Eq. (1–17). (See Problem 1–13.)

Lagrange's equations are more useful than Newton's equations in many problems because it is usually much easier to write down an expression for the potential energy in some appropriate coordinate system than it is to recognize all the various forces. The Lagrangian formalism is based on the potential energy of the system, whereas the Newtonian approach is based on the forces acting on the system.

To illustrate the utility of the Lagrangian approach, we shall redo Example 3, the two-dimensional motion of a particle in a coulombic force field.

Example 3′. The kinetic energy is

$$K = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2)$$

and the potential energy is $U = -K/r$. The Lagrangian, then, is

$$L(r, \theta, \dot{r}, \dot{\theta}) = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{K}{r} \tag{1-18}$$

The two Lagrangian equations of motion are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) = \frac{\partial L}{\partial r}$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) = \frac{\partial L}{\partial \theta}$$

or using Eq. (1–18) for L,

$$\frac{d}{dt}(m\dot{r}) = mr\dot{\theta}^2 - \frac{K}{r^2}$$

$$\frac{d}{dt}(mr^2\dot{\theta}) = 0$$

These two equations are just Eqs. (1–11) and (1–13). Note, however, that they were obtained in a much more straightforward manner than were Eqs. (1–11) and (1–13). Problems 1–10 through 1–12 further illustrate the utility of the Lagrangian formulation. Other problems involve the motion of one and two particles in central force fields.

Equations (1–17) are three second-order ordinary differential equations. To completely specify the solutions, we need three initial velocities $\dot{q}_1(0)$, $\dot{q}_2(0)$, $\dot{q}_3(0)$, and three initial positions $q_1(0)$, $q_2(0)$, $q_3(0)$. These six initial conditions along with Lagrange's equations completely determine the future (and past) trajectory of the system. If there were N particles in the system, there would be $3N$ Lagrange equations and $6N$ initial conditions.

There is another formulation of classical mechanics that involves $6N$ first-order differential equations. Although this formulation is not as convenient as Lagrange's for solving problems, it is more convenient from a theoretical point of view, particularly in quantum mechanics and statistical mechanics. This is the Hamiltonian formulation.

HAMILTONIAN APPROACH

We define a generalized momentum by

$$p_j = \frac{\partial L}{\partial \dot{q}_j} \qquad j = 1, 2, \ldots, 3N \tag{1–19}$$

This generalized momentum is said to be *conjugate* to q_j. Note that Eq. (1–19) is simply $p_x = m\dot{x}$, and so on, in Cartesian coordinates.

We now define the Hamiltonian function for a system containing just one particle (for simplicity) by

$$H(p_1, p_2, p_3, q_1, q_2, q_3) = \sum_{j=1}^{3} p_j \dot{q}_j - L(\dot{q}_1, \dot{q}_2, \dot{q}_3, q_1, q_2, q_3) \tag{1–20}$$

It is understood here that the $\dot{q}_j$'s have been eliminated in favor of the p_j's by means of Eq. (1–19).

An important difference between the Lagrangian approach and the Hamiltonian approach is that the Lagrangian is considered to be a function of the generalized velocities $\dot{q}_j$ and the generalized coordinates q_j, whereas the Hamiltonian is considered to be a function of the generalized momenta p_j and the conjugate generalized coordinates q_j. This may appear to be a fine distinction at this point, but it will turn out to be important later on. It also may seem, at this time, that the definition Eq. (1-20) is rather obscure, but we shall give a motivation for its form in Section 1-4. (See Problem 1-38.)

For the kinds of systems that we shall treat in this book, the kinetic energy is of the form

$$K=\sum_{j=1}^{3N} a_j(q_1, q_2, \ldots, q_{3N})\dot{q}_j^{\,2} \tag{1-21}$$

that is, a quadratic function of the generalized velocities. The coefficients a_j are, in general, functions of generalized coordinates but not an explicit function of time. If, furthermore, the potential energy is a function only of the generalized coordinates, then the p_j occurring in Eq. (1-20) are given by

$$p_j = \frac{\partial L}{\partial \dot{q}_j} = \frac{\partial K}{\partial \dot{q}_j} = 2a_j \dot{q}_j$$

where the last equality comes from Eq. (1-21). Substituting this into Eq. (1-20) gives the important result

$$H = K + U = \text{total energy} \tag{1-22}$$

We shall now show that if the Lagrangian is not an explicit function of time, then $dH/dt = 0$. We begin with the definition of H, that is, Eq. (1-20).

$$dH = \sum_j \dot{q}_j \, dp_j + \sum_j p_j \, d\dot{q}_j - \sum \frac{\partial L}{\partial \dot{q}_j} d\dot{q}_j - \sum \frac{\partial L}{\partial q_j} dq_j$$

But if we use Eqs. (1-17) and (1-19), we see that

$$dH = \sum \dot{q}_j \, dp_j - \sum \dot{p}_j \, dq_j \tag{1-23}$$

The total derivative of H is (assuming no explicit dependence on time)

$$dH = \sum \left(\frac{\partial H}{\partial p_j}\right) dp_j + \sum \left(\frac{\partial H}{\partial q_j}\right) dq_j \tag{1-24}$$

Comparing Eqs. (1-23) and (1-24), we get Hamilton's equations of motion:

$$\frac{\partial H}{\partial p_j} = \dot{q}_j \qquad \frac{\partial H}{\partial q_j} = -\dot{p}_j \qquad j = 1, 2, \ldots, 3N \tag{1-25}$$

Hamilton's equations are $6N$ first-order differential equations. It is easy to show from Eqs. (1-24) and (1-25) that $dH/dt = 0$. (See Problem 1-14.) This along with Eq. (1-22) says that energy is conserved in such systems.

Since the Hamiltonian is so closely related to the energy, and it is the total energy which is usually the prime quantity in quantum and statistical mechanics, the Hamiltonian formalism will turn out to be the most useful from a conceptual point of view. Fortunately, however, we shall never have to solve the equations of motion for macroscopic systems. The role of statistical mechanics is to avoid doing just that.

1-3 QUANTUM MECHANICS

In the previous section we have seen that a knowledge of the initial velocities and coordinates of a particle or a system of particles was sufficient to determine the future course of the system if the equations of motion, essentially the potential field that the system experiences, are known. If the state (its velocities and coordinates) of the system is known at time t_0, then classical mechanics provides us with a method of calculating the state of the system at any other time t_1.

By the 1920s it was realized that such a calculation was too detailed in principle. The Heisenberg uncertainty principle states that it is impossible to precisely specify both the momentum and position of a particle simultaneously. Consequently, the prescription given by classical mechanics had to be modified to include the principle of uncertainty. This modification resulted in the development of quantum mechanics.

There are a number of levels of introducing the central ideas of quantum mechanics, but for most of the material in this text, we need consider only the most elementary. A fundamental concept of quantum mechanics is the so-called wave function $\Psi(\mathbf{q}, t)$, where $\mathbf{q}$ represents the set of coordinates necessary to describe the system. The wave function is given the physical interpretation that the probability that at time t the system is found between q_1 and $q_1 + dq_1$, q_2 and $q_2 + dq_2$, and so on, is

$$\Psi^*(\mathbf{q}, t)\Psi(\mathbf{q}, t)\, dq_1\, dq_2 \cdots dq_{3N}$$

We shall often write $dq_1 \cdots dq_{3N}$ as $d\mathbf{q}$. The uncertainty principle dictates that $\Psi(\mathbf{q}, t)$ is the most complete description of the system that can be obtained. Since the system is sure to be somewhere, we have

$$\int \Psi^*(\mathbf{q}, t)\Psi(\mathbf{q}, t)\, d\mathbf{q} = 1 \tag{1-26}$$

If Eq. (1–26) is satisfied, Ψ is said to be normalized.

A central problem of quantum mechanics is the calculation of $\Psi(\mathbf{q}, t)$ for any system of interest. We denote the time-independent part of $\Psi(\mathbf{q}, t)$ by $\psi(\mathbf{q})$. The state of the system described by a particular $\psi(\mathbf{q})$ is said to be a stationary state. Throughout this book we shall deal with stationary states only.

For our purpose, the wave function ψ is given as the solution of the Schrödinger equation

$$\mathscr{H}\psi = E\psi \tag{1-27}$$

where $\mathscr{H}$ is the Hamiltonian operator, and E is a scalar quantity corresponding to the energy of the system. The Hamiltonian operator is

$$\begin{aligned}\mathscr{H} &= -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + U(x, y, z)\\ &= -\frac{\hbar^2}{2m}\nabla^2 + U(x, y, z)\end{aligned} \tag{1-28}$$

where $\hbar$ is $h/2\pi$, that is, Planck's constant divided by 2π. The first term here corresponds to the kinetic energy, and the second term is the potential energy. The Hamiltonian operator, then, corresponds to the total energy. There is a quantum mechanical operator and an equation similar to Eq. (1–27) corresponding to every quantity of classical mechanics, but we shall need only the one for the energy, namely, the Schrödinger equation.

Given certain physical boundary conditions of the system, a knowledge of $\mathscr{H}$ alone is sufficient to determine ψ and E. The wave function ψ is called an eigenfunction of the operator $\mathscr{H}$, and E is called an eigenvalue. There will usually be many ψ's and E's that satisfy Eq. (1–28), and this is indicated by labeling ψ and E with one or more subscripts. Generally, then, we have

$$\mathscr{H}\psi_j = E_j\psi_j \tag{1-29}$$

Equation (1–29) is a partial differential equation for ψ_j. The application of the boundary conditions often limits the values of E_j to only certain discrete values. Some simple examples are

1. a particle in a one-dimensional infinite well:

$$\mathscr{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}$$

$$\varepsilon_n = \frac{h^2 n^2}{8ma^2} \qquad n = 1, 2, \ldots \tag{1–30}$$

2. a simple harmonic oscillator:

$$\mathscr{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2}kx^2$$

$$\varepsilon_n = (n + \tfrac{1}{2})\hbar\omega \qquad n = 0, 1, 2, \ldots \tag{1–31}$$

where $\omega = (k/m)^{1/2}$.

3. a rigid rotor (see Problem 1–21):

$$\mathscr{H} = -\frac{\hbar^2}{2I}\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right\}$$

$$\varepsilon_J = \frac{J(J+1)\hbar^2}{2I} \qquad J = 0, 1, 2, \ldots \tag{1–32}$$

Here I is the moment of inertia of the rotor (see Problem 1–15 for a treatment of the classical counterpart of this system).

The rigid rotor illustrates another important concept of quantum mechanics, namely, that of degeneracy. It happens that there may be a number of eigenfunctions or states of the system having the same eigenvalue or energy. The number of eigenfunctions having this energy is called the degeneracy of the system. For the rigid rotor, the degeneracy, ω_J, is $2J + 1$. The particle in a one-dimensional infinite well and the simple harmonic oscillator are nondegenerate, that is, the ω_n are unity. The concept of energy states and degeneracy plays an important role in statistical thermodynamics.

Consider the energy states of a particle in a three-dimensional infinite well. These are given by

$$\varepsilon_{n_x n_y n_z} = \frac{h^2}{8ma^2}(n_x^2 + n_y^2 + n_z^2) \qquad n_x, n_y, n_z = 1, 2, 3, \ldots \tag{1–33}$$

The degeneracy is given by the number of ways that the integer $M = 8ma^2\varepsilon/h^2$ can be written as the sum of the squares of three positive integers. In general, this is an erratic and discontinuous function of M (the number of ways will be zero for many values of M), but it becomes smooth for large M, and it is possible to derive a simple expression for it. Consider a three-dimensional space spanned by n_x, n_y, and n_z. There is a one-to-one correspondence between energy states given by Eq. (1–33) and the points in this n_x, n_y, n_z space with coordinates given by positive integers. Figure 1–1 shows a two-dimensional version of this space. Equation (1–33) is an equation for a sphere of radius $R = (8ma^2\varepsilon/h^2)^{1/2}$ in this space

$$n_x^2 + n_y^2 + n_z^2 = \frac{8ma^2\varepsilon}{h^2} = R^2$$

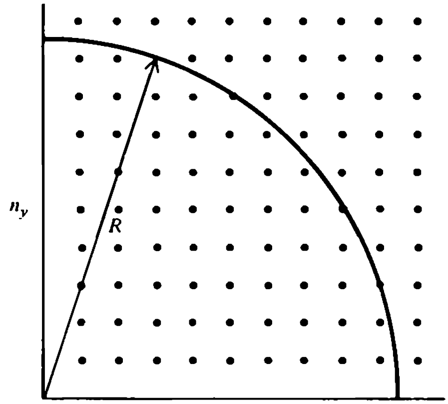

Figure 1-1. **A two-dimensional version of the (n_x, n_y, n_z) space, the space with the quantum numbers n_x, n_y, and n_z as axes.**

We wish to calculate the number of lattice points that are at some fixed distance from the origin in this space. In general, this is very difficult, but for large R we can proceed as follows. We treat R or ε as a continuous variable and ask for the number of lattice points between ε and $\varepsilon + \Delta\varepsilon$. To calculate this quantity, it is convenient to first calculate the number of lattice points consistent with an energy $\leq \varepsilon$. For large ε, it is an excellent approximation to equate the number of lattice points consistent with an energy $\leq \varepsilon$ with the volume of one octant of a sphere of radius R. We take only one octant, because n_x, n_y, and n_z are restricted to be positive integers. If we denote the number of such states by $\Phi(\varepsilon)$, we can write

$$\Phi(\varepsilon) = \frac{1}{8}\left(\frac{4\pi R^3}{3}\right) = \frac{\pi}{6}\left(\frac{8ma^2\varepsilon}{h^2}\right)^{3/2} \tag{1-34}$$

The number of states between ε and $\varepsilon + \Delta\varepsilon$ ($\Delta\varepsilon/\varepsilon \ll 1$) is

$$\begin{aligned}\omega(\varepsilon, \Delta\varepsilon) &= \Phi(\varepsilon + \Delta\varepsilon) - \Phi(\varepsilon) \\ &= \frac{\pi}{4}\left(\frac{8ma^2}{h^2}\right)^{3/2} \varepsilon^{1/2}\,\Delta\varepsilon + O((\Delta\varepsilon)^2)\end{aligned} \tag{1-35}$$

If we take $\varepsilon = 3kT/2$, $T = 300°\text{K}$, $m = 10^{-22}$ g, $a = 10$ cm, and $\Delta\varepsilon$ to be 0.01ε (in other words a 1% band around ε), then $\omega(\varepsilon, \Delta\varepsilon)$ is $O(10^{28})$.* So even for a system as simple as a particle-in-a-box, the degeneracy can be very large at room temperature.

For an N-particle system, the degeneracy is tremendously greater than $0(10^{28})$. To see this, consider a system of N noninteracting particles in a cube. The energy of this system is

$$E = \frac{h^2}{8ma^2}\sum_{j=1}^{N}(n_{xj}^2 + n_{yj}^2 + n_{zj}^2) = \frac{h^2}{8ma^2}\sum_{j=1}^{3N} s_j^2$$

where n_{xj}, n_{yj}, n_{zj}, and s_j are positive integers. The degeneracy of this system can be calculated by generalizing the above derivation for one particle. Using the volume of

---

* We use the notation $O(10^{28})$ say, to mean of the order of magnitude 10^{28}. This differs from standard mathematical notation, but there should be no confusion.

an N-dimensional sphere from Problem 1–24, the number of states with energy $\leq E$ is*

$$\Phi(E) = \frac{1}{\Gamma(N+1)\Gamma[(3N/2)+1]}\left(\frac{2\pi ma^2 E}{h^2}\right)^{3N/2} \tag{1–36}$$

where $\Gamma(n)$ here is the gamma function. (See Problem 1–58.) The number of states between E and $E + \Delta E$ is

$$\Omega(E, \Delta E) = \frac{1}{\Gamma(N+1)\Gamma(3N/2)}\left(\frac{2\pi ma^2}{h^2}\right)^{3N/2} E^{(3N/2-1)}\,\Delta E \tag{1–37}$$

In this case, $E = 3NkT/2$. If we take $T = 300°\text{K}$, $m = 10^{-22}$ g, $a = 10$ cm, $N = 6.02 \times 10^{23}$, and ΔE equal to $0.01E$, we get $\Omega(E, \Delta E)$ to be $O(10^N)$ (see Problem 1–23), an extremely large number. This shows that as the number of particles in the system increases, the quantum mechanical degeneracy becomes enormous. Although we have shown this only for a system of noninteracting particles confined to a cubical box, that is, an ideal gas, the result is generally true. We shall see in the next chapter that the concept of the degeneracy of a macroscopic system is very important.

There is another quantum mechanical result that we shall use later on. It often happens that the Hamiltonian of a many-body system can be written either exactly or approximately as a summation of one-particle or few-particle Hamiltonians, that is,

$$\mathscr{H} = \mathscr{H}_\alpha + \mathscr{H}_\beta + \mathscr{H}_\gamma + \cdots \tag{1–38}$$

Let the eigenvalues of $\mathscr{H}_j$ be ε_j, and the eigenfunctions be ψ_j, where $j = \alpha, \beta, \gamma, \ldots$ To solve the many-body Schrödinger equation, we let $\psi = \psi_\alpha\psi_\beta\psi_\gamma \ldots$. Then

$$\begin{aligned}
\mathscr{H}\psi &= (\mathscr{H}_\alpha + \mathscr{H}_\beta + \mathscr{H}_\gamma + \cdots)\psi_\alpha\psi_\beta\psi_\gamma \\
&= \psi_\beta\psi_\gamma \cdots \mathscr{H}_\alpha\psi_\alpha + \psi_\alpha\psi_\gamma \quad \mathscr{H}_\beta\psi_\beta + \cdots \\
&= \psi_\beta\psi_\gamma\varepsilon_\alpha\psi_\alpha + \psi_\alpha\psi_\gamma\varepsilon_\beta\psi_\beta + \cdots \\
&= (\varepsilon_\alpha + \varepsilon_\beta + \cdots)\psi_\alpha\psi_\beta\psi_\gamma \cdots = E\psi
\end{aligned} \tag{1–39}$$

In other words, the energy of the entire system is the sum of the energies of the individual particles if they do not interact. This is a very important result and will allow us to reduce a many-body problem to a one-body problem if the interactions are weak enough to ignore, such as in the case of a dilute gas. We shall see a number of cases where, even though the interactions are too strong to be ignored (such as in a solid), it is possible to formally or mathematically write the Hamiltonian in the form of Eq. (1–38). This will lead to defining quasi-particles like phonons and photons.

The last quantum mechanical topic we shall discuss here is that of the symmetry of wave functions with respect to the interchange of identical particles. Consider a system of N identical particles, described by a wave function $\psi(1, 2, 3, \ldots, N)$, where 1 denotes the coordinates of particle 1, and so on. If we interchange the position of any two of the particles, say particles 1 and 2, the wave function must either remain the same or change sign. (See Problem 1–26.) Thus if we let P_{12} be an operator that exchanges the two identical particles 1 and 2, then

$$\begin{aligned}
P_{12}\psi(1, 2, 3, \ldots, N) &= \psi(2, 1, 3, \ldots, N) \\
&= \pm\psi(1, 2, 3, \ldots, N)
\end{aligned} \tag{1–40}$$

---

* The extra factor of $\Gamma(N+1)$ occurs here because of the indistinguishability of the N particles. This will be discussed fully in Chapter 4.

It turns out that whether the wave function remains the same or changes sign is a function of the nature of the two identical particles that are exchanged. For particles with an integral spin (such as the He-4 nucleus, photons, ...), the wave function remains the same. In this case the wave function is called symmetric, and such particles are called bosons. For particles with half-integral spin (such as electrons, ...), the wave function is called antisymmetric, and the particles are called fermions. Chapter 4 considers the consequences of this symmetry requirement of wave functions.

1–4 THERMODYNAMICS

In this section we shall not attempt to review thermodynamics, but shall simply state the three laws and briefly discuss their consequences. Problems 1–27 through 1–36 review some of the equations and manipulations that arise in thermodynamics. Two topics that are not often treated in elementary physical chemistry are presented here, namely, Legendre transformations and Euler's theorem. Both of these topics will be used later on.

The pressure-volume work done by a system on its surroundings in going from state A to state B is

$$w = \int_A^B p\, dV$$

where p is the pressure exerted by the surroundings on the system. The differential quantity δw is positive if dV is positive.

The heat absorbed by the system from the surroundings during the change of the system from state A to state B is

$$q = \int_A^B \delta q$$

The first law of thermodynamics states that even though w and q depend upon the path taken from A to B, their difference does not. Their difference, then, is a function only of the two states A and B, or, namely, is a state function. This function is called the internal energy or thermodynamic energy and is denoted by E.

The first law of thermodynamics is

$$\begin{aligned}\Delta E &= E_B - E_A = q - w \\ &= \int_A^B \delta q - \int_A^B p\, dV \qquad (1\text{–}41)\end{aligned}$$

For simplicity, we consider only p–V work.

A reversible change is one in which the driving force (a difference in pressure, a difference in temperature, and so on) is infinitesimal. Any other change is called irreversible or spontaneous. Problem 1–27 asks the reader to show that for an isothermal process, $w_{\text{rev}} > w_{\text{irrev}}$ and $q_{\text{rev}} > q_{\text{irrev}}$.

The first law of thermodynamics is nothing but a statement of the law of conservation of energy. The second law is somewhat more abstract and can be stated in a number of equivalent ways. One of them is: There is a quantity S, called entropy, which is a state function. In an irreversible process, the entropy of the system and its surroundings increases. In a reversible process, the entropy of the system and its surroundings remains constant. The entropy of the system and its surroundings never decreases. The system and its surroundings are often referred to as the universe.

The mathematical expression for the difference in entropy between states A and B of a system is given by

$$\Delta S = \int_A^B \frac{dq_{\text{rev}}}{T} \tag{1–42}$$

Note that the heat appearing here is that associated with a reversible process. To compute ΔS between two states A and B, we must take the system from A to B in a reversible manner.

Another statement of the second law is: Along any reversible path, there exists an integrating factor T, common to all systems such that

$$dS = \frac{dq_{\text{rev}}}{T} \tag{1–43}$$

is an exact differential, that is, that S is a state function. Thus

$$\Delta S = \int_A^B \frac{dq_{\text{rev}}}{T}$$

For all other processes

$$\Delta S > \int_A^B \frac{dq}{T}$$

where T is the temperature of the surroundings.

The third law of thermodynamics states: If the entropy of each element in some crystalline state be taken as zero at the absolute zero of temperature, every substance has a finite positive entropy, but at the absolute zero of temperature, the entropy may become zero, and does become so in the case of perfect crystalline substances.

The second law is concerned with only the difference in the entropy between two states. The third law allows us to calculate the absolute entropy of a substance by means of the expressions

$$S - S_0 = \int_0^T \frac{dq_{\text{rev}}}{T} \quad \text{and} \quad S_0 = 0 \tag{1–44}$$

Problem 1–36 asks you to calculate the absolute entropy of gaseous nitromethane at its boiling point.

For simple one-component systems, the first law can be written in the form

$$dE = T\,dS - p\,dV \tag{1–45}$$

This implies that

$$\left(\frac{\partial E}{\partial S}\right)_V = T \quad \text{and} \quad \left(\frac{\partial E}{\partial V}\right)_S = -p \tag{1–46}$$

The simplicity of these partial derivatives implies that E is a "natural" function of S and V. For example, if we were to consider E to be a function of V and T (see Problem 1–30), we would get

$$dE = \left[T\left(\frac{\partial p}{\partial T}\right)_V - p\right] dV + C_V\,dT$$

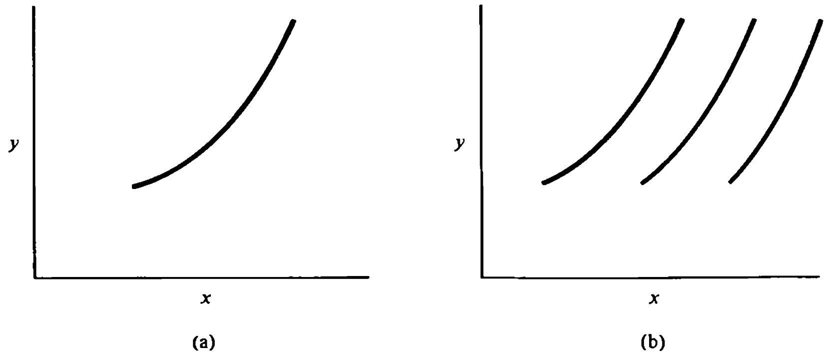

Figure 1–2. **(a) shows the function $y(x)$ and (b) shows a family of functions, all of which give the same value of y for any fixed value of p.**

Note that in this case the coefficients of dV and dT are not as simple as the coefficients of dV and dS obtained when E is expressed as a function of S and V. The "simplicity" of the expression $dE = T\ dS - p\ dV$ suggests that S and V are the "natural" variables for E. The quantities S and V (especially S) are difficult to control in the laboratory and consequently are not always the most desirable independent variables. A more useful pair might be (T, V) or (T, p). An important question that arises, then, is the existence of other thermodynamic state functions whose natural variables are (T, V) or (T, p), and so on. Furthermore, how would one find them if they do exist. This leads us to the topic of Legendre transformations.

We shall discuss a function of one variable in some detail and then simply present the generalization to a function of many variables. Consider a function $y = y(x)$, and let its slope be $p = p(x)$. We wish to describe the function $y(x)$ in terms of its slope. Figure 1–2, however, shows that the slope alone is not sufficient to completely specify $y(x)$. Figure 1–2(a) shows the curve $y(x)$, and Fig. 1–2(b) shows a family of curves, all of which give the same value of y for any one value of p. In order to uniquely describe the curve in Fig. 1–2(a), we must select one member of the family of curves in Fig. 1–2(b). We do this by specifying the intercepts of the tangent lines with the y-axis. Let the intercept be $\phi(p)$. Instead of describing the curve in Fig. 1–2(a) by y versus x, then, we can equally well represent it by specifying the slope at each point along with the intercept of the slope with the y-axis. Figure 1–3 shows these two representations. One sees that either representation can be used to describe the function. The relation between the two representations can be obtained by referring to Fig. 1–4. This figure shows that the slope p at any point is given by

$$p = \frac{y - \phi}{x - 0}$$

The result that we are after is

$$\phi(p) = y - px \tag{1–47}$$

The function $\phi(p)$ is the Legendre transformation of y. It is completely equivalent to $y(x)$, but considers p to be the independent variable instead of x. This may not be clear from the notation in Eq. (1–47), but it is understood there that y and x have been eliminated in favor of p by using the equations $y = y(x)$ and $x = x(p)$.

Let us apply this to the thermodynamic energy $E(S, V)$. We seek a function of T and V that is completely equivalent to E. Equation (1–46) shows that $T = (\partial E/\partial S)_V$,

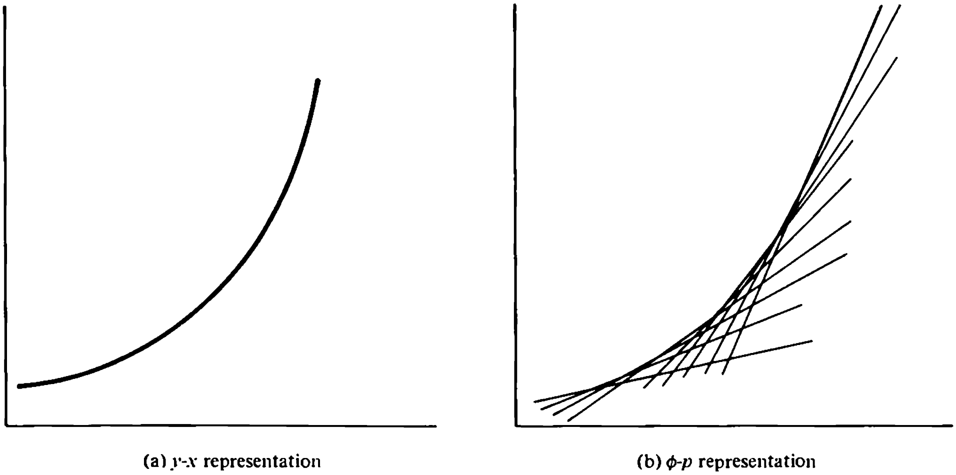
(a) y-x representation (b) ϕ-p representation

Figure 1–3. **In (a), the function is represented by the locus of points, y versus x. In (b), the same function is given by the envelope of its tangent curves.**

and so we are in a position to apply Eq. (1–47) directly. This can be treated as a one-variable problem, since V is held fixed throughout. Therefore the Legendre transformation of E that considers T and V to be the independent variables is $E - TS$. Of course, this is the Helmholtz free energy

$$A(T, V) = E - TS \tag{1-48}$$

whose differential form is

$$dA = -S\,dT - p\,dV \tag{1-49}$$

This shows that the natural variables of A are T and V. Another motivation for saying this is that the condition for equilibrium at constant T and V is that A assume its minimum value, or that $\Delta A \leq 0$ for a spontaneous process at constant T and V To prove this, write

$$\begin{aligned} dA &= dE - T\,dS - S\,dT \\ &= \delta q - p\,dV - \delta q_{\text{rev}} - S\,dT \\ &= \delta q - \delta q_{\text{rev}} \end{aligned} \tag{1-50}$$

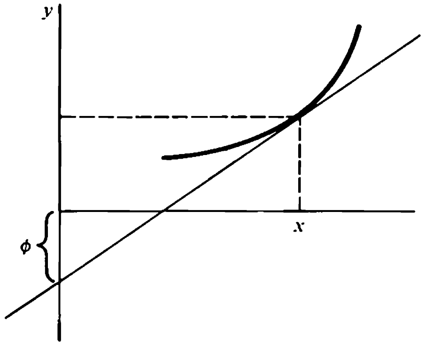

Figure 1–4. **The diagram used to derive the connection between the y–x representation and the ϕ–p representation.**

at constant T and V But $\delta q \leq \delta q_{\text{rev}}$ (see Problem 1–27), and so $\Delta A \leq 0$ at constant T and V

In elementary physical chemistry, the function $A = E - TS$ is often presented as an a priori definition. But it should be apparent now that this form is dictated by the Legendre transformation if one specifies T and V to be the independent variables.

A function, whose natural variables are S and p, can be obtained in the same manner. Equation (1–46) shows that $p = -(\partial E/\partial V)_S$, so Eq. (1–47) gives that $E + pV$ is a thermodynamic state function, whose natural variables are S and p. This function is, of course, the enthalpy.

The generalization of Eq. (1–47) to more than one variable is simply

$$\phi(p) = y - \sum_j p_j x_j \tag{1–51}$$

where the x_j's are the independent variables of y, and $p_j = (\partial y/\partial x_j)$. We can use Eq. (1–51) to construct a thermodynamic state function, whose natural variables are T and p. Using Eqs. (1–46) and Eq. (1–51), we see that such a function is $E - TS + pV$, the Gibb's free energy. Its differential form $dG = -S\,dT + V\,dp$ and the fact that $\Delta G \leq 0$ for a spontaneous change at constant T and p suggest that T and p are the natural variables of G.

Up to this point we have considered only closed one-component systems. In general, E, H, A, and G depend upon the number of moles or molecules of each component. If we let N_j be the number of moles of component j, we have

$$dE = T\,dS - p\,dV + \sum_j \left(\frac{\partial E}{\partial N_j}\right)_{S,V,N_K,\,j\neq K} dN_j \tag{1–52}$$

$$= T\,dS - p\,dV + \sum_j \mu_j\,dN_j \tag{1–53}$$

where the second line defines μ_j. By adding $d(pV)$ to both sides of Eq. (1–53), we get

$$dH = T\,dS + V\,dp + \sum_j \mu_j\,dN_j \tag{1–54}$$

If we subtract $d(TS)$ from both sides of Eq. (1–53), we get

$$dA = -S\,dT - p\,dV + \sum_j \mu_j\,dN_j \tag{1–55}$$

Similar manipulations give

$$dG = -S\,dT + V\,dp + \sum_j \mu_j\,dN_j \tag{1–56}$$

Equations (1–52) through (1–56) show that

$$\mu_j = \left(\frac{\partial E}{\partial N_j}\right)_{S,V,\ldots} = \left(\frac{\partial H}{\partial N_j}\right)_{S,p,\ldots} = \left(\frac{\partial A}{\partial N_j}\right)_{V,T,\ldots} = \left(\frac{\partial G}{\partial N_j}\right)_{p,T,\ldots} \tag{1–57}$$

The quantity μ_j is called the chemical potential.

There is a mathematical theorem, called Euler's theorem, which is very useful in thermodynamics. Before discussing Euler's theorem, however, we must define extensive and intensive variables. Extensive properties are additive; their value for the whole system is equal to the sum of their values for the individual parts. Examples are the volume, mass, and entropy. Intensive properties are not additive. Examples are temperature and pressure. The temperature of any small part of a system in equilibrium is the same as the temperature of the whole system. Euler's theorem deals with extensive and intensive variables. If

$$f(\lambda x_1, \lambda x_2, \ldots, \lambda x_N) = \lambda^n f(x_1, x_2, \ldots, x_N) \tag{1–58}$$

f is said to be a homogeneous function of order n. The functions $f(x) = 3x^2$ and $f(x, y, z) = xy^2 + z^3 - 6x^4/y$ are homogeneous functions of degree 2 and 3, respectively, whereas $f(x) = x^2 + 2x - 3$ and $f(x, y) = xy - e^{xy}$ are not homogeneous. Euler's theorem states that if $f(x_1, \ldots, x_N)$ is a homogeneous function of order n, then

$$nf(x_1, \ldots, x_N) = x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + \cdots + x_N \frac{\partial f}{\partial x_N} \tag{1–59}$$

The proof of Euler's theorem is simple. Differentiate Eq. (1–58) with respect to λ:

$$\begin{aligned} n\lambda^{n-1} f(x_1, x_2, \ldots, x_N) &= \left(\frac{\partial f}{\partial \lambda x_1}\right)\left(\frac{\partial \lambda x_1}{\partial \lambda}\right) + \left(\frac{\partial f}{\partial \lambda x_2}\right)\left(\frac{\partial \lambda x_2}{\partial \lambda}\right) + \cdots + \left(\frac{\partial f}{\partial \lambda x_N}\right)\left(\frac{\partial \lambda x_N}{\partial \lambda}\right) \\ &= x_1\left(\frac{\partial f}{\partial \lambda x_1}\right) + x_2\left(\frac{\partial f}{\partial \lambda x_2}\right) + \cdots + x_N\left(\frac{\partial f}{\partial \lambda x_N}\right) \end{aligned}$$

Euler's theorem is proved by letting $\lambda = 1$.

Extensive thermodynamic variables are homogeneous of degree 1. Let us apply Euler's theorem to the Gibb's free energy.

$$G(T, p, \lambda N_1, \lambda N_2, \ldots) = \lambda G(T, p, N_1, N_2, \ldots)$$

The variables T and p here can be treated as constants. Equation (1–59) gives that

$$G = \sum_j N_j \left(\frac{\partial G}{\partial N_j}\right)_{T, p, \ldots} = \sum_j N_j \mu_j \tag{1–60}$$

Taking the derivative of this at constant T and p,

$$dG = \sum_j N_j \, d\mu_j + \sum_j \mu_j \, dN_j \qquad \text{(constant } T \text{ and } p)$$

But using Eq. (1–56) at constant T and p, we have

$$\sum_j N_j \, d\mu_j = 0 \qquad \text{(constant } T \text{ and } p) \tag{1–61}$$

This is called the Gibbs-Duhem equation and is very useful in the thermodynamic study of solutions. (See *Physical Chemistry*, 4th ed., by W. J. Moore, p. 235, under "Additional Reading," for a simple application of the Gibbs-Duhem equation.)

We shall conclude this section on thermodynamics with a brief discussion of the application of thermodynamics to chemical equilibria. Consider the general reaction

$$\nu_A A + \nu_B B + \cdots \rightleftharpoons \nu_D D + \nu_E E + \cdots \tag{1–62}$$

The capital letters represent the formulas of the compounds, and the ν_j represent stoichiometric coefficients. It is more convenient to write Eq. (1–62) mathematically as

$$\nu_D D + \nu_E E + \cdots - \nu_A A - \nu_B B - \cdots = 0 \tag{1–63}$$

Define the extent of reaction λ, such that $dN_j = \nu_j \, d\lambda$ for all j, where the ν's for products are positive, and those for reactants are negative.

At constant T and p, we have

$$dG = \sum_j \mu_j \, dN_j = \left(\sum_j \mu_j \nu_j\right) d\lambda \qquad \text{(constant } T \text{ and } p)$$

At equilibrium, G must be a minimum with respect to λ, so we write

$$\sum_j \mu_j \nu_j = \nu_D \mu_D + \nu_E \mu_E + \cdots - \nu_A \mu_A - \nu_B \mu_B - \cdots = 0 \tag{1-64}$$

at equilibrium. The equilibrium between phases can be considered to be a chemical reaction of the form $A \rightleftharpoons B$, and so Eq. (1–64) gives that $\mu_A = \mu_B$ from the equilibrium condition between two pure phases.

Now consider the application of Eq. (1–64) to a chemical reaction between gases dilute enough to be considered ideal. Let the reaction be $\nu_A A + \nu_B B \rightleftharpoons \nu_C C + \nu_D D$. At constant temperature,

$$dG = V\,dp \qquad \text{(constant } T\text{)}$$

and so

$$G - G^0 = \int_{p_0}^{p} V\,dp = \int_{p_0}^{p} \frac{NkT}{p}\,dp = NkT \ln \frac{p}{p_0} \tag{1-65}$$

In this equation G^0 is the standard free energy of the gas, the standard state being the gas at a pressure p_0. Usually p_0 is taken to be 1 atmosphere. If we take N to be 1 mole, then G and G^0 become μ and μ^0. Each component in the reactive gas mixture will have an equation of the form of Eq. (1–65), and so we have

$$\mu_j(T, p) = \mu_j^0(T) + RT \ln \frac{p_j}{p_{0j}} \tag{1-66}$$

The total free energy change is

$$\begin{aligned} \Delta\mu &= \nu_C \mu_C + \nu_D \mu_D - \nu_A \mu_A - \nu_B \mu_B \\ &= \Delta\mu^0 + RT \ln \frac{(p_C')^{\nu_C}(p_D')^{\nu_D}}{(p_A')^{\nu_A}(p_B')^{\nu_B}} \end{aligned} \tag{1-67}$$

In this equation the (p')'s are p/p_0, that is, they are the pressures relative to the standard states. These (p')'s are unitless. The argument of the logarithm here has the *form* of an equilibrium constant, but is not equal to the equilibrium constant unless the pressures are those which exist at chemical equilibrium. Equation (1–67) gives the change in free energy of the conversion of reactants at *arbitrary* pressures to products at *arbitrary* pressures.

At equilibrium, $\Delta\mu = 0$, and we have

$$\begin{aligned} \Delta\mu^0 &= -RT \ln \left[\frac{(p_C')^{\nu_C}(p_D')^{\nu_D}}{(p_A')^{\nu_A}(p_B')^{\nu_B}} \right]_{\text{equilibrium}} \\ &= -RT \ln K_p \end{aligned} \tag{1-68}$$

There are extensive tabulations of μ^0's, and so $\Delta\mu^0$ is a simple matter to calculate. We see that if $\Delta\mu^0 < 0$, then $K_p > 1$, that is, the conversion of reactants *in their standard states* to products *in their standard states* proceeds spontaneously. On the other hand, if $\Delta\mu^0 > 0$, then $K_p < 1$, and we conclude that the reaction does not proceed spontaneously. It should be understood, however, that this applies only to reactants and products in their standard states. In general, it is $\Delta\mu$ along with Eq. (1–67) that determines the extent of a chemical reaction. (See Problem 1–34.)

1–5 MATHEMATICS

In this section we shall discuss several mathematical techniques or results that are repeatedly used in statistical thermodynamics. The topics we shall discuss here are random variables and distribution functions, Stirling's approximation, the binomial and multinomial coefficients, the Lagrange method of undetermined multipliers, and the behavior of binomial and multinomial coefficients for large numbers.

PROBABILITY DISTRIBUTIONS

Let u be a variable which can assume the M discrete values $u_1, u_2, \ldots, u_M$ with corresponding probabilities $p(u_1), p(u_2), \ldots, p(u_M)$. The variable u is said to be a discrete random variable, and $p(u)$ is said to be a discrete distribution. The mean, or average, value of u is

$$\bar{u} = \frac{\sum_{j=1}^{M} u_j p(u_j)}{\sum_{j=1}^{M} p(u_j)}$$

Since $p(u_j)$ is a probability, $p(u_j)$ must be normalized, that is, the summation in the denominator must equal unity. The mean of any function of u, $f(u)$, is given by

$$\overline{f(u)} = \sum_{j=1}^{M} f(u_j)p(u_j) \tag{1–69}$$

If $f(u) = u^m$, $\overline{f(u)}$ is called the mth moment of the distribution $p(u)$. If $f(u) = (u - \bar{u})^m$, $\overline{f(u)}$ is called the mth central moment, that is, the mth moment about the mean. In particular, the mean of $(u - \bar{u})^2$ is called the variance, and is a measure of the spread of the distribution. The square root of the variance is the standard deviation.

A very commonly occurring and useful discrete distribution is the Poisson distribution:

$$P(m) = \frac{a^m e^{-a}}{m!} \qquad m = 0, 1, 2, \ldots \tag{1–70}$$

This distribution has been applied to shot noise in electron tubes, the distribution of galaxies in space, aerial search, and many others.* (See Problem 1–42.)

If the random variable U is continuous rather than discrete, then we interpret $p(u)\,du$ as the probability that the random variable U lies between the values u and $u + du$. The mean of any function of U is

$$\overline{f(u)} = \int f(u)p(u)\,du \tag{1–71}$$

The limits of the integral are over the entire range of U.

The most important continuous probability distribution is the Gaussian distribution:

$$p(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{(x-\bar{x})^2}{2\sigma^2}\right\} \qquad -\infty \le x \le \infty \tag{1–72}$$

The quantity σ^2, which is the variance, controls the width of the Gaussian distribution. The smaller the σ, the narrower the Gaussian distribution becomes. In the limit $\sigma \to 0$, Eq. (1–72) becomes a delta function (this is one representation of a delta function of Appendix B). Problems 1–43 through 1–45 involve some important results based on Eq. (1–72).

---

* See *Modern Probability Theory and Its Applications* by E. Parzen (New York: Wiley, 1960).

STIRLING'S APPROXIMATION

In statistical thermodynamics we often encounter factorials of very large numbers, such as Avogadro's number. The calculation and mathematical manipulation of factorials become awkward for large N. Therefore it is desirable to find an approximation for $N!$ for large N. Problems of this sort occur often in mathematics and are called asymptotic approximations, that is, an approximation to a function which improves as the argument of that function increases. Since $N!$ is actually a product, it is convenient to deal with $\ln N!$ because this is a sum. The asymptotic approximation to $\ln N!$ is called Stirling's approximation, which we now derive.

Since $N! = N(N-1)(N-2)\cdots(2)(1)$, $\ln N!$ is

$$\ln N! = \sum_{m=1}^{N} \ln m \tag{1-73}$$

Figure 1-5 shows $\ln x$ plotted versus x. The sum of the areas under these rectangles up to N is $\ln N!$. Figure 1-5 also shows the continuous curve $\ln x$ plotted on the same graph. Thus $\ln x$ is seen to form an envelope to the rectangles, and this envelope becomes a steadily smoother approximation to the rectangles as x increases. We can approximate the area under these rectangles by the integral of $\ln x$. The area under $\ln x$ will poorly approximate the rectangles only in the beginning. If N is large enough (we are deriving an asymptotic expansion), this area will make a negligible contribution to the total area. We may write, then,

$$\ln N! = \sum_{m=1}^{N} \ln m \approx \int_{1}^{N} \ln x \, dx = N \ln N - N \qquad (N \text{ large}) \tag{1-74}$$

which is Stirling's approximation to $\ln N!$. The lower limit could just as well have been taken as 0 in Eq. (1-74), since N is large. (Remember that $x \ln x \to 0$ as $x \to 0$.)

A more refined derivation of Stirling's approximation gives $\ln N! \approx N \ln N - N + \ln(2\pi N)^{1/2}$, but this additional term is seldom necessary. (See Problem 1-59.)

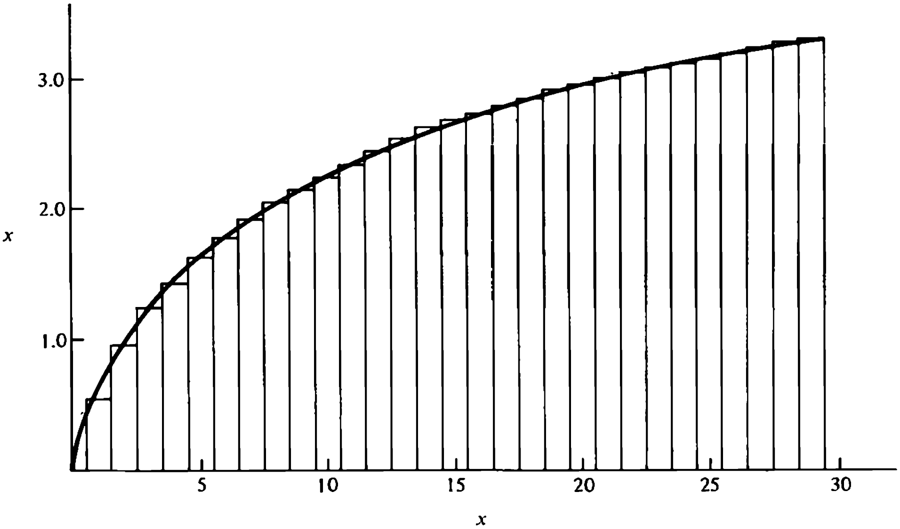

Figure 1-5. **A plot of $\ln x$ versus x, showing how the summation of $\ln m$ can be approximated by the integral of $\ln x$.**

BINOMIAL AND MULTINOMIAL DISTRIBUTION

During the course of our discussion of the canonical ensemble, we shall encounter the problem of determining how many ways it is possible to divide N distinguishable systems into groups such that there are n_1 systems in the first group, n_2 systems in the second group, and so on, and such that $n_1 + n_2 + \quad = N$, that is, all the systems are accounted for. This is actually one of the easiest problems in combinatorial analysis. To solve this, we first calculate the number of permutations of N distinguishable objects, that is, the number of possible different arrangements or ways to order N distinguishable objects. Let us choose one of the N objects and place it in the first position, one of the $N - 1$ remaining objects and place it in the second position, and so on, until all N objects are ordered. Clearly there are N choices for the first position, $N - 1$ choices for the second position, and so on, until finally there is only one object left for the Nth position. The total number of ways of doing this is then the product of all the choices,

$$N(N-1)(N-2)\cdots(2)(1) \equiv N! \qquad \text{(distinguishable objects)}$$

Next we calculate the number of ways of dividing N distinguishable objects into two groups, one group containing N_1 objects, say, and the other containing the remaining $N - N_1$. There are $N(N-1)\cdots(N - N_1 + 1)$ ways to form the first group, and $N_2! = (N - N_1)!$ ways to form the second group. The total number is, then, the product

$$N(N-1)\cdots(N-N_1+1) \times (N-N_1)! = \frac{N!}{(N-N_1)!} \times (N-N_1)! = N!$$

But this has overcounted the situation drastically, since the order in which we place N_1 members in the first group and N_2 in the second group is immaterial to the problem as stated. All $N_1!$ orders of the first group and $N_2!$ orders of the second group correspond to just one division of N objects into N_1 objects and N_2 objects. Therefore the desired result is

$$\frac{N!}{N_1!(N-N_1)!} = \frac{N!}{N_1!N_2!} \tag{1–75}$$

Since the combination of factorials in Eq. (1–75) occurs in the binomial expansion,

$$(x+y)^N = \sum_{N_1=0}^{N} \frac{N!x^{N-N_1}y^{N_1}}{N_1!(N-N_1)!} = \sum_{N_1 N_2}^{*} \frac{N!x^{N_1}y^{N_2}}{N_1!N_2!} \tag{1–76}$$

$N!/N_1!(N-N_1)!$ is called a binomial coefficient. The asterisk on the second summation in Eq. (1–76) signifies the restriction $N_1 + N_2 = N$.

The generalization of Eq. (1–75) to the division of N into r groups, the first containing N_1, and so on, is easily seen to be

$$\frac{N!}{N_1!N_2!\cdots N_r!} = \frac{N!}{\prod_{j=1}^{r} N_j!} \tag{1–77}$$

where $N_1 + N_2 + \cdots + N_r = N$. This is known as a multinomial coefficient, since it occurs in the expansion

$$(x_1 + x_2 + \cdots + x_r)^N = \sum_{N_1=0}^{N} \sum_{N_2=0}^{N} \quad \sum_{N_r=0}^{N}{}^{*} \frac{N!x_1^{N_1}\cdots x_r^{N_r}}{\prod_{j=1}^{r} N_j!} \tag{1–78}$$

where this time the asterisk signifies the restriction $N_1 + N_2 + \cdots + N_r = N$.

There are a number of other combinatorial formulas that are useful in statistical thermodynamics, but Eq. (1–77) is the most useful for our purposes. Combinatorial formulas can become rather demanding to derive. We refer to Appendix AVII of Mayer and Mayer* which contains a collection of formulas.

METHOD OF LAGRANGE MULTIPLIERS

It will be necessary, later, to maximize Eq. (1–77) with the constraint $N_1 + N_2 + \cdots + N_r = \text{constant}$. This brings us to the mathematical problem of maximizing a function of several (or many) variables $f(x_1, x_2, \ldots, x_r)$ when the variables are connected by other equations, say $g_1(x_1, \ldots, x_r) = 0$, $g_2(x_1, \ldots, x_r) = 0$, and so on. This type of problem is readily handled by the method of Lagrange undetermined multipliers.

If it were not for the constraints, $g_j(x_1, x_2, \ldots, x_r) = 0$, the maximum of $f(x_1, \ldots, x_r)$ would be given by

$$\delta f = \sum_{j=1}^{r} \left(\frac{\partial f}{\partial x_j}\right)_0 \delta x_j = 0 \tag{1–79}$$

where the zero subscript indicates that this equation equals zero only when the r partial derivatives are evaluated at the maximum (or minimum) of f. Denote these values of x_j by x_j^0 If there were no constraints, each of the δx_j would be able to be varied independently and arbitrarily, and so we would conclude that $(\partial f/\partial x_j) = 0$ for every j, since δf must equal zero. This would give r equations from which the values of the $r x_j^0$ could be obtained.

On the other hand, if there is some other relation between the x's, such as $g(x_1, x_2, \ldots, x_r) = 0$, we have the additional equation

$$\delta g = \sum_{j=1}^{r} \left(\frac{\partial g}{\partial x_j}\right)_0 \delta x_j = 0 \tag{1–80}$$

This equation serves as a constraint that the δx_j must satisfy, thus making one of them depend upon the other $r - 1$. In the Lagrange method, one multiplies Eq. (1–80) by some parameter, say λ, and adds the result to Eq. (1–79) to get

$$\sum_{j=1}^{r} \left(\frac{\partial f}{\partial x_j} - \lambda \frac{\partial g}{\partial x_j}\right)_0 \delta x_j = 0 \tag{1–81}$$

The δx_j are still not independent, because of Eq. (1–80), and so they cannot be varied independently. Equation (1–80), however, can be treated as an equation giving one of the δx_j in terms of the other $r - 1$ independent ones. Pick any one of the $r\ \delta x_j$ as the dependent one. Let this be δx_μ.

The trick now is that we have not specified λ yet. We set it equal to $(\partial f/\partial x_\mu)_0/(\partial g/\partial x_\mu)_0$, making the coefficient of δx_μ in Eq. (1–81) vanish. The subscript zero here indicates that $(\partial f/\partial x_\mu)$ and $(\partial g/\partial x_\mu)$ are to be evaluated at values of the x_j such that f is at its maximum (or minimum) under the constraint of Eq. (1–80). Of course, we do not know these values of x_j yet, but we can nevertheless formally define λ in this manner. This leaves a sum of terms in Eq. (1–81) involving only the independent δx_j, which can be varied independently, yielding that

$$\left(\frac{\partial f}{\partial x_j}\right)_0 - \lambda\left(\frac{\partial g}{\partial x_j}\right)_0 = 0 \qquad j = 1, 2, \ldots, \mu - 1, \mu + 1, \ldots, r$$

---

* See Mayer and Mayer, *Statistical Mechanics* (New York: Wiley, 1940).

If we combine these $r - 1$ equations with our choice for λ, we have

$$\left(\frac{\partial f}{\partial x_j}\right)_0 - \lambda\left(\frac{\partial g}{\partial x_j}\right)_0 = 0 \tag{1-82}$$

for all j.

As we said above, the choice of λ here is certainly formal, since both $(\partial f/\partial x_\mu)_0$ and $(\partial g/\partial x_\mu)_0$ must be evaluated at these values of x_j which maximizes f, but these are known from Eq. (1–82) only in terms of λ. But this presents no difficulty, since in practice λ is determined by physical requirements. Examples of this will occur in the next two chapters.

Lagrange's method becomes no more difficult in the case in which there are several constraints. Let $g_1(x_1, \ldots, x_r), g_2(x_1, \ldots, x_r), \ldots$ be a set of constraints. We introduce a Lagrange multiplier for each $g_i(x_1, \ldots, x_r)$ and proceed as above to get

$$\frac{\partial f}{\partial x_j} - \lambda_1 \frac{\partial g_1}{\partial x_j} - \lambda_2 \frac{\partial g_2}{\partial x_j} - \cdots = 0 \tag{1-83}$$

BINOMIAL DISTRIBUTION FOR LARGE NUMBERS

Lastly, there is one other mathematical observation we need here in order to facilitate the discussion in the next chapter. This observation concerns the shape of the multinomial coefficient [Eq. (1–78)] as a function of the N_j's, as the N_j's become very large. To simplify notation, we shall consider only the binomial coefficient, but this will not affect our conclusions. Let us first find the value of N_1 for which $f(N_1) = N!/N_1!(N - N_1)!$ reaches its maximum value. Since N_1 and N are both very large, we treat them as continuous variables. Also since $\ln x$ is a monotonic function of x, we can maximize $f(N_1)$ by maximizing $\ln f(N_1)$. This allows us to use Stirling's approximation. The maximum of $f(N_1)$ is found, then, from

$$\frac{d \ln f(N_1)}{dN_1} = 0$$

to be located at $N_1{}^* = N/2$. Let us now expand $\ln f(N_1)$ about this point. The Taylor expansion is

$$\ln f(N_1) = \ln f(N_1{}^*) + \frac{1}{2}\left(\frac{d^2 \ln f(N_1)}{dN_1{}^2}\right)_{N_1 = N_1{}^*} (N_1 - N_1{}^*)^2 + \cdots \tag{1-84}$$

The linear term in $N_1 - N_1{}^*$ is missing, because the first derivative of $\ln f(N)$ is zero at $N_1 = N_1{}^*$. The second derivative appearing in Eq. (1–84) is equal to $-4/N$. Thus if we ignore higher-order terms (see Problem 1–53), Eq. (1–84) can be written in the form of a Gaussian curve

$$f(N_1) = f(N_1{}^*)\exp\left\{-\frac{2(N_1 - N_1{}^*)^2}{N}\right\} \tag{1-85}$$

Comparison of this with the standard form of the Gaussian function

$$f(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{(x - x^*)^2}{2\sigma^2}\right\} \tag{1-86}$$

shows that the standard deviation is of the order of $N^{1/2}$. Equation (1–85) is, therefore, a bell-shaped function, centered at $N_1{}^* = N/2$ and having a width of a few multiples of $N^{1/2}$. Problem 1–43 establishes the well-known fact that a Gaussian function goes essentially to zero when x differs from x^* by a few σ's. Since we are interested only in large values of N_1 (or N), say numbers of the order of 10^{20}, we have a bell-shaped curve that is contained between 10^{20} ± a few multiples of 10^{10}, which, if plotted, would for all practical purposes look like a delta function centered at $N_1{}^* = N/2$. Thus we have shown that the binomial coefficient peaks very strongly at the point $N_1 = N_2 = N/2$. This same behavior occurs for a multinomial coefficient as well. If there are s N_j's, the multinomial coefficient has a very sharp maximum at the point $N_1 = N_2 = \cdots = N_s = N/s$. (See Problem 1–50.) This peak becomes sharper as the N_j's become larger, and become a delta function in the limit $N_j \to \infty$ for all j.

MAXIMUM TERM METHOD

Another important result, which is a consequence of the large numbers encountered in statistical mechanics, is the *maximum-term method.* It says that under appropriate conditions the logarithm of a summation is essentially equal to the logarithm of the maximum term in the summation. To see how this goes, consider the sum

$$S = \sum_{N=1}^{M} T_N$$

where $T_N > 0$ for all N. Since all the terms are positive, the value of S must be greater than the value of the largest term, say $T_{\max}$, and less than the product of the number of terms and the value of the largest term. Thus we can write

$$T_{\max} \le S \le M T_{\max}$$

Taking logarithms gives

$$\ln T_{\max} \le \ln S \le \ln T_{\max} + \ln M$$

We shall see that it is often the case in statistical mechanics that $T_{\max}$ will be $O(e^M)$. Thus we have

$$O(M) \le \ln S \le O(M) + \ln M$$

For large M, $\ln M$ is negligible with respect to M itself, and so we see that $\ln S$ is bounded from above and below by $\ln T_{\max}$, and so

$$\ln S = \ln T_{\max}$$

This is a rather remarkable theorem, and like a number of other theorems used in statistical mechanics, its validity results from the large numbers involved.

ADDITIONAL READING

1–2

Davidson, N. 1962. *Statistical mechanics.* New York: McGraw-Hill. Chapter 2.
Goldstein, H. 1950. *Classical mechanics.* Reading, Mass.: Addison-Wesley.
Symon, K. R. 1960. *Mechanics.* Reading, Mass.: Addison-Wesley.
Tolman, R. C. 1938. *Statistical mechanics.* London: Oxford University Press. Chapters 1 and 2.

1–3

Hameka, H. F. 1967. *Introduction to quantum theory.* New York: Harper & Row.
Hanna, M. W. 1969. *Quantum mechanics in chemistry,* 2nd ed. New York: Benjamin.
Hill, T. L. 1956. *Statistical mechanics.* New York: McGraw-Hill. Section 8.
Karplus, M., and Porter, R. N. 1970. *Atoms and molecules.* New York: Benjamin.
Kestin, J., and Dorfman, J. R. 1971. *A course in statistical thermodynamics.* New York: Academic. Chapter 3.
Pilar, F. L. 1968. *Elementary quantum chemistry.* New York: McGraw-Hill.
Tolman, R. C. 1938. *Statistical mechanics.* London: Oxford University Press. Chapters 7 and 8.

1–4

ANDREWS, F. C. 1971. *Thermodynamics.* New York: Wiley.
CALLEN, H. B. 1960. *Thermodynamics.* New York: Wiley.
MAHAN, B. H. 1963. *Elementary chemical thermodynamics.* New York: Benjamin.
MOORE, W. J. 1972. *Physical chemistry,* 4th ed. Englewood Cliffs, N.J.: Prentice-Hall.
REISS, H. 1965. *Methods of thermodynamics.* Boston, Mass.: Blaisdell.
WASER, J. 1966. *Basic chemical thermodynamics.* New York: Benjamin.

1–5

ABRAMOWITZ, M., and STEGUN, I. A. "Handbook of Mathematical Functions," *Natl. Bur. Stan. Appl. Math. Series,* **55,** 1964.
ARFKEN, G., 1970. *Mathematical methods for physicists,* 2nd ed. New York: Academic.
KESTIN, J., and DORFMAN, J. R. 1971. *A course in statistical thermodynamics.* New York: Academic. Chapter 4.
KREYSZIG, E. 1962. *Advanced engineering mathematics,* 2nd ed. New York, Wiley.
REIF, F. 1965. *Statistical and thermal physics.* New York: McGraw-Hill, Chapter 1.

PROBLEMS

1–1. Solve the equation of motion of a body of mass m dropped from a height h. Assume that there exists a viscous drag on the body that is proportional to and in the opposite direction to the velocity of the body. (Let the proportionality constant be γ.) Solve for the so-called terminal velocity, that is, the limiting velocity as $t \to \infty$.

1–2. Calculate the trajectory of a shell shot out of a cannon with velocity v_0, assuming no aerodynamic resistance and that the cannon makes an angle θ with the horizontal axis.

1–3. Remembering that the potential energy is given by

$$V(x) = -\int_0^x F(\xi)\,d\xi = \tfrac{1}{2}kx^2$$

for a simple harmonic oscillator, derive an expression for the total energy as a function of time. Discuss how the kinetic and potential energy behave as a function of time.

1–4. Solve the equation for a harmonic oscillator of mass m and force constant k that is driven by an external force of the form $F(t) = F_0 \cos \omega_0 t$.

1–5. Show that

$$\xi(t) = A \sin \omega t + B \cos \omega t$$

can be written as

$$\xi(t) = C \sin(\omega t + \phi)$$

1–6. Show that the total linear momentum is conserved for a system of N particles with an interaction potential which depends only on the distance between particles.

1–7. When does $p = \partial L/\partial \dot{q}$ but $\neq \partial K/\partial \dot{q}$?

1–8. Consider a system of two-point particles with masses m_1 and m_2 moving in two dimensions. It is very common for their potential of interaction to depend upon their relative coordinates $(x_1 - x_2, y_1 - y_2)$ only. Thus the total energy is

$$E = \frac{m_1}{2}(\dot{x}_1^2 + \dot{y}_1^2) + \frac{m_2}{2}(\dot{x}_2^2 + \dot{y}_2^2) + U(x_1 - x_2, y_1 - y_2)$$

Now introduce four new variables

$$X = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2} \qquad Y = \frac{m_1 y_1 + m_2 y_2}{m_1 + m_2}$$

$$x_{12} = x_1 - x_2 \qquad y_{12} = y_1 - y_2$$

and show that this two-body problem can be reduced to two one-body problems, one involving the center of mass of the system and one involving the *relative* motion of the two particles.

Give a physical interpretation of the ratio $m_1m_2/(m_1 + m_2)$ that arises naturally in the relative motion. What is this quantity called? This result is easily extended to three dimensions.

1–9. Extend the development of Problem 1–8 to the case in which each particle also experiences an external potential energy, say $U(x_1, y_1, z_1)$ and $U(x_2, y_2, z_2)$. Interpret the resulting equations.

1–10. Derive Lagrange's equations for a particle moving in two dimensions under a central potential $u(r)$. Which of these equations illustrates the law of conservation of angular momentum? Is angular momentum conserved if the potential depends upon θ as well?

1–11. For a particle moving in three dimensions under the influence of a spherically symmetrical potential $U = U(r)$, write down the Lagrangian and the equations of motion in spherical coordinates (r, θ, ϕ). Show that $H = K + V$ from

$$H = \sum p_i \dot{q}_i - L$$

for this potential.

1–12. Solve the equation of motion of two masses m_1 and m_2 connected by a harmonic spring with force constant k.

1–13. Start with Lagrange's equations in Cartesian coordinates, that is,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = \frac{\partial L}{\partial x}$$

and so on. Now introduce three generalized coordinates q_1, q_2, and q_3 which are related to the Cartesian coordinates by $x = x(q_1, q_2, q_3)$, and so on. Show that by transforming Lagrange's equations from $x, \dot{x}, y, \dot{y}, z$, and $\dot{z}$ as independent variables to $q_1, \dot{q}_1, q_2, \dot{q}_2, q_3$, and $\dot{q}_3$ we get

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_1}\right) = \frac{\partial L}{\partial q_1}$$

and so on.

1–14. If H, the classical Hamiltonian, does not depend explicitly on time, show that $dH/dt = 0$. What does this mean physically? Is this true if H does depend explicitly upon time?

1–15. Consider the rotation of a diatomic molecule with a fixed internuclear separation l and masses m_1 and m_2. By employing center of mass and relative coordinates, show that the rotational kinetic energy can be written in spherical coordinates as

$$\tfrac{1}{2}I(\dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta)$$

and from this derive the rotational Hamiltonian

$$H_{\text{rot}} = \frac{1}{2I}\left(p_\theta^2 + \frac{p_\phi^2}{\sin^2 \theta}\right)$$

In these equations, $I = \mu l^2$, where μ is the reduced mass. This Hamiltonian is useful for studying the rotation of diatomic molecules.

1–16. Show that the motion of a particle under a central force law takes place entirely in a single plane.

1–17. What is the expectation (average) value for the linear momentum p_x of a particle in a one-dimensional box p_x^2? Briefly discuss your results.

1–18. Show that the energy eigenvalues of a free particle confined to a cube of length a are given by

$$\varepsilon = \frac{h^2}{8ma^2}(n_x^2 + n_y^2 + n_z^2) \qquad n_x, n_y, n_z = 1, 2, \ldots$$

1–19. Show that the energy eigenvalues of a free particle confined to a rectangular parallepiped of lengths a, b, and c are given by

$$\varepsilon = \frac{h^2}{8m}\left(\frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2}\right) \qquad n_x, n_y, n_z = 1, 2, \ldots$$

1–20. Calculate the energy eigenvalues of a particle confined to a ring of radius a.

1–21. Show that the Hamiltonian operator of a rigid rotor is given by Eq. (1–32).

1–22. Calculate the degeneracy of the first few levels of a free particle confined to a cube of length a.

1–23. Verify the calculation that follows Eq. (1–37) which shows that the quantum mechanical degeneracy of a macroscopic system is $O(10^N)$.

1–24. We need to know the volume of an N-dimensional sphere in order to derive Eq. (1–36). This can be determined by the following device. Consider the integral

$$I = \int_{-\infty}^{\infty} \quad \int e^{-(x_1^2 + x_2^2 + \cdots + x_N^2)}\, dx_1\, dx_2 \cdots dx_N$$

First show that $I \quad \pi^{N/2}$. Now one can formally transform the volume element $dx_1\, dx_2 \cdots dx_N$ to N-dimensional spherical (hyperspherical) coordinates to get

$$\int_{\text{angles}} dx_1\, dx_2 \cdots dx_N \rightarrow r^{N-1} S_N\, dr$$

where S_N is the factor that arises upon integration over the angles. Show that $S_2 - 2\pi$ and $S_3 = 4\pi$. S_N can be determined for any N by writing I in hyperspherical coordinates:

$$I \quad \int_0^{\infty} e^{-r^2} r^{N-1} S_N\, dr$$

Show that $I = S_N \Gamma(N/2)/2$, where $\Gamma(x)$ is the gamma function (see Problem 1–58). Equate these two values for I to get

$$S_N = \frac{2\pi^{N/2}}{\Gamma(N/2)}$$

Show that this reduces correctly for $N = 2$ and 3. Lastly now, convince yourself that the volume of an N-dimensional sphere of radius a is given by

$$V_N = \int_0^a S_N r^{N-1}\, dr = \frac{\pi^{N/2}}{\Gamma\left(\frac{N}{2} + 1\right)} a^N$$

and show that this reduces correctly for $N = 2$ and 3.

1–25. Derive an expression for the density of translational quantum states for a two-dimensional ideal gas.

1–26. Prove that a many-body wave function must be either symmetric or antisymmetric under the interchange of any two particles. Hint: Apply the exchange operation twice.

1–27. Show for an isothermal process that $w_{\text{rev}} > w_{\text{irrev}}$ and $q_{\text{rev}} > q_{\text{irrev}}$.

1–28. Derive the thermodynamic equation

$$C_p - C_V = \left[p + \left(\frac{\partial E}{\partial V}\right)_T\right]\left(\frac{\partial V}{\partial T}\right)_p$$

and evaluate this difference for an ideal gas and a gas that obeys the van der Waals equation.

1–29. Derive the thermodynamic equation of state

$$\left(\frac{\partial E}{\partial V}\right)_T - T\left(\frac{\partial p}{\partial T}\right)_V = -p$$

1–30. Derive the equation

$$dE = \left[T\left(\frac{\partial p}{\partial T}\right)_V - p\right] dV + C_V\, dT$$

and from this show that $(\partial E/\partial V)_T = a/V^2$ for a van der Waals gas.

1–31. Show that

$$\left(\frac{\partial E}{\partial V}\right)_{\mu/T,\,1/T} + \frac{1}{T}\left(\frac{\partial p}{\partial(1/T)}\right)_{\mu/T,\,V} = -p$$

1–32. Derive an expression for $\partial \ln K/\partial T$ in terms of ΔH, the heat of reaction, and in terms of C_p, the heat capacity at constant pressure.

1–33. Consider the "water-gas" reaction

$$CO + H_2O(g) \rightarrow H_2 + CO_2$$

where

$$K_p = \frac{P_{H_2} P_{CO_2}}{P_{CO} P_{H_2O}}$$

and given the following data:

Substance	(kcal/mole)	a	$b \times 10^3$	$c \times 10^7$	ΔH^0_{298}(kcal/m)
CO	−32.81	6.42	1.67	1.96	−26.4157
$H_2O(g)$	−54.64	7.26	2.30	2.83	−57.7979
CO_2	94.26	6.21	10.40	−35.45	−94.0518
H_2	0.00	6.95	−0.20	4.81	0.00

where the heat capacity of the gases in cal deg^{-1} $mole^{-1}$ is given by

$$C_p = a + bT + cT^2$$

Calculate K_p at 298°K and 800°K.

1–34. Calculate the free energy change at 700°C for the conversion of carbon monoxide at 10 atm and water vapor at 5 atm to carbon dioxide and hydrogen at partial pressures of 1.5 atm each. The equilibrium constant K_p for this reaction is 0.71. Is this process theoretically feasible?

1–35. It is illustrated in Chapter 17 that the speed of sound c_0 propagated through a gas is

$$c_0 = (m\rho\kappa_S)^{-1/2}$$

where κ_S is the adiabatic compressibility

$$\kappa_S = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_S$$

Show that this is equivalent to

$$c_0 = V\left\{-\frac{\gamma}{M}\left(\frac{\partial p}{\partial V}\right)_T\right\}^{1/2}$$

where $\gamma = C_p/C_V$, and M is the molecular weight of the gas. Using the above result, show that

$$c_0 = \left(\gamma \frac{RT}{M}\right)^{1/2}$$

for an ideal gas.

1–36. Jones and Giauque obtained the following values for C_p of nitromethane.

°K	15	20	30	40	50	60	70	80	90	100
C_P	0.89	2.07	4.59	6.90	8.53	9.76	10.70	11.47	12.10	12.62
°K	120	140	160	180	200	220	240	260	280	300
C_P	13.56	14.45	15.31	16.19	17.08	17.98	18.88	25.01	25.17	25.35

The melting point is 244.7°K, heat of fusion 2319 cal/mole. The vapor pressure of the liquid at 298.1°K is 3.666 cm. The heat of vaporization at 298.0°K is 9147 cal/mole. Calculate the third-law entropy of CH_2NO_2 gas at 298.1°K and 1 atm pressure (assuming ideal gas behavior).

1–37. Derive the Legendre transformation of E in which $S \rightarrow T$ and $N \rightarrow \mu$.

1–38. Apply a Legendre transformation to the Lagrangian $L(q_j, \dot{q}_j)$ to eliminate the generalized velocities in favor of generalized momenta, defined by $p_j \equiv \partial L/\partial \dot{q}_j$. What function does this turn out to be?

1–39. Find the natural function of V, E, and μ. Hint: Start with the natural function of V, E, and N, namely, S, and transform $N \rightarrow \mu$.

1–40. Derive the Legendre transformation of E in which $S \rightarrow T$, $N \rightarrow \mu$, and $V \rightarrow p$. What peculiar thing happens when all the extensive variables are transformed out?

1–41. Show that $\overline{(x-\bar{x})^2} = \overline{x^2} - \bar{x}^2$.

1–42. Show that the Poisson distribution $P(m) = a^m e^{-a}/m!$ is normalized. Calculate $\bar{m}$ and the variance. What is the significance of the parameter a?

1–43. Sketch the Gaussian distribution as σ (or even $\sigma/\bar{x}$) becomes smaller and smaller. To what type of distribution does a Gaussian go in the limit $\sigma \rightarrow 0$. Discuss the meaning of this distribution.

1–44. For the Gaussian distribution $p(x)$ show that

(a)

$$\int_{-\infty}^{\infty} p(x)\, dx = 1$$

(b) Calculate the nth central moment where $n = 0, 1, 2,$ and 3.

(c) In the limit $\sigma \rightarrow 0$ what kind of distribution is approached where

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right)$$

1–45. The quantity $\overline{(x-\bar{x})^j}$ is called the jth central moment. Show that all odd central moments of a Gaussian vanish. What about the even ones? Relate the $j = 2$ central moment to the parameter σ.

1–46. Let $f(x, y)$ be a joint probability density, that is, $f(x, y)\, dx\, dy$ is the probability that X lies between x and $x + dx$ *and* Y lies between y and $y + dy$. If X and Y are *independent*, then

$$f(x, y)\, dx\, dy = f_1(x) f_2(y)\, dx\, dy$$

If X and Y are independent, show that the mean and variance of their sum is equal to the sum of the means and variances, respectively, of X and Y; that is, show that if $W = X + Y$, then

$$\bar{W} = \bar{X} + \bar{Y}$$

$$\overline{(W-\bar{W})^2} = \overline{(X-\bar{X})^2} + \overline{(Y-\bar{Y}^2)}$$

1–47. Let X be a random variable on the positive numbers, $0 \leq x < \infty$, and let $p(x)$ be its probability density function. The function $\phi(s)$ defined by

$$\phi(s) = \int_0^{\infty} e^{-sx} p(x)\, dx$$

is called the characteristic function of $p(x)$. Find the relation between $\phi(s)$ and the moments of $p(x)$. Is knowledge of all the moments of $p(x)$ (assuming they exist) sufficient to specify $p(x)$ itself? Why or why not?

1–48. Show that the characteristic function of the density function of the sum of two independent random variables is the product of the characteristic functions of the densities of the two random variables themselves. What is the density function of $W = X + Y$?

1–49. Maximize

$$W(N_1, N_2, \ldots, N_M) = \frac{N!}{\prod_{j=1}^{M} N_j!}$$

with respect to each N_j under the constraints that

$$\sum N_j = N \;= \text{a fixed constant}$$

$$\sum E_j N_j = \mathscr{E} = \text{another fixed constant}$$

Hint: Consider the N_j's to be continuous, large enough to use Stirling's approximation of $N_j!$, and leave your answer in terms of the two undetermined multipliers.

1–50. Show that the maximum of a multinomial distribution is given when $N_1 = N_2 = \cdots N_s = N/s$.

1–51. Use the method of undetermined multipliers to show that

$$-\sum_{j=1}^{N} P_j \ln P_j$$

subject to the condition

$$\sum_{j=1}^{N} P_j = 1$$

is a maximum when $P_j = \text{constant}$.

1–52. Consider the sum

$$\sum_{N=0}^{M} \frac{M!x^N}{N!(M-N)!}$$

where $x = O(1)$, and M and N are $O(10^{20})$. First show that $\ln \sum = M \ln(1 + x)$ *exactly*, and then calculate the logarithm of the maximum term. Hint: Remember the binomial expansion.

1–53. Show that the higher terms that were dropped in the expansion of $\ln f(N)$ in Eq. (1–84) are completely negligible for large values of N and M.

1–54. The Planck blackbody distribution law

$$\rho(\omega, T)\, d\omega = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3\, d\omega}{\exp(\beta\hbar\omega) - 1}$$

gives the blackbody radiation energy density between frequencies ω and $\omega + d\omega$. ($\hbar \equiv h/2\pi$, $\omega \equiv 2\pi\nu$, and $\varepsilon = h\nu = \hbar\omega$.) Substitute this into

$$\frac{E}{V} = \int_0^\infty \rho(\omega, T)\, d\omega$$

to derive the temperature dependence of E/V. Do this by expressing your result as a group of factors multiplying a dimensionless integral. You do not need to evaluate this integral.

1–55. Show that $e^x/(1 \pm e^x)^2$ is an even function of x.

1–56. The heat capacity of the Einstein model of a crystal is given by

$$C_V = 3Nk \left(\frac{\Theta_E}{T}\right)^2 \frac{e^{\Theta_E/T}}{(e^{\Theta_E/T} - 1)^2}$$

where Θ_E is the "characteristic temperature" of the crystal. Determine both the high- and low-temperature limiting expressions for the heat capacity. Do the same thing for the Debye model of crystals, in which

$$C_V = 9Nk\left(\frac{T}{\Theta_D}\right)^3 \int_0^{\Theta_D/T} \frac{x^4 e^x\,dx}{(e^x-1)^2}$$

where Θ_D is the Debye temperature of the crystal.

1–57. Recognizing it as a geometric series, sum the following series in closed form:

$$S = \sum_{n=0}^{\infty} e^{-}$$

Compare this result to

$$I = \int_0^{\infty} e^{-an}\,dn$$

Under what conditions are these two results the same?

1–58. One often encounters the gamma function in statistical thermodynamics. It was introduced by Euler as a function of x, which is continuous for positive values of x and which reduces to $n!$ when $x - n$, an integer. The gamma function $\Gamma(x)$ is defined by

$$\Gamma(x) = \int_0^{\infty} e^{-t}t^{x-1}\,dt$$

First show by integrating by parts that

$$\Gamma(x+1) = x\Gamma(x)$$

Using this, show that $\Gamma(n+1) = n!$ for n an integer. Show that

$$\Gamma(\tfrac{1}{2}) = \sqrt{\pi}$$

Evaluate $\Gamma(\frac{3}{2})$ using the recurrence formula $\Gamma(x+1) = x\Gamma(x)$. Lastly show that

$$\begin{aligned}\Gamma\left(n+\frac{1}{2}\right) &= \frac{1\cdot 3\cdots(2n-1)}{2^n}\,\Gamma\left(\frac{1}{2}\right)\\ &= \frac{(2n)!}{2^{2n}n!}\sqrt{\pi}\end{aligned}$$

For a discussion of the gamma function, see G. Arfken, *Mathematical Methods for Physicists*, 2nd ed. (New York: Academic, 1970).

1–59. We can derive Stirling's approximation from an asymptotic approximation to the gamma function $\Gamma(x)$. From the previous problem

$$\begin{aligned}\Gamma(N+1) = N! &= \int_0^{\infty} e^{-x}x^N\,dx\\ &= \int_0^{\infty} e^{Ng(x)}\,dx\end{aligned}$$

where $g(x) = \ln x - x/N$. If $g(x)$ possesses a maximum at some point, say x_0, then for large N, $\exp(Ng(x))$ will be extremely sharply peaked at x_0. Under this condition, the integral for $N!$ will be dominated by the contribution of the integrand from the point x_0. First show that $g(x)$ does, in fact, possess at maximum at the point $x_0 = N$. Expand $g(x)$ about this point, keeping terms only up to and including $(x-N)^2$ to get

$$g(x) \approx g(N) - \frac{(x-N)^2}{2} + \cdots$$

Why is there no linear term in $(x - N)$? Substitute this expression for $g(x)$ into the integral for $N!$ and derive the asymptotic formula

$$\ln N! \approx N \ln N - N + \ln(2\pi N)^{1/2}$$

1–60. Verify the energy conversion factors in Appendix A. (The one labeled "temperature" means that temperature required to give an energy equal to kT, where k is the Boltzmann constant.)

1–61. An integral that appears often in statistical mechanics and particularly in the kinetic theory of gases is

$$I_n = \int_0^\infty x^n e^{-ax^2}\, dx$$

This integral can be readily generated from two basic integrals. For even values of n, we first consider

$$I_0 = \int_0^\infty e^{-ax^2}\, dx$$

The standard trick to evaluate this integral is to square it, and then transform the variables into polar coordinates.

$$\begin{aligned} I_0^2 &= \int_0^\infty \int_0^\infty e^{-ax^2} e^{-ay^2}\, dx\, dy \\ &= \int_0^\infty \int_0^{\pi/2} e^{-ar^2} r\, dr\, d\theta \\ & \frac{\pi}{4a} \\ I_0 &= \frac{1}{2}\left(\frac{\pi}{a}\right)^{1/2} \end{aligned}$$

Using this result, show that for even n

$$I_n = \frac{1 \cdot 3 \cdot 5 \cdots (n-1)}{2(2a)^{n/2}} \left(\frac{\pi}{a}\right)^{1/2} \qquad n \text{ even}$$

For odd values of n, the basic integral I_1 is easy. Using I_1, show that

$$I_n = \frac{\Gamma\left(\frac{n+1}{2}\right)}{2a^{(n+1)/2}} \qquad n \text{ odd}$$

1–62. Show that a Gaussian distribution is extremely small beyond a few multiples of σ.

1–63. Another function that occurs frequently in statistical mechanics is the Riemann zeta function, defined by

$$\zeta(s) = \sum_{k=1}^{\infty} k^{-s}$$

First show that $\zeta(1) = \infty$, but that $\zeta(s)$ is finite for $s > 1$. Show that another definition of $\zeta(s)$ is

$$\zeta(s) = \frac{1}{\Gamma(s)} \int_0^\infty \frac{x^{s-1}\, dx}{(e^x - 1)}$$

that is, show that this is identical to the first definition. In addition, show that

$$\eta(s) = \sum_{k=1}^{\infty} (-1)^{k-1} k^{-s} = (1 - 2^{1-s})\zeta(s)$$

$$\lambda(s) = \sum_{k=0}^{\infty} (2k+1)^{-k} = (1 - 2^{-s})\zeta(s)$$

The evaluation of $\zeta(s)$ for integral s can be done using Fourier series, and some results are $\zeta(2) = \pi^2/6$ and $\zeta(4) = \pi^4/90$.
For a discussion of the Riemann zeta function, see G. Arfken, *Mathematical Methods for Physicists*, 2nd ed. (New York: Academic, 1970).

CHAPTER 2

THE CANONICAL ENSEMBLE

In this chapter we shall introduce the basic concepts and assumptions of statistical thermodynamics, and then apply them to a system which has fixed values of V and N and is in thermal equilibrium with its environment. We shall derive the fundamental connection between the quantum mechanical energy levels available to an N-body system and its thermodynamic functions. This link is effected by a function, called the partition function, which is of central importance in statistical thermodynamics. In Section 2–4 we discuss the relevance of the statistical thermodynamic equations to the second and third laws of classical thermodynamics.

2–1 ENSEMBLE AVERAGES

Our goal is to calculate thermodynamic properties in terms of molecular properties. Given the structure of the individual molecules of our system and the form of the intermolecular potential, we wish to be able to calculate thermodynamic properties, such as entropy and free energy. We shall do this first with respect to mechanical properties (such as pressure, energy, volume), which are quantum mechanical or classical mechanical quantities, and then we shall bring nonmechanical thermodynamic variables (such as entropy, free energy) into our discussion by appealing to the equations of thermodynamics. One useful distinction between mechanical and nonmechanical properties is that mechanical properties are defined without appealing to the concept of temperature, whereas the definitions of nonmechanical properties involve the temperature.

Consider some macroscopic system of interest, such as a liter of water or a salt solution. From a macroscopic point of view, we can completely specify a system by a few parameters, say the volume, concentration or density, and temperature. Regardless of the complexity of the system, it requires only a small number of parameters to describe it. From a microscopic point of view, on the other hand, there will be an enormous number of quantum states consistent with the fixed macroscopic properties.

We saw in Chapter 1 that the degeneracy of an isolated N-body system is of the order of 10^N for all but the very lowest energies. This means that the liter of water or the salt solution could be in any one of the order of 10^N possible quantum states. It would be impossible for us to ever determine which of the order of 10^N possible states the system is in. The state of the system must be known, however, in order to calculate a mechanical thermodynamic property, such as the pressure, since the values of that property in each of the possible quantum states would, in general, be different. Thus we are faced with what appears to be an impossible task.

It is at this point that we appeal to the work of Maxwell, Boltzmann, and particularly Gibbs. The modern (postquantum) version of their approach is that in order to calculate the value of any mechanical thermodynamic property (say, the pressure), one calculates the value of that mechanical property in each and every one of the quantum states that is consistent with the few parameters necessary to specify the system in a macroscopic sense. The average of these mechanical properties is then taken, giving each possible quantum state the same weight. We then *postulate* that this average mechanical property corresponds to a parallel thermodynamic property. For example, we postulate that the average energy corresponds to the thermodynamic energy and that the average pressure corresponds to the thermodynamic pressure. It turns out that the calculation of a mechanical property averaged over all the consistent quantum states can be readily performed. Before doing this, however, we shall introduce some concepts that will make this procedure clearer.

We first discuss the concept of an ensemble of systems, first introduced by Gibbs. An ensemble is a (mental or virtual) collection of a very large number of systems, say $\mathscr{A}$, each constructed to be a replica on a thermodynamic (macroscopic) level of the particular thermodynamic system of interest. For example, suppose the system has a volume V, contains N molecules of a single component, and is known to have an energy E. That is, it is an isolated system with N, V, and E fixed. Then the ensemble would have a volume $\mathscr{A}V$, contain $\mathscr{A}N$ molecules, and have a total energy $\mathscr{E} = \mathscr{A}E$. Each of the systems in this ensemble is a quantum mechanical system of N interacting atoms or molecules in a container of volume V The values of N and V, along with the force law between the molecules, are sufficient to determine the energy eigenvalues E_j of the Schrödinger equation along with their associated degeneracies $\Omega(E_j)$. These energies are the only energies available to the N-body system. Hence the fixed energy E must be one of these E_j's and, consequently, there is a degeneracy $\Omega(E)$. Note that there are $\Omega(E)$ different quantum states consistent with the only things we know about our macroscopic system of interest, namely, the values of N, V, and E. Although all the systems in the ensemble are identical from a thermodynamic point of view, they are not necessarily identical on a molecular level. So far we have said nothing about the distribution of the members of the ensemble with respect to the $\Omega(E)$ possible quantum states.

We shall further restrict our ensemble to obey the *principle of equal a priori probabilities*. That is to say, we require that each and every one of the $\Omega(E)$ quantum states is represented an equal number of times in the ensemble. Since we have no information to consider any one of the $\Omega(E)$ quantum states to be more important than any other, we must treat each of them equally, that is, we must utilize the principle of equal a priori probabilities. All of the $\Omega(E)$ quantum states are consistent with the given values of N, V, and E, the only information we have about the system. Clearly, the number of systems in the ensemble must be an integral multiple of $\Omega(E)$. The number of systems in an ensemble is a very large number and can be made arbitrarily large by simply doubling, tripling, and so on, the size of the ensemble. An alternative interpretation of the principle of equal a priori probabilities is that *an isolated system* (*N, V, and E fixed*) *is equally likely to be in any of its $\Omega(E)$ possible quantum states.*

We now define an ensemble average of a mechanical property as the average value of this property over all the members of the ensemble, utilizing the principle of equal a priori probabilities. We postulate that the ensemble average of a mechanical property can be equated to its corresponding thermodynamic property.

There are two complications in the above treatment that we should mention; neither of them, fortunately, is of any practical consequence. We have assumed that the isolated system that we have been using as an example has precisely the energy E. We know, however, from quantum mechanics that there always exists a small uncertainty ΔE in the value of E. For all thermodynamic purposes, this complication is completely inconsequential, and we shall therefore ignore it. The explanation of the other complication involves a greater knowledge of quantum mechanics than is generally required in this book. We have assumed that the systems of the ensemble are in one of the $\Omega(E)$ degenerate eigenstates having the eigenvalue E. The choice of these $\Omega(E)$ eigenfunctions, however, is somewhat arbitrary since any linear combination of these is also an eigenfunction with energy E. Moreover, a quantum mechanical system will, in general, not be in one of the $\Omega(E)$ selected states, but will be some linear combination of them. Thus we have tacitly assumed that any system with N, V, and E given will be a "pure state," whereas a system with N, V, and E given will most likely be in a "mixed state," that is, in a state described by a linear combination of the pure states we have chosen. In any event, this complication need not be considered, since the results do not differ appreciably from those obtained from the simpler and more naive point of view which we have presented above and now adopt.

Let us summarize this section by stating that we wish to calculate the ensemble average of some mechanical property, and then show that this can be set equal to the corresponding thermodynamic property. We have stated above that the calculation of the ensemble average is not difficult, and now we shall address ourselves to that problem. As Schrödinger says in his book:* "There is, essentially, only one problem in statistical thermodynamics, the distribution of a given amount of energy E over identical systems. Or perhaps better, to determine the distribution of an assembly of identical systems over the possible states in which the system can find itself, given that the energy of the assembly is a constant E."

So far, in this section, we have focused our attention on an ensemble whose members have N, V, and E fixed. This is called the *microcanonical ensemble* and is useful for theoretical discussions. For more practical applications, however, we consider not isolated systems, but those in which the temperature rather than the energy is fixed. The most commonly used ensemble in statistical thermodynamics is the *canonical ensemble*, in which the individual systems have N, V, and T fixed. The remainder of this chapter will deal with the canonical ensemble. There are many other types of ensembles, in fact, one for each set of thermodynamic variables that are used to specify an individual member of the ensemble. We shall discuss some of these other ensembles in the next chapter.

2–2 METHOD OF THE MOST PROBABLE DISTRIBUTION

Consider an experimental system with N, V, and T as its independent thermodynamic variables. We can mentally construct an ensemble of such systems in the following manner. We enclose each system in a container of volume V with walls that are heat

---

* E. Schrödinger, *Statistical Thermodynamics* (Cambridge: Cambridge University Press, 1952).

conducting but impermeable to the passage of molecules. The entire ensemble of systems is then placed in a very large heat bath at temperature T. When equilibrium is reached, the entire ensemble is at a uniform temperature T. Since the containing walls of each system are heat conducting, each and every system of the ensemble has the same fixed values of N, V, and T. Now, the entire ensemble is surrounded by thermal insulation, thus making the ensemble itself an isolated system with volume $\mathscr{A}V$, number of molecules $\mathscr{A}N$, and some total energy $\mathscr{E}$. (The actual value of $\mathscr{E}$ is not important.) Each of the $\mathscr{A}$ members of the canonical ensemble finds itself in a large heat bath at temperature T

Because each of the systems of the canonical ensemble is not isolated but is at a fixed temperature, the energy of each system is not fixed at any set value. Thus we shall have to consider the entire spectrum of energy states for each member of the canonical ensemble. Let the energy eigenvalues of the quantum states of a system be $E_1(N, V)$, $E_2(N, V)\ldots$, ordered such that $E_{j+1} \geq E_j$. It is important to understand here that any particular energy, say E_i, is repeated according to its degeneracy, that is, occurs $\Omega(E_i)$ times. Any particular system might be found in any of these quantum states. We shall show later that the *average energy* or the *probability* that some system has a certain energy depends upon the temperature; however, any of the set of energies $\{E_j\}$ is possible, and so must be considered.

We can specify a state of the entire ensemble by saying that a_1, a_2, a_3, of the systems are in states 1, 2, 3, ..., respectively, with energies E_1, E_2, E_3, ... Thus we can describe any one state of the ensemble by writing

State No.	1, 2, 3, $l\ldots$
Energy	$E_1, E_2, E_3, \ldots, E_l \ldots$
Occupation No.	$a_1, a_2, a_3, \ldots, a_l \ldots$

Occupation Number means the number of systems of the ensemble in that particular state. The set of occupation numbers is called a distribution. We shall often denote the set $\{a_j\}$ by $\mathbf{a}$.

Of course, the occupation numbers satisfy the two conditions:

$$\sum_j a_j = \mathscr{A} \tag{2–1}$$

$$\sum_j a_j E_j = \mathscr{E} \tag{2–2}$$

The first condition simply accounts for all the members of the ensemble, and the second represents the fact that the entire canonical ensemble is an isolated system, and hence has some fixed energy $\mathscr{E}$.

Since the canonical ensemble has been isolated from its surroundings by thermal insulation, we can apply the principle of equal a priori probabilities to this isolated system. In the form that we wish to use here, the principle of equal a priori probabilities says that every possible state of the canonical ensemble, that is, every distribution of occupation numbers a_1, a_2, consistent with Eqs. (2–1) and (2–2) is equally probable and must be given equal weight in performing ensemble averages.

The number of ways $W(\mathbf{a}) \equiv W(a_1, a_2, a_3, \ldots)$ that any particular distribution of the a_j's can be realized is the number of ways that $\mathscr{A}$ *distinguishable* objects can be arranged into groups, such that a_1 are in the first group, a_2 in the second, and so on [see Eq. (1–77)]:

$$W(\mathbf{a}) = \frac{\mathscr{A}!}{a_1!a_2!a_3!\cdots} = \frac{\mathscr{A}!}{\prod_k a_k!} \tag{2–3}$$

The systems are distinguishable since they are macroscopic systems, which we could, in principle, furnish with labels.

In general, there are very many distributions which are consistent with Eqs. (2–1) and (2–2). In any particular distribution, $a_j/\mathscr{A}$ is the fraction of systems or members of the canonical ensemble in the jth energy state (with energy E_j). The overall probability P_j that a system is in the jth quantum state is obtained by averaging $a_j/\mathscr{A}$ over all the allowed distributions, giving equal weight to each one according to the principle of equal a priori probabilities. Thus P_j is given by

$$P_j = \frac{\bar{a}_j}{\mathscr{A}} = \frac{1}{\mathscr{A}} \frac{\sum_{\mathbf{a}} W(\mathbf{a}) a_j(\mathbf{a})}{\sum_{\mathbf{a}} W(\mathbf{a})} \tag{2-4}$$

In Eq. (2–4), the notation $a_j(\mathbf{a})$ signifies that the value of a_j depends upon the distribution, and the summations are over all distributions that satisfy Eqs. (2–1) and (2–2). We shall later let $\mathscr{A} \to \infty$, but the ratio $\bar{a}_j/\mathscr{A}$ will remain finite since $\bar{a}_j \to \infty$ as well.

Given the probability that a system with fixed values of N, V, and T is in the jth quantum state, one can calculate the canonical ensemble average of any mechanical property from

$$\bar{M} = \sum_j M_j P_j \tag{2-5}$$

where M_j is the value of M in the jth quantum state. Thus the prescription for calculating the ensemble average of any mechanical property is given by Eqs. (2–4) and (2–5) and is, in principle, complete. The summations involved in Eq. (2–4), however, are very difficult to perform mathematically, and thus in practice Eqs. (2–4) and (2–5) are too complicated to use.

The fact that we can let $\mathscr{A} \to \infty$, however, allows us to appeal to the results of Section 1–5. We have seen there that multinomial coefficients, such as $W(\mathbf{a})$, are extremely peaked about their maximum value if all the variables a_j are large. In Eqs. (2–1) through (2–4), each of the a_j's can be made arbitrarily large since $\mathscr{A}$ can be made arbitrarily large. Thus we can use an argument here very similar to that used in Section 1–5. We need make only one modification or extension. We have shown that $W(\mathbf{a})$ is a maximum when all the a_j's are equal, under the one constraint Eq. (2–1). We have now an additional constraint Eq. (2–2) on the a_j's. So instead of peaking at the point at which all the a_j's are equal, it will peak at some other set of a_j's but the spread, nevertheless, will be arbitrarily small. We shall determine this set of a_j's shortly. Let us denote this distribution by $\mathbf{a}^* = \{a_j^*\}$.

The spread of $W(\mathbf{a})$ about its maximum value can be made arbitrarily narrow by taking the a_j, that is, $\mathscr{A}$, to be arbitrarily large. Thus the $W(\mathbf{a})$ in Eq. (2–4) at any set of a_j's other than the set $\mathbf{a}^*$, which maximizes $W(\mathbf{a})$, are completely negligible. We can replace the summations in Eq. (2–4) over all distributions by just one term, evaluated at $\mathbf{a}^*$. Thus we can write

$$P_j = \frac{1}{\mathscr{A}} \frac{\sum_{\mathbf{a}} W(\mathbf{a}) a_j(\mathbf{a})}{\sum_{\mathbf{a}} W(\mathbf{a})} = \frac{1}{\mathscr{A}} \frac{W(\mathbf{a}^*) a_j^*}{W(\mathbf{a}^*)} = \frac{a_j^*}{\mathscr{A}} \qquad (\lim a_j \to \infty) \tag{2-6}$$

where a_j^* is the value of a_j in that distribution that maximizes $W(\mathbf{a})$, that is, the most probable distribution. The name of this section, the method of the most probable distribution, is derived from Eq. (2–6). Comparing Eqs. (2–6) with (2–4), we have

$$P_j = \frac{\bar{a}_j}{\mathscr{A}} = \frac{a_j^*}{\mathscr{A}} \tag{2-7}$$

Thus, to calculate the probabilities to be used in ensemble averages, we need determine only that distribution $\mathbf{a}^*$ that maximizes $W(\mathbf{a})$ under the two constraints Eqs. (2–1) and (2–2). This is the problem to which we now turn.

As this is a problem of maximizing a function of many variables with given constraints on the variables, we have a direct application of Lagrange's method of undetermined multipliers. Following Section 1–5, the set of a_j's that maximizes $W(\mathbf{a})$, subject to Eqs. (2–1) and (2–2), is found from

$$\frac{\partial}{\partial a_j}\left\{\ln W(\mathbf{a}) - \alpha \sum_k a_k - \beta \sum_k a_k E_k\right\} = 0, \qquad j = 1, 2, \ldots \tag{2–8}$$

where α and β are the undetermined multipliers. Using Eq. (2–3) for $W(\mathbf{a})$ along with Stirling's approximation (which is exact here since each of the a_j's can be made arbitrarily large), one gets

$$-\ln a_j^* - \alpha - 1 - \beta E_j = 0 \qquad j = 1, 2, \ldots \tag{2–9}$$

or

$$a_j^* = e^{-\alpha'} e^{-\beta E_j} \qquad j = 1, 2, \ldots \tag{2–10}$$

where $\alpha' = \alpha + 1$. (See Problem 2–3.) This gives us the most probable distribution in terms of α and β. We now evaluate α' and β by using Eqs. (2–1) and (2–2) along with physical arguments.

2–3 THE EVALUATION OF THE UNDETERMINED MULTIPLIERS, α AND β

We can obtain an expression for α (or α') in terms of β by summing both sides of Eq. (2–10) over j and using Eq. (2–1) to get

$$e^{\alpha'} = \frac{1}{\mathscr{A}} \sum_j e^{-\beta E_j} \tag{2–11}$$

Equation (2–7) thus becomes

$$P_j = \frac{a_j^*}{\mathscr{A}} = \frac{e^{-\beta E_j(N,V)}}{\sum_j e^{-\beta E_j(N,V)}} \tag{2–12}$$

Substituting this into Eq. (2–5), with E_j taken to be the mechanical property, gives

$$\bar{E} = \bar{E}(N, V, \beta) = \frac{\sum_j E_j(N, V) e^{-\beta E_j(N,V)}}{\sum_j e^{-\beta E_j(N,V)}} \tag{2–13}$$

According to the postulate of the ensemble method of Gibbs, this average energy $\bar{E}(N, V, \beta)$ corresponds to the thermodynamic energy E.

The pressure is another important mechanical variable. When a system is in the state j, $dE_j = -p_j\, dV$ is the work done on the system when its volume is increased by dV (keeping the number of particles in the system fixed). Thus the pressure in the state j is given by

$$p_j = -\left(\frac{\partial E_j}{\partial V}\right)_N \tag{2–14}$$

The canonical ensemble average of p_j is

$$\bar{p} = \sum_j p_j P_j = -\frac{\sum_j \left(\frac{\partial E_j}{\partial V}\right) e^{-\beta E_j}}{\sum_j e^{-\beta E_j}} \tag{2-15}$$

We postulate that $\bar{p}$ corresponds to the thermodynamic pressure.

The sum in the denominator of Eqs. (2–13) and (2–15) occurs throughout the equations of the canonical ensemble. Let this be denoted by $Q(N, V, \beta)$:

$$Q(N, V, \beta) = \sum_j e^{-\beta E_j(N,V)} \tag{2-16}$$

We shall see that this function $Q(N, V, \beta)$ is the central function of the canonical ensemble.

We have made two connections with thermodynamics:

$$\begin{matrix} p \leftrightarrow \bar{p} \\ E \leftrightarrow \bar{E} \end{matrix} \quad \text{(ensemble postulate of Gibbs)}$$

Equation (2–13) gives E as a function of β. In principle, one could solve this equation for β as a function of E, but in practice this is not feasible. Fortunately β turns out to be a more convenient quantity than E, so much so that it is preferable to have E as a function of β rather than the inverse. We shall now evaluate β in two different ways.

We differentiate Eq. (2–13) with respect to V, keeping N and β fixed:

$$\left(\frac{\partial \bar{E}}{\partial V}\right)_{N,\beta} = -\bar{p} + \beta\overline{Ep} - \beta\bar{E}\bar{p} \tag{2-17}$$

In this equation,

$$\overline{Ep} = \frac{\sum_j p_j E_j e^{-\beta E_j}}{Q} = -\frac{\sum_j \left(\frac{\partial E_j}{\partial V}\right) E_j e^{-\beta E_j}}{Q}$$

and

$$\bar{E}\bar{p} = \frac{\sum_j E_j e^{-\beta E_j}}{Q} \cdot \frac{\sum_j p_j e^{-\beta E_j}}{Q}$$

Similarly, we can differentiate Eq. (2–15) to get

$$\left(\frac{\partial \bar{p}}{\partial \beta}\right)_{N,V} = \bar{E}\bar{p} - \overline{Ep} \tag{2-18}$$

From Eqs. (2–17) and (2–18) we get

$$\left(\frac{\partial \bar{E}}{\partial V}\right)_{N,\beta} + \beta\left(\frac{\partial \bar{p}}{\partial \beta}\right)_{N,V} = -\bar{p} \tag{2-19}$$

Note that $\bar{E}$ is a function of N, V, and β, whereas the E_j's are functions of N and V only. This is an important distinction, that should be clearly and completely understood.

Let us now compare Eq. (2–19) with the purely thermodynamic equation. (See Problem 1–29.)

$$\left(\frac{\partial E}{\partial V}\right)_{T,N} - T\left(\frac{\partial p}{\partial T}\right)_{N,V} = -p \tag{2-20}$$

which we rewrite in terms of $1/T$ instead of T:

$$\left(\frac{\partial E}{\partial V}\right)_{N,1/T} + \frac{1}{T}\left(\frac{\partial p}{\partial 1/T}\right)_{N,V} = -p \tag{2–21}$$

A comparison of Eq. (2–19) with Eq. (2–21) allows us to deduce that $\beta = \text{const}/T$. It is customary to write $\beta = 1/kT$, where k is a constant, whose value could possibly vary from substance to substance. We shall show now, however, that k has the same value for all substances, that is, k is a universal constant.

Consider two closed systems A and B, each having its own kind of particles and energy states, but in thermal contact with each other and immersed in a heat bath of temperature T. We now construct a canonical ensemble of systems AB (as shown in Fig. 2–1) representative of a thermodynamic AB system at temperature T and apply the method of the most probable distribution to the AB system. Let the number of molecules and volume of the A and B systems be N_A, V_A, and N_B, V_B, respectively, and let their energy states be denoted by $\{E_{jA}\}$ and $\{E_{jB}\}$. If a_j denotes the number of A systems in state E_{jA}, and b_j denotes the number of B systems in state E_{jB}, then the number of states of the AB ensemble with compound distribution $\{a_j\}$ and $\{b_j\}$ is

$$W(\mathbf{a}, \mathbf{b}) = \frac{\mathscr{A}!}{\prod_j a_j!} \cdot \frac{\mathscr{B}!}{\prod_k b_k!} \tag{2–22}$$

where $\mathscr{A}$ and $\mathscr{B}$ ($\mathscr{A} = \mathscr{B}$) are the number of A and B system, respectively. Equation (2–22) turns out to be a product of the separate A and B factors, because we can arrange the A systems over their possible quantum states independently of the B systems, and vice versa. The a_j's and b_j's must satisfy the three relations:

$$\begin{gathered}\sum_j a_j = \mathscr{A} \\ \sum_j b_j = \mathscr{B} = \mathscr{A} \\ \sum_j (a_j E_{jA} + b_j E_{jB}) = \mathscr{E}\end{gathered} \tag{2–23}$$

We now apply the method of the most probable distribution to Eqs. (2–22) and (2–23) to get Problem 2–9 for the simultaneous probability that the AB system has its A part in the ith quantum state and its B part in the jth quantum state:

$$P_{ij} = \frac{e^{-\beta E_{iA}}}{Q_A} \cdot \frac{e^{-\beta E_{jB}}}{Q_B} = P_{iA} P_{jB} \tag{2–24}$$

where

$$Q_A = \sum_k e^{-\beta E_{kA}} \quad \text{and} \quad Q_B = \sum_k e^{-\beta E_{kB}} \tag{2–25}$$

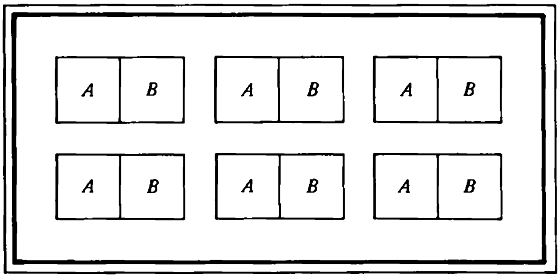

Figure 2–1. **Canonical ensemble of composite *AB* systems.**

Thus we have shown that two arbitrary systems in thermal contact have the same β. But we have seen from Eqs. (2–19) and (2–21) that $\beta = 1/kT$, and so the two systems must have the same value of k. Since the nature of the two systems is completely arbitrary, k must have the same value for all systems. Thus k is a universal constant and can therefore be evaluated using any convenient system. The most convenient is an ideal gas, and one can determine from the equation of state of an ideal gas [*cf.* Eq. (5–18)] that $k = 1.3806 \times 10^{-16}$ erg-deg^{-1}, where the temperature is in units of degrees Kelvin.

There is an alternative way to determine β which utilizes the fact that $1/T$ is an integrating factor of dq_{rev}. We shall present this argument here, since it will bring the nonmechanical property of entropy into our formalism.

The argument based around Fig. 2–1 shows that if two systems are in thermal contact at equilibrium, they have the same value of β. Since the two systems can be quite arbitrary, this implies that β must be some function of the temperature. We shall now show that $\beta\, dq_{\text{rev}}$ is an exact differential.

Consider the function $f = \ln Q$. We regard f as a function of β and all the E_j's:

$$f(\beta, E_1, E_2, \ldots) = \ln\left\{\sum_j e^{-\beta E_j}\right\} \tag{2–26}$$

The total derivative f is

$$df = \left(\frac{\partial f}{\partial \beta}\right)_{E_j\text{'s}} d\beta + \sum_k \left(\frac{\partial f}{\partial E_k}\right)_{\beta, E_i\text{'s}} dE_k \tag{2–27}$$

The partial derivatives occurring here are determined from Eq. (2–26) to be

$$\left(\frac{\partial f}{\partial \beta}\right)_{E_j\text{'s}} = \frac{-\sum_j E_j e^{-\beta E_j}}{Q} = -\bar{E}$$

$$\left(\frac{\partial f}{\partial E_k}\right)_{\beta, E_j\text{'s}} = \frac{-\beta e^{-\beta E_k}}{Q} = -\beta P_k$$

Thus Eq. (2–27) becomes

$$df = -\bar{E}\, d\beta - \beta \sum_j P_j\, dE_j$$

which can be written as

$$d(f + \beta\bar{E}) = \beta\left(d\bar{E} - \sum_j P_j\, dE_j\right) \tag{2–28}$$

We now subject the ensemble of systems to the following physical process. We change the volume of all the systems by dV, changing, of course, the E_j's for all of them alike in order to still have an ensemble of macroscopically identical systems. We also change the temperature of the ensemble by dT by coupling it with a large heat bath (of the same temperature), changing the temperature slightly and then isolating the ensemble from the heat bath.

If initially there were a_j systems of the ensemble in the energy state j with energy E_j, then $a_j\, dE_j$ is the work done on all these systems in changing the energy from E_j to $E_j + dE_j$. The total work done on the ensemble is $\sum a_j\, dE_j$ and $\sum_j P_j\, dE_j$ is the ensemble average reversible work that we do on the systems. And since $d\bar{E}$ is the average energy increase, the term enclosed in parentheses on the right-hand side of Eq. (2–28) is the average reversible heat supplied to a system. Thus Eq. (2–28) is

$$d(f + \beta\bar{E}) = \beta\, \delta q_{\text{rev}} \tag{2–29}$$

which says that $\beta\,\delta q_{\text{rev}}$ is the derivative of a state function, that is, that β is an integrating factor of δq_{rev}. One statement of the second law of thermodynamics says that β must be equal to constant/T, or $1/kT$.

The left-hand side of Eq. (2–29), therefore, must be dS/k, and so we can write that

$$S = \frac{\bar{E}}{T} + k \ln Q + \text{constant} \tag{2–30}$$

where the constant is independent of T and of the parameters (N, V, and so on) on which the E_j's depend. Since thermodynamics deals with ΔS only, the constant will always drop out of any calculations of entropy changes for chemical and/or physical changes. We shall, therefore, set this constant to zero and discuss the implications of this at the end of the chapter.

In the above argument that β was an integrating factor of δq_{rev}, we used the fact that the average work done on a system was $\sum_j P_j\, dE_j$. We do work, then, by changing the energies slightly, but keeping the population of these states fixed (the P_j's do not change). A molecular interpretation of thermodynamic work, then, is a change in the quantum mechanical energy states of the system, keeping the population over them fixed. That a molecular interpretation of the absorption of heat is the inverse of this can be seen from

$$\begin{aligned} d\bar{E} &= \sum_j E_j\, dP_j + \sum_j P_j\, dE_j \\ &= \delta q_{\text{rev}} - \delta\omega_{\text{rev}} \end{aligned}$$

Thus when a small quantity of heat is absorbed from the surroundings, the energy states of the system do not change (N and V are fixed), but the population of these states does.

2–4 THERMODYNAMIC CONNECTION

We now complete the connection between thermodynamics and the canonical ensemble. Equation (2–13) for $\bar{E}$ can be written as (see Problem 2–10):

$$E = kT^2 \left(\frac{\partial \ln Q}{\partial T}\right)_{N,V} \tag{2–31}$$

and we can also easily derive (see Problem 2–10)

$$p = kT \left(\frac{\partial \ln Q}{\partial V}\right)_{N,T} \tag{2–32}$$

from Eq. (2–15). Equation (2–30) is an equation for the entropy S in terms of Q:

$$S = kT \left(\frac{\partial \ln Q}{\partial T}\right)_{N,V} + k \ln Q \tag{2–33}$$

We have E, p, and S now as functions of Q, and so it is possible to derive expressions for all the thermodynamic functions in terms of Q. The function Q is the central statistical thermodynamic function of the canonical ensemble (N, V, and T fixed) and is called the *canonical* (*ensemble*) *partition function*:

$$Q(N, V, T) = \sum_j e^{-E_j(N,V)/kT} \tag{2–34}$$

The partition function serves as a bridge between the quantum mechanical energy states of a macroscopic system and the thermodynamic properties of that system. If we can obtain Q as a function of N, V, and T, we can calculate thermodynamic properties in terms of quantum mechanical and molecular parameters. Although the E_j's are the energy states of an N-body system and consequently appear to be unobtainable in practice, we shall see that in a great many cases, we shall be able to reduce the N-body problem to a one-body, two-body, three-body problem, and so on, or approximate the system by classical mechanics. Both of these routes turn out to be very useful. For now, however, we need only assume that there is such a set of energies.

We can derive an equation for the Helmholtz free energy A in terms of Q by using Eqs. (2–31) and (2–33) along with the fact that $A = E - TS$. The result is

$$A(N, V, T) = -kT \ln Q(N, V, T) \tag{2–35}$$

Notice that of all the thermodynamic functions, it is A that is directly proportional to $\ln Q(N, V, T)$, and that A is the thermodynamic potential whose natural independent variables are those of the canonical ensemble. Equation (2–35) can be considered to be the most important connection between thermodynamics and the canonical partition function, since it is possible to derive many equations starting with its differential form (see Problem 2–11). Table 3–1 contains a summary of the formulas of the canonical ensemble.

In this chapter, we have developed the connection between thermodynamics and the quantum mechanical states available to a macroscopic system characterized by N, V, and T. This connection can be summarized by Eq. (2–35). Before concluding this chapter, we shall discuss the second and third laws of thermodynamics from a statistical thermodynamic point of view. A statement of the second law of thermodynamics for closed, isothermal systems is that $\Delta A < 0$ for a spontaneous process. We wish to derive this inequality starting with Eq. (2–35). To do this, we first write Eq. (2–34) in a slightly different form

Consider Eq. (2–34) for $Q(N, V, T)$. The summation is over all the possible quantum states of the N-body system. In carrying out the summation, a particular value of $\exp(-E_j/kT)$ will occur $\Omega(E_j)$ times, where $\Omega(E_j)$ is the degeneracy. Instead of listing $\exp(-E_j/kT)$ $\Omega(E_j)$ times, we could simply write $\Omega(E_j)\exp(-E_j/kT)$, and then sum over different values of E. If we do this, Eq. (2–34) is

$$Q(N, V, T) = \sum_E \Omega(N, V, E)e^{-E(N, V)/kT} \tag{2–36}$$

where we have dropped the no longer necessary j subscript of E_j. In Eq. (2–34), we sum over the *states* of the system. In Eq. (2–36) we sum over *levels*. Equation (2–36) is a more useful form for discussing the second law of thermodynamics.

Consider a typical spontaneous processes, such as the expansion of a gas into a vacuum. Figure 2–2 shows the initial and final states of such a process. For simplicity, we consider the entire system to be isolated. Initially the gas might be confined to one half of the container. After removing the barrier, the gas occupies the entire container.

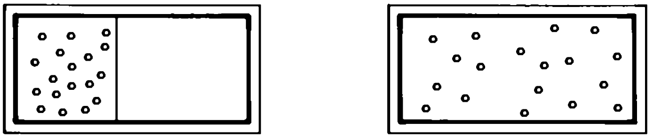

Figure 2–2. **The initial and final states of the expansion of a gas into a vacuum.**

Equation (1–36) for $\Omega(N, V, E)$ of an ideal gas shows that the number of states is proportional to V^N. For the process illustrated in Fig. 2–2, the gas goes from a thermodynamic state of energy E, number of particles N, and volume $V/2$ to one with the same energy E (the system is isolated), the same number of particles N, but with volume V Thus according to Eq. (1–36), the number of quantum states available or accessible to the system is increased.

Another example of a spontaneous process is the following. Initially we have an isolated system containing a mixture of hydrogen and oxygen gases. Although hydrogen and oxygen react to form water, in the absence of a catalyst the reaction is so slow that we can ignore it. Since the rate of this reaction (uncatalyzed) is very slow compared to any thermodynamic measurement, we can consider the mixture of hydrogen and oxygen to be simply a mixture of two gases in equilibrium. If we now add a small amount of catalyst to the system, the hydrogen and oxygen will readily form water, so that the system contains hydrogen, oxygen, and water. Thus the addition of a small amount of catalyst makes all the energy states associated with water molecules available or accessible to the system, and, hence, the system proceeds spontaneously to populate these states. Since the originally accessible states are still accessible (there is still some hydrogen and oxygen in the system), this spontaneous process is associated with an increase in the number of states accessible to the system by the removal of some constraint. In this case, the constraint was a high activation energy barrier, which was removed by the addition of the catalyst.

Both of the spontaneous processes that we have discussed occurred, because some restraint, inhibition, or barrier was removed which made additional quantum states accessible to the system. In general, any spontaneous process in an isolated system can be viewed in this manner. The removal of some constraint allows a greater number of quantum states to be accessible to the system, thus the "flow" of the system into these states is observed as a spontaneous process.

The above discussion is limited to isolated systems. In order to discuss the condition $\Delta A < 0$, we must now consider isothermal processes. When a system is in a heat bath rather than isolated, we must include all possible energy states or levels of the system. When a restraint is removed, the number of accessible quantum states of each and every energy E cannot decrease, and will usually increase, since the original states are still available. Thus we have that $\Omega_2(N, V, E) \geq \Omega_1(N, V, E)$ for all E, where the subscripts 1 and 2 denote the initial and final states, respectively. We now use this inequality along with Eq. (2–36) to show that $\Delta A < 0$. Since no term can be negative and many are positive, we have

$$Q_2 - Q_1 = \sum_E \{\Omega_2(N, V, E) - \Omega_1(N, V, E)\}e^{-E/kT} > 0 \tag{2–37}$$

In Eq. (2–37) we sum over all the levels available to the final state. It follows immediately from the inequality in Eq. (2–37) that

$$\Delta A = A_2 - A_1 = -kT \ln \frac{Q_2}{Q_1} < 0 \tag{2–38}$$

for a spontaneous isothermal process, and thus we have written the second law of thermodynamics in terms of Eq. (2–35).

Lastly, we consider the implications of putting the "constant" of Eq. (2–30) equal to zero. We shall see that this gives us a statistical thermodynamic version of the third law of thermodynamics. If we write Eq. (2–33) for S more explicitly, we get

$$S = k \ln \sum_j e^{-E_j/kT} + \frac{1}{T}\frac{\sum_j E_j e^{-E_j/kT}}{\sum_j e^{-E_j/kT}} \tag{2–39}$$

We wish to study the behavior of this equation as $T \to 0$. Assume for generality that the first n states have the same energy ($E_1 = E_2 = \cdots = E_n$) and that the next m states have the same energy ($E_{n+1} = E_{n+2} = \cdots = E_{n+m}$), and so on. Then in the limit of small T, Eq. (2–39) becomes (see Problem 2–19):

$$S = k \ln n + \frac{km}{n} e^{-(E_{n+1}-E_1)/kT} + \frac{m}{nT}(E_{n+1} - E_1)e^{-(E_{n+1}-E_1)/kT}$$

and so

$$\lim_{T\to 0} S = k \ln n \tag{2–40}$$

Thus as $T \to 0$, S is proportional to the logarithm of the degeneracy of the lowest level. Unless n is very large, Eq. (2–40) says that S is practically zero. For example, if the system were a gas of N-point particles, and the degeneracy of the lowest level were of the order of N, $k \ln N$ would be practically zero compared to a typical order of magnitude of the entropy, namely, Nk. Thus setting the "constant" of Eq. (2–30) equal to zero is equivalent to adopting the convention that the entropy of most systems is zero at the absolute zero of temperature [*cf.* Eq. (1–44)].

ADDITIONAL READING

General

Andrews, F. C. 1963. *Equilibrium statistical mechanics.* New York: Wiley. Chapters 9–11.

Hill, T. L. 1960. *Statistical thermodynamics.* Reading, Mass.: Addison-Wesley. Sections 1–1 through 1–5.

Kestin, J., and Dorfman, J. R. 1971. *A course in statistical thermodynamics.* New York: Academic. Sections 5–1 through 5–3 and 5–11 through 5–17.

Kittel, C. 1969. *Thermal physics.* New York: Wiley. Chapters 1–4.

Knuth, E. L. 1966. *Statistical thermodynamics.* New York: McGraw-Hill. Chapters 2, 4, and 5.

Mayer, J. E., and Mayer, M. G. 1940. *Statistical mechanics.* New York: Wiley. Chapters 3, 4, and 10.

Schrödinger, E. 1952. *Statistical thermodynamics.* Cambridge: Cambridge University Press. Chapters 1–3.

PROBLEMS

2–1. From statistical mechanics we have shown

$$\left(\frac{\partial \bar{E}}{\partial V}\right)_{N,\beta} + \beta\left(\frac{\partial \bar{p}}{\partial \beta}\right)_{N,V} = -\bar{p}$$

and from the thermodynamics we have

$$\left(\frac{\partial E}{\partial V}\right)_{N,T} - T\left(\frac{\partial p}{\partial T}\right)_{N,V} = -p$$

Why can't β be linearly proportional to the temperature? That is, $\beta = \text{constant} \times T$.

2–2. To investigate the replacement of $\bar{n}_j$ by n_j^*, that is, the replacement of the average number of systems in state j by the most probable number in state j, consider the simple example in which $\Omega(\mathbf{n})$ is just a binomial distribution

$$\Omega(\mathbf{n}) = \frac{n!}{n_1!\,(n - n_1)!}$$

and actually calculate n_1^* and $\bar{n}_1$. Hint: Recall that

$$(1+x)^n = \sum_{n_1=0}^{n} \frac{n!x^{n_1}}{n_1!(n-n_1)!}$$

2–3. Show that Eq. (2–9) follows from Eq. (2–8). Note that in deriving this result, we have written ln $W(\mathrm{a})$ as $\mathscr{A} \ln \mathscr{A} - \mathscr{A} - \sum_j a_j \ln a_j + \mathscr{A}$ and have considered $\mathscr{A}$ to be a constant. Show that Eq. (2–10) is independent of this assumption, that is, derive Eq. (2–10) treating $\mathscr{A}$ as $\sum_j a_j$.

2–4. Starting with Eq. (2–31), prove that the Boltzmann constant k must be positive, using the fact that the heat capacity C_V is always positive.

2–5. Show that the entropy can be written as

$$S = -k \sum_j P_j \ln P_j$$

where P_j is given by Eq. (2–12).

2–6. Maximize the function defined as "information" in information theory.

$$I = \sum_j P_j \ln P_j$$

subject to the two constraints

$$\sum_j P_j = 1$$

and

$$\sum_j E_j P_j = E = \text{fixed}$$

Compare this result to that of Problem 1–51.

2–7. Obtain the most probable distribution of N molecules of an ideal gas contained in two equal and connected volumes at the same temperature by minimizing the Helmholtz free energy for the two systems.

2–8. Differentiate Eq. (2–16) with respect to β to derive Eq. (2–13).

2–9. Derive Eq. (2–24).

2–10. Derive Eqs. (2–31) and (2–32).

2–11. Derive Eqs. (2–31) through (2–33) by starting with $A = -kT \ln Q$.

2–12. We can derive Eq. (2–36) directly by the method of Lagrange multipliers. We label the *levels* rather than the states by a subscript l. The degeneracy of the lth level, whose energy is E_l, is Ω_l. The number of ways of distributing systems over levels, with degeneracy Ω_l, is

$$W(\mathrm{a}) = \frac{\mathscr{A}!}{\prod_l a_l} \prod_l \Omega_l^{a_l}$$

where a_l is the number of systems in the lth level. Maximize this, subject to the constraints

$$\sum_l a_l = \mathscr{A}$$

$$\sum_l a_l E_l = \mathscr{E}$$

to get

$$a_l^* = \frac{\Omega_l e^{-E_l/kT}}{\sum_l \Omega_l e^{-E_l/kT}}$$

2–13. Show that for a particle confined to a cube of length a that

$$p_j = \frac{2}{3} \frac{E_j}{V}$$

By taking the ensemble average of both sides, we have

$$\bar{p}=\frac{2}{3}\frac{\bar{E}}{V}$$

If we use the fact that $\bar{E}=\frac{3}{2}NkT$ (to be proved in Chapter 5), we get the ideal gas equation of state.

2–14. We shall show in Chapter 5 that the partition function of a monatomic ideal gas is

$$Q(N, V, T)=\frac{1}{N!}\left(\frac{2\pi mkT}{h^2}\right)^{3N/2}V^N$$

Derive expressions for the pressure and the energy from this partition function. Also show that the ideal gas equation of state is obtained if Q is of the form $f(T)V^N$, where $f(T)$ is any function of temperature.

2–15. In Chapter 11 we shall approximate the partition function of a crystal by

$$Q=\left(\frac{e^{-h\nu/2kT}}{1-e^{-h\nu/kT}}\right)^{3N}e^{U_0/kT}$$

where $h\nu/k\equiv\Theta_E$ is a constant characteristic of the crystal, and U_0 is the sublimation energy of the crystal. Calculate the heat capacity from this simple partition function and show that at high temperatures, one obtains the law of Dulong and Petit, namely, that $C_V\rightarrow 3Nk$ as $T\rightarrow\infty$.

2–16. In Chapter 13 of this author's textbook *Statistical Thermodynamics*, it is shown that the partition function of an ideal gas of diatomic molecules in an external electric field $\mathcal{E}$ is

$$Q(N, V, T, \mathcal{E})=\frac{[q(V, T, \mathcal{E})]^N}{N!}$$

where

$$q(V, T, \mathcal{E})=V\left(\frac{2\pi mkT}{h^2}\right)^{3/2}\left(\frac{8\pi^2 IkT}{h^2}\right)\frac{e^{-h\nu/2kT}}{(1-e^{-h\nu/kT})}\left(\frac{kT}{\mu\mathcal{E}}\right)\sinh\left(\frac{\mu\mathcal{E}}{kT}\right)$$

Here I is the moment of inertia of the molecule; ν is its fundamental vibrational frequency; and μ is its dipole moment. Using this partition function along with the thermodynamic relation,

$$dA=-S\,dT-p\,dV-M\,d\mathcal{E}$$

where $M=N\bar{\mu}$, where $\bar{\mu}$ is the average dipole moment of a molecule in the direction of the external field $\mathcal{E}$, show that

$$\bar{\mu}=\mu\left[\coth\left(\frac{\mu\mathcal{E}}{kT}\right)-\frac{kT}{\mu\mathcal{E}}\right]$$

Sketch this result versus $\mathcal{E}$ from $\mathcal{E}=0$ to $\mathcal{E}=\infty$ and interpret it.

2–17. In Chapter 14 we shall derive an *approximate* partition function for a dense gas, which is of the form

$$Q(N, V, T)=\frac{1}{N!}\left(\frac{2\pi mkT}{h^2}\right)^{3N/2}(V-Nb)^N e^{aN^2/VkT}$$

where a and b are constants that are given in terms of molecular parameters. Calculate the equation of state from this partition function. What equation of state is this? Calculate the thermodynamic energy and the heat capacity and compare it to Problem 1–30.

2–18. From electrostatics, the displacement vector D is given by $D=\mathcal{E}+4\pi P$, where $\mathcal{E}$ is the electric field, and P is the polarization, i.e., the total dipole moment M per unit volume.

The dielectric constant ε is defined by $D = \varepsilon\mathscr{E}$. In the simple case of a parallel plate capacitor, D is the field produced by a set of charges on the plates, and so we can consider it to be external field; M is the total moment (both permanent and induced) of the substance between the plates; and $\mathscr{E}$ is the field between the plates, that is, the force that an infinitesimal change would feel. If only a vacuum existed between the plates, D and $\mathscr{E}$ would be the same. A real substance modifies D such that $D = \mathscr{E} + 4\pi(M/V)$, where V is the volume. Since D, $\mathscr{E}$, and M are all in the same direction (at least for simple fluids), D must be $\geq \mathscr{E}$, which says that $\varepsilon \geq 1$.

When an external electric field D is present, the first law of thermodynamics becomes

$$dE = T\,dS - p\,dV - M\,dD + \mu\,dN$$

Problem: Describe *concisely* how one would calculate ε (at least in principle) from statistical mechanics.

2–19. Derive equation 2–40.

CHAPTER 3

OTHER ENSEMBLES AND FLUCTUATIONS

In Chapter 2 we considered an ensemble in which N, V, and T are held fixed for each system. This ensemble is one of many possible ensembles that can be constructed. For example, if we allow the walls of the containers to be permeable to molecular transport, N is no longer fixed for each system, and we no longer have a canonical ensemble. The ensemble in this case is called a grand canonical ensemble and is discussed in Section 3–1. In Section 3–2 we discuss two other ensembles that are often used in statistical thermodynamics: the microcanonical ensemble, in which N, V, and E are fixed, and the isothermal-isobaric ensemble, in which N, T, and p are fixed. The last section, Section 3–3, is devoted to an investigation of fluctuations in statistical thermodynamics.

One of our basic assumptions is that the ensemble average of a mechanical property can be equated to the corresponding thermodynamic function; hence it is important that we investigate the expected spread about the mean value. We show in Section 3–3 that for macroscopic systems the probability distribution of observing some mechanical property is a very narrow Gaussian distribution whose mean is the ensemble average. One important deduction from this result is that the various ensembles are essentially equivalent and that one can choose to work with a partition function on the basis of mathematical convenience rather than on the basis of which thermodynamic variables are used to specify the system of interest.

3–1 GRAND CANONICAL ENSEMBLE

In the previous chapter we treated the canonical ensemble, in which each system is enclosed in a container whose walls are heat conducting, but impermeable to the passage of molecules. The entire ensemble is placed in a heat bath at temperature T until equilibrium is reached, and then is isolated from its surroundings. Each system of the ensemble is specified by N, V, and T. In this section we shall treat a grand canonical ensemble. In a grand canonical ensemble, each system is enclosed in a container whose walls are both heat conducting and permeable to the passage of molecules. The number of molecules in a system, therefore, can range over all possible values, that is, each system is open with respect to the transport of matter. We construct a grand canonical ensemble by placing a collection of such systems in a large heat bath at temperature T and a large reservoir of molecules. After equilibrium is reached, the entire ensemble is isolated from its surroundings. Since the entire ensemble is at equilibrium with respect to the transport of heat and matter, each system is specified by V, T, and μ, where μ is the chemical potential. (If there is more than one component, the chemical potential of each component is the same from system to system.) Figure 3-1 shows a schematic picture of a grand canonical ensemble.

We proceed now in the same manner as in the treatment of the canonical ensemble. In this case, however, we must specify a system not only by which quantum state it is in but also by the number of molecules in the system. For each value of N, there is a set of energy states $\{E_{Nj}(V)\}$. We let a_{Nj} be the number of systems in the ensemble that contain N molecules and are in the state j. Each value of N has a particular set of levels associated with it, so we first specify N and then j. The set of occupation numbers $\{a_{Nj}\}$ is a distribution. By the postulate of equal a priori probabilities, we assume that all states associated with all possible distributions are to be given equal weight or equal probability of occurrence in the ensemble. Each possible distribution must satisfy the following three conditions:

$$\sum_N \sum_j a_{Nj} = \mathscr{A} \tag{3-1}$$

$$\sum_N \sum_j a_{Nj} E_{Nj} = \mathscr{E} \tag{3-2}$$

$$\sum_N \sum_j a_{Nj} N = \mathscr{N} \tag{3-3}$$

The three symbols $\mathscr{A}$, $\mathscr{E}$, and $\mathscr{N}$ denote the number of systems in the ensemble, the total energy of the ensemble (the ensemble is isolated), and the total number of molecules in the ensemble.

For any possible distribution, the number of states is given by

$$W(\{a_{Nj}\}) = \frac{\mathscr{A}!}{\prod_N \prod_j a_{Nj}!} \tag{3-4}$$

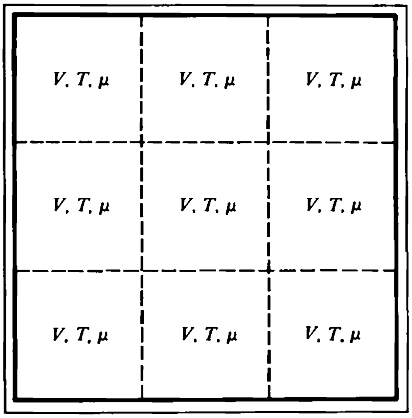

Figure 3-1. **A schematic picture of a grand canonical ensemble. Each system has a fixed volume and temperature, but is open with respect to molecular transport.**

As in the treatment of the canonical ensemble, the distribution that maximizes W subject to the appropriate constraints completely dominates all others. Thus we maximize Eq. (3–4) under the constraints of Eqs. (3–1) through (3–3), respectively, and we get (see Problem 3–1)

$$a_{Nj}^* = e^{-\alpha} e^{-\beta E_{Nj}(V)} e^{-\gamma N} \tag{3–5}$$

As before, the parameter α is easily determined in terms of the other parameter(s). We sum both sides of Eq. (3–5) over N and j and use Eq. (3–1) to get

$$P_{Nj}(V, \beta, \gamma) = \frac{a_{Nj}^*}{\mathscr{A}} = \frac{e^{-\beta E_{Nj}(V)} e^{-\gamma N}}{\sum_N \sum_j e^{-\beta E_{Nj}(V)} e^{-\gamma N}} \tag{3–6}$$

where $P_{Nj}(V, \beta, \gamma)$ is the probability that any randomly chosen system contains N molecules and be in the jth energy state, with energy $E_{Nj}(V)$.

The averages of the mechanical properties E, P, and N are

$$\begin{aligned} \bar{E}(V, \beta, \gamma) &= \frac{1}{\Xi} \sum_N \sum_j E_{Nj}(V) e^{-\beta E_{Nj}(V)} e^{-\gamma N} \\ &= -\left(\frac{\partial \ln \Xi}{\partial \beta}\right)_{V, \gamma} \end{aligned} \tag{3–7}$$

$$\begin{aligned} \bar{p}(V, \beta, \gamma) &= \frac{1}{\Xi} \sum_N \sum_j \left(-\frac{\partial E_{Nj}}{\partial V}\right) e^{-\beta E_{Nj}(V)} e^{-\gamma N} \\ &= \frac{1}{\beta}\left(\frac{\partial \ln \Xi}{\partial V}\right)_{\beta, \gamma} \end{aligned} \tag{3–8}$$

$$\begin{aligned} \bar{N}(V, \beta, \gamma) &= \frac{1}{\Xi} \sum_N \sum_j N e^{-\beta E_{Nj}(V)} e^{-\gamma N} \\ &= -\left(\frac{\partial \ln \Xi}{\partial \gamma}\right)_{V, \beta} \end{aligned} \tag{3–9}$$

where

$$\Xi(V, \beta, \gamma) = \sum_N \sum_j e^{-\beta E_{Nj}(V)} e^{-\gamma N} \tag{3–10}$$

We now determine β and γ. In our treatment of the canonical ensemble, one of the methods used to determine β was to derive an equation that related $(\partial \bar{E}/\partial V)_{N,\beta}$ to $(\partial \bar{p}/\partial \beta)_{N,V}$ and to compare this with a purely thermodynamic equation relating $(\partial E/\partial V)_{N,T}$ to $(\partial p/\partial T)_{N,V}$ [*cf.* Eqs. (2–17) to (2–21)]. This comparison suggested that β was proportional to $1/T$. We then showed that any two systems at the same temperature have the same value of β, thus proving that $\beta = 1/kT$, where k is a universal constant. We can do the same thing here (Problem 3–2), but it is not necessary.

A grand canonical ensemble can be considered to be a collection of canonical ensembles in thermal equilibrium with each other but with all possible values of N. Each of the systems has the same value of β, regardless of the number of molecules it contains. That β has the same value as in the canonical ensemble can be seen by imagining that we suddenly make the walls of the containers impermeable to the molecules but still heat conducting. This gives us a collection of canonical ensembles with V, N, and T fixed, and the arguments of Chapter 2 can be used to show that $\beta = 1/kT$

The value of γ can be found by using the same method that we used in Chapter 2 to show that β was an integrating factor of δq_{rev}. Consider the function

$$f(\beta, \gamma, \{E_{Nj}(V)\}) = \ln \Xi = \ln \sum_N \sum_j e^{-\beta E_{Nj}(V)} e^{-\gamma N}$$

As the notation indicates, we regard f to be a function of β, γ, and the E_{Nj}'s. The total derivative of f is

$$df = \left(\frac{\partial f}{\partial \beta}\right)_{\gamma, \{E_{Nj}\}} d\beta + \left(\frac{\partial f}{\partial \gamma}\right)_{\beta, \{E_{Nj}\}} d\gamma + \sum_N \sum_j \left(\frac{\partial f}{\partial E_{Nj}}\right)_{\beta, \gamma, E_{Nj},s} dE_{Nj}$$

Using Eqs. (3–7) through (3–10), we have

$$df = -\bar{E}\, d\beta - \bar{N}\, d\gamma - \beta \sum_N \sum_j P_{Nj}\, dE_{Nj}$$

The last term here is the ensemble average reversible work done by the systems. For simplicity, we assume only $p-V$ work to get

$$df = -\bar{E}\, d\beta - \bar{N}\, d\gamma + \beta \bar{p}\, dV$$

Paralleling our development in Chapter 2, we add $d(\beta\bar{E}) + d(\gamma\bar{N})$ to both sides of this equation:

$$d(f + \beta\bar{E} + \gamma\bar{N}) = \beta\, d\bar{E} + \beta\bar{p}\, dV + \gamma\, d\bar{N}$$

If we compare this to the purely thermodynamic equation

$$T\, dS = dE + p\, dV - \mu\, dN$$

and use the fact that $\beta = 1/kT$, we can conclude that

$$\gamma = \frac{-\mu}{kT} \tag{3–11}$$

$$S = \frac{\bar{E}}{T} - \frac{\bar{N}\mu}{T} + k \ln \Xi \tag{3–12}$$

In Eq. (3–12), we have set the constant of integration equal to zero in accord with the third law of thermodynamics (see Problem 3–5).

We have now brought the entropy, a nonmechanical property, into our discussion. Equation (3–12), along with Eqs. (3–6) through(3–9), allows us to express any thermodynamic function of interest in a grand canonical ensemble in terms of $\Xi(V, T, \mu)$. This function is called the *grand (canonical ensemble) partition function*:

$$\Xi(V, T, \mu) = \sum_N \sum_j e^{-E_{Nj}(V)/kT} e^{\mu N/kT} \tag{3–13}$$

As the canonical partition function is the connection between thermodynamics and statistical thermodynamics for closed, isothermal systems (N, V, and T fixed), the grand partition function serves as the link for open, isothermal systems (V, T, and μ fixed). If we can determine Ξ for a system, we can calculate its thermodynamic properties.

By summing over j for fixed N in Eq. (3–13), we see that it is possible to write Ξ in the form

$$\Xi(V, T, \mu) = \sum_N Q(N, V, T) e^{\mu N/kT} \tag{3–14}$$

The term $e^{\mu/kT}$ is often denoted by λ. Since $\mu = kT \ln \lambda$, λ is an absolute activity, for the difference in chemical potentials between two states is given by $\Delta\mu = kT \ln(a_2/a_1)$, where a_1 and a_2 are activities.

Since we take the number of systems in an ensemble to be arbitrarily large, the number of particles in an ensemble becomes arbitrarily large, and hence the possible number of particles in any one system can approach infinity. Therefore the summation in Eq. (3–14) can be taken from 0 to ∞:

$$\Xi(V, T, \mu) = \sum_{N=0}^{\infty} Q(N, V, T)\lambda^N \tag{3–15}$$

Even though it may appear from Eq. (3–15) that Ξ would be more difficult to obtain than Q, it actually turns out in many problems that Ξ is easier to obtain, since the constraint of constant N is often mathematically awkward. This constraint can be avoided by using a grand partition function, that is, by summing over all values of N (see Section 4–2). Furthermore, there are many systems in which the many-body problem can be reduced to a one-body, two-body problem, and so on. In these cases, the grand partition function is particularly useful.

To complete our discussion of the grand canonical ensemble, we shall show that pV is the thermodynamic characteristic function of $\ln \Xi$. To see this, compare Eq. (3–12) with the thermodynamic equation

$$G = \mu N = E + pV - TS$$

Thus we have

$$pV = kT \ln \Xi(V, T, \mu) \tag{3–16}$$

Problem 1–37 shows that pV is the thermodynamic function whose natural variables are V, T, and μ. Equations (3–8), (3–9), and (3–12) can be derived from Eq. (3–16) and the thermodynamic equation $d(pV) = S\,dT + N\,d\mu + p\,dV$ (see Problem 3–6). Table 3–1 summarizes the formulas of the grand canonical ensemble.

3–2 OTHER ENSEMBLES

We could go on to consider other ensembles. For example, we could construct an ensemble of systems in which the containing walls of each system are heat conducting and flexible, so that each system of the ensemble is described by N, T, and p. The constraints would be on the total energy and total volume of the ensemble, and the partition function would turn out to be (see Problem 3–9):

$$\Delta(N, T, p) = \sum_{E}\sum_{V} \Omega(N, V, E)e^{-E/kT}e^{-pV/kT} \tag{3–17}$$

whose characteristic thermodynamic function is the Gibbs free energy, that is,

$$G = -kT \ln \Delta(N, T, p) \tag{3–18}$$

Equation (3–17) is called the isothermal-isobaric partition function. Notice that the natural variables of G are N, T, and p, the variables associated with this ensemble.

If we compare Eq. (3–17) with the two other partition functions that we have derived [Eqs. (2–34) and (3–13)], we see that all three can be obtained by starting with $\Omega(N, V, E)$, multiplying by some appropriate exponential, and summing over one or two of the variables N, V, and E. In a sense, $\Omega(N, V, E)$ is fundamental to all ensembles and, in fact, is itself the partition function for conceptually the most simple ensemble, the one representative of isolated systems. This is called the microcanonical ensemble.

We can apply the results of the previous section to a treatment of an isolated system. The grand canonical ensemble represents a collection of systems whose containing walls allow heat and molecules to pass freely from one system to another. From a physical point of view, the entire grand canonical ensemble is equivalent to one isolated system of volume $\mathscr{A}V$, containing $\mathscr{N}$ molecules and having energy $\mathscr{E}$. The partitions in Fig. 3–1 can be considered to be a conceptual division of one isolated system into $\mathscr{A}$ subsystems. The entropy of the entire ensemble S_e, considered as one isolated system, is $\mathscr{A}S$, where S is the entropy of each of the open, isothermal systems. This entropy is given by Eq. (3–12):

$$S = k(\beta\bar{E} + \gamma\bar{N} + \ln \Xi) \tag{3–19}$$

where we use the notation β and γ for convenience. We use Eqs. (3–7) and (3–9) for $\bar{E}$ and $\bar{N}$:

$$\begin{aligned} S &= k \ln \Xi + k\left(\sum_{N,j} \beta E_{Nj} \frac{e^{-\beta E_{Nj}} e^{-\gamma N}}{\Xi} + \sum_{N,j} \gamma \frac{N e^{-\beta E_{Nj}} e^{-\gamma N}}{\Xi}\right) \\ &= k \ln \Xi + k \sum_{N,j} (\beta E_{Nj} + \gamma N) \frac{e^{-\beta E_{Nj}} e^{-\gamma N}}{\Xi} \\ &= k \ln \Xi - k \sum_{N,j} (\ln a_{Nj}{}^* + \ln \Xi - \ln \mathscr{A}) \frac{a_{Nj}{}^*}{\mathscr{A}} \end{aligned} \tag{3–20}$$

where we have used Eq. (3–6) to write the last line. We can perform the summation over the second two terms in parentheses in Eq. (3–20):

$$S = k \ln \Xi - \frac{k}{\mathscr{A}} \sum_{N,j} a_{Nj}{}^* \ln a_{Nj}{}^* - k \ln \Xi + k \ln \mathscr{A}$$

or

$$\begin{aligned} S_e = \mathscr{A}S &= k\mathscr{A} \ln \mathscr{A} - k \sum_{N,j} a_{Nj}{}^* \ln a_{Nj}{}^* \\ &= k \ln W(\{a_{Nj}{}^*\}) \end{aligned} \tag{3–21}$$

We see that for an isolated system, the entropy is proportional to the logarithm of the number of states available to the system. In another notation, we can write

$$S = k \ln \Omega(N, V, E) \tag{3–22}$$

Equation (3–22) shows that the more states there are available to an isolated system, the higher is its entropy. This equation serves as the basis for qualitative statements concerning entropy and disorder, randomness, and so on. In practice, Eq. (3–22) is not used for the calculation of thermodynamic functions since N, V, and E are all mechanical variables.

The argument leading to Eq. (2–38) incidentally can be immediately applied to Eq. (3–22). For any spontaneous process in an isolated system,

$$\Delta S = k \ln \frac{\Omega_2}{\Omega_1} > 0$$

where 1 and 2 represent the initial and final states, respectively.

Equation (3–22) is due to Boltzmann and is possibly the best-known equation in statistical thermodynamics, mainly for historical reasons. Of course, Boltzmann (1844–1906) did not express his famous equation in terms of quantum states, but rather in a classical mechanical framework. We shall take up classical statistical mechanics in Chapter 7. Boltzmann, in fact, was a great contributor to both equilibrium and nonequilibrium statistical mechanics. He was one of the first to see clearly how probability ideas could be combined with mechanics. Equation (3–22) is carved on his tombstone in the Zentralfriedhoff in Vienna, although the equation is not often used today. However, his contribution to nonequilibrium statistical mechanics is such that to this day the so-called Boltzmann equation (Chapters 18 and 19) still is the fundamental equation describing the transport of dilute gases. It is interesting to note that Boltzmann, who contributed so much to understanding macroscopic phenomena in terms of molecular mechanics, lived at a time when the atomic theory was not so generally accepted as it is today, and his work was severely criticized by some of the leading physicists of the day. He committed suicide in 1906 (for reasons not entirely clear) and never lived to see the full acceptance of his work in statistical mechanics.

Although $\Omega(N, V, E)$ is not generally available, we have determined it for an ideal gas in Section 1–3 [*cf.* Eq. (1–37)]. If we calculate $k \ln \Omega$, neglecting terms of order less than $O(N^{-1})$, we get (see Problem 3–11):

$$S = Nk \ln\left[\left(\frac{2\pi mkT}{h^2}\right)^{3/2} \frac{Ve^{5/2}}{N}\right] \tag{3–23}$$

We shall see later that this equation gives excellent agreement with experiment, but now we simply show that if we use

$$dS = \frac{1}{T}dE + \frac{p}{T}dV - \frac{\mu}{T}dN \tag{3–24}$$

to get

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

and substitute Eq. (3–23) into this, we find

$$pV = NkT$$

which is the ideal gas equation of state. See Table 3–1 for a summary of the formulas related to the microcanonical ensemble.

It is possible to derive partition functions appropriate to other sets of independent variables, but the four that we have considered above are sufficient for most applications. We shall show that in the limit of large systems in equilibrium, one can choose an ensemble and its partition function on the basis of mathematical convenience rather than on the basis of which thermodynamic variables are used to describe the system. This result will come out of a study of fluctuations, which we turn to now.

3–3 FLUCTUATIONS

The methods that we have developed allow us to calculate ensemble averages of mechanical variables, which we then equate to thermodynamic functions. Equations such as Eq. (2–12) or (3–6) are the probability distributions over which these ensemble

Table 3–1. **A summary of formulas for several types of ensemble**

---

microcanonical ensemble, $\Omega(N, V, E)$

$$S = k \ln \Omega$$

$$dS = \frac{1}{T}dE + \frac{p}{T}dV - \frac{\mu}{T}dN$$

$$\frac{1}{kT} = \left(\frac{\partial \ln \Omega}{\partial E}\right)_{N,V} \tag{3–25}$$

$$\frac{p}{kT} = \left(\frac{\partial \ln \Omega}{\partial V}\right)_{N,E} \tag{3–26}$$

$$\frac{\mu}{kT} = -\left(\frac{\partial \ln \Omega}{\partial N}\right)_{V,E} \tag{3–27}$$

canonical ensemble, $Q(N, V, T)$

$$A = -kT \ln Q$$

$$dA = -S\,dT - p\,dV + \mu\,dN$$

$$S = k \ln Q + kT\left(\frac{\partial \ln Q}{\partial T}\right)_{N,V} \tag{3–28}$$

$$p = kT\left(\frac{\partial \ln Q}{\partial V}\right)_{N,T} \tag{3–29}$$

$$\mu = -kT\left(\frac{\partial \ln Q}{\partial N}\right)_{V,T} \tag{3–30}$$

$$E = kT^2\left(\frac{\partial \ln Q}{\partial T}\right)_{N,V} \tag{3–31}$$

grand canonical ensemble, $\Xi(V, T, \mu)$

$$pV = kT \ln \Xi$$

$$d(pV) = S\,dT + N\,d\mu + p\,dV$$

$$S = k \ln \Xi + kT\left(\frac{\partial \ln \Xi}{\partial T}\right)_{V,\mu} \tag{3–32}$$

$$N = kT\left(\frac{\partial \ln \Xi}{\partial \mu}\right)_{V,T} \tag{3–33}$$

$$p = kT\left(\frac{\partial \ln \Xi}{\partial V}\right)_{\mu,T} = kT\frac{\ln \Xi}{V} \tag{3–34}$$

isothermal-isobaric ensemble, $\Delta(N, T, p)$

$$G = -kT \ln \Delta$$

$$dG = -S\,dT + V\,dp + \mu\,dN$$

$$S = k \ln \Delta + kT\left(\frac{\partial \ln \Delta}{\partial T}\right)_{N,p} \tag{3–35}$$

$$V = -kT\left(\frac{\partial \ln \Delta}{\partial p}\right)_{N,T} \tag{3–36}$$

$$\mu = -kT\left(\frac{\partial \ln \Delta}{\partial N}\right)_{T,p} \tag{3–37}$$

---

averages are taken. In Section 1–5 we saw that the average is the first of a family of moments. Another important moment is the second central moment or the variance, $\overline{(x-\bar{x})^2}$, which is a measure of the spread of a probability distribution about the mean value. Furthermore, we saw toward the end of Section 1–5 that the most meaningful measure of the spread of a distribution is the square root of the variance, that is, the standard deviation, relative to the mean value. A standard deviation of 10^{10} may be large as an absolute number, but it is extremely small if the mean of the probability distribution is 10^{20}. In this section we shall calculate the variances of several mechanical variables and compare these to the mean values.

Any deviation of a mechanical variable from its mean value is called a fluctuation, and the investigation of the probability of such deviations is called fluctuation theory. Fluctuation theory is important in statistical mechanics for a number of reasons. The most obvious reason is to determine to what extent we expect to observe deviations from the mean values that we calculate. If the spread about these is large, then experimentally we would observe a range of values, whose mean or average is given by statistical thermodynamics. We shall see, however, that the probability of observing any value other than the mean value is extremely remote. As a corollary to this important result, we shall see that all of the ensembles that we have considered earlier are equivalent for all practical purposes. In addition, there are several statistical thermodynamical theories of solutions and light scattering based on fluctuation theory, and one formulation of the statistical mechanical theory of transport focuses on the rate of decay of spontaneous fluctuations.

Let us consider first fluctuations in a canonical ensemble. In a canonical ensemble, N, V, and T are held fixed, and we can investigate fluctuations in the energy, pressure, and related properties since these are the ones that vary from system to system. It is important to be aware of the properties that can vary and those properties that are fixed in each ensemble. We shall consider fluctuations in the energy. Thus we use Eq. (2–12) for the probability distribution of the energy and write for the variance

$$\begin{aligned}\sigma_E{}^2 &= \overline{(E-\bar{E})^2} = \overline{E^2} - \bar{E}^2 \\ &= \sum_j E_j{}^2 P_j - \bar{E}^2\end{aligned} \tag{3–38}$$

where

$$P_j = \frac{e^{-\beta E_j}}{Q(N, V, \beta)} \tag{3–39}$$

We can write Eq. (3–38) in a more convenient form by noting that

$$\begin{aligned}\sum_j E_j{}^2 P_j &= \frac{1}{Q}\sum_j E_j{}^2 e^{-\beta E_j} = -\frac{1}{Q}\frac{\partial}{\partial\beta}\sum_j E_j e^{-\beta E_j} \\ &= -\frac{1}{Q}\frac{\partial}{\partial\beta}(\bar{E}Q) = -\frac{\partial\bar{E}}{\partial\beta} - \bar{E}\frac{\partial \ln Q}{\partial\beta} \\ &= kT^2\frac{\partial\bar{E}}{\partial T} + \bar{E}^2\end{aligned} \tag{3–40}$$

Thus Eq. (3–38) becomes

$$\sigma_E{}^2 = kT^2\left(\frac{\partial\bar{E}}{\partial T}\right)_{N,V} \tag{3–41}$$

and if we associate $\bar{E}$ with the thermodynamic energy, we have

$$\sigma_E{}^2 = kT^2C_V \tag{3–42}$$

where C_V is the molar heat capacity.

To explore the *relative* magnitude of this spread, we look at

$$\frac{\sigma_E}{\bar{E}} = \frac{(kT^2C_V)^{1/2}}{\bar{E}} \tag{3–43}$$

To get an order-of-magnitude estimate of this ratio, we use the values of $\bar{E}$ and C_V for an ideal gas, namely, $O(NkT)$ and $O(Nk)$, respectively. If we use these values in Eq. (3–43), we find that $\sigma_E/\bar{E}$ is $O(N^{-1/2})$, showing that in a typical macroscopic system, the relative deviations from the mean are extremely small. The probability distribution of the energy may, therefore, be regarded as a Gaussian distribution which is practically a delta function.

We can derive a Gaussian distribution approximation to $P(E)$, the probability of observing a particular value of E in a canonical ensemble. According to Eq. (2–36), $P(E)$ is given by $C\Omega(E)e^{-E/kT}$, where C is a normalization factor which is independent of E. Since $\Omega(E)$ is an increasing function of E, and $e^{-E/kT}$ is a decreasing function of E, their product $P(E)$ peaks at some value of E, say E^*. But we have just seen above that the spread about the maximum value is extremely small, and so E^* and $\bar{E}$ are essentially the same point. The width of $P(E)$ is $O(N^{-1/2})$, and so E^* and $\bar{E}$ differ by $O(N^{-1/2})$.

Let us now expand $P(E)$ in a Taylor series about E^*, or $\bar{E}$. As in Section 1–5, it is more convenient to work with $\ln P(E)$. From the definition of $E^*(\approx \bar{E})$ as the value of E at the maximum in $P(E)$,

$$\left(\frac{\partial \ln P}{\partial E}\right)_{E=E^*=\bar{E}} = \left(\frac{\partial \ln \Omega}{\partial E}\right)_{E=E^*=\bar{E}} - \beta = 0 \tag{3–44}$$

Equation (3–44) determines $\bar{E}$ as a function of β. The second derivative of $\ln P(E)$ is

$$\left(\frac{\partial^2 \ln P}{\partial E^2}\right) = \left(\frac{\partial^2 \ln \Omega}{\partial E^2}\right)$$

which is to be evaluated at $E = E^* = \bar{E}$. Since

$$\begin{aligned}\left(\frac{\partial^2 \ln \Omega}{\partial E^2}\right)_{E=\bar{E}} &= \frac{\partial^2 \ln \Omega(\bar{E})}{\partial \bar{E}^2} = \frac{\partial}{\partial \bar{E}}\left(\frac{\partial \ln \Omega(\bar{E})}{\partial \bar{E}}\right) \\ &= \frac{\partial}{\partial \bar{E}}\left(\frac{\partial \ln \Omega}{\partial E}\right)_{E=\bar{E}} = \frac{\partial \beta}{\partial \bar{E}}\end{aligned}$$

where the last term follows from Eq. (3–44), we have

$$\left(\frac{\partial^2 \ln P}{\partial E^2}\right)_{E=E^*=\bar{E}} = \frac{\partial \beta}{\partial \bar{E}} = -\frac{1}{kT^2}\frac{\partial T}{\partial \bar{E}} = -\frac{1}{kT^2C_V} \tag{3–45}$$

The Taylor expansion of $\ln P(E)$ through quadratic terms is

$$\ln P(E) = \ln P(\bar{E}) - \frac{(E-\bar{E})^2}{2kT^2C_V} + \cdots \tag{3–46}$$

or

$$P(E) = P(\bar{E})\exp\left\{-\frac{(E-\bar{E})^2}{2kT^2C_V}\right\} \tag{3–47}$$

Problem 3–16 involves showing that terms beyond the quadratic terms can be ignored in Eq. (3–46).

If we compare Eq. (3–47) to the standard form of a Gaussian distribution [Eq. (1–72)], we see that $\sigma_E^2 = kT^2C_V$ (in agreement with Eq. (3–42)] and that the normalization constant $P(\bar{E})$ is $(2\pi\sigma_E^2)^{-1/2}$. Equation (3–47) can be used to calculate the probability of observing a value of E that differs from $\bar{E}$. For example, the probability of observing an energy that differs by 0.1 percent from the average energy of 1 mole of an ideal gas $O(e^{-10^6})$, an extremely small number. (See Problem 3–12.)

Incidentally, the derivation of Eq. (3–47) is a case where we must be careful not to confuse the variable E in $P(E)$ with E^*, $\bar{E}$, or the thermodynamic quantity E, which unfortunately is also called "E." This is especially true of Eq. (3–44), where E is a variable, and $\bar{E}$ is that particular value of the variable for which the quantity $\partial \ln \Omega/\partial E$, a function of E, is equal to the preassigned value of β (N, V, T are given in a canonical ensemble).

We could also calculate the fluctuations in the pressure in a canonical ensemble, but this is left to Problem 3–18. Instead, we consider the fluctuations in a grand canonical ensemble. In a grand canonical ensemble, V, T, and μ are held fixed, while the energy and number of particles in each system are allowed to vary. We can calculate the fluctuation in the number of particles in the same manner as we treated the fluctuation in energy in a canonical ensemble. If σ_N^2 is the variance in the number of particles, then

$$\sigma_N^2 = \overline{N^2} - \bar{N}^2 = \sum_{N,j} N^2 P_{Nj} - \bar{N}^2 \tag{3–48}$$

where

$$P_{Nj} = \frac{e^{-\beta E_{Nj}} e^{-\gamma N}}{\Xi(V, \beta, \gamma)}$$

We treat $\overline{N^2}$ in analogy to Eq. (3–40):

$$\begin{aligned}\sum_{N,j} N^2 P_{Nj} &= \frac{1}{\Xi} \sum_{N,j} N^2 e^{-\beta E_{Nj}} e^{-\gamma N} = -\frac{1}{\Xi}\frac{\partial}{\partial \gamma} \sum_{N,j} N e^{-\beta E_{Nj}} e^{-\gamma N} \\ &= -\frac{1}{\Xi}\frac{\partial}{\partial \gamma}(\bar{N}\Xi) = -\frac{\partial \bar{N}}{\partial \gamma} - \bar{N}\frac{\partial \ln \Xi}{\partial \gamma} \\ &= kT\left(\frac{\partial \bar{N}}{\partial \mu}\right)_{V,T} + \bar{N}^2\end{aligned} \tag{3–49}$$

Thus Eq. (3–48) becomes

$$\sigma_N^2 = kT\left(\frac{\partial \bar{N}}{\partial \mu}\right) \tag{3–50}$$

The right-hand side of this equation can be written in a more familiar form by thermodynamic manipulations. Problem 3–26 proves that

$$\left(\frac{\partial \mu}{\partial \bar{N}}\right)_{V,T} = -\frac{V^2}{\bar{N}^2}\left(\frac{\partial p}{\partial V}\right)_{N,T}$$

and so

$$\sigma_N^2 = \frac{\bar{N}^2 kT\kappa}{V} \tag{3–51}$$

where κ is the isothermal compressibility

$$\kappa \equiv -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_{N,T} \tag{3–52}$$

The value of σ_N relative to $\overline{N}$ is

$$\frac{\sigma_N}{\overline{N}} = \left(\frac{kT\kappa}{V}\right)^{1/2} \tag{3–53}$$

To get an order-of-magnitude estimate of this ratio, we use the fact that $\kappa = 1/p$ for an ideal gas to get $\sigma_N/\overline{N} = N^{-1/2}$. Again we find that relative deviations from the mean are very small. The result, $0(N^{-1/2})$, is typical of fluctuations in statistical thermodynamics.

Since V is fixed in the grand canonical ensemble, the fluctuation in the number of particles is proportional to the fluctuation in the density ρ, and so

$$\frac{\sigma_\rho}{\overline{\rho}} = \frac{\sigma_N}{\overline{N}} = \left(\frac{kT\kappa}{V}\right)^{1/2} \tag{3–54}$$

There is a condition under which the fluctuations in density are not negligible. At the critical point of a substance, $(\partial p/\partial V)_{N,T}$ is zero, and hence its isothermal compressibility is infinite. Thus there are large fluctuations in the density from point to point in a fluid at its critical point. This is observed macroscopically by the phenomenon of critical opalescence, in which a pure substance becomes turbid at its critical point.

We can also derive a Gaussian approximation to $P(N)$. Let $N^*(=\overline{N})$ be the value of N at the peak in $P(N)$. We have

$$P(N) = CQ(N, V, T)e^{\beta\mu N}$$

where C is a normalization constant. Then

$$\left(\frac{\partial \ln P}{\partial N}\right)_{N=N^*} = \left(\frac{\partial \ln Q}{\partial N}\right)_{N=\overline{N}} + \beta\mu = 0$$

This equation determines $\overline{N}$ as a function of $\beta\mu$. Also,

$$\left(\frac{\partial^2 \ln P}{\partial N^2}\right)_{N=\overline{N}} = \left(\frac{\partial^2 \ln Q}{\partial N^2}\right)_{N=\overline{N}} = \frac{\partial}{\partial \overline{N}}\frac{\partial \ln Q(\overline{N}, V, T)}{\partial \overline{N}} = -\frac{\partial \beta\mu}{\partial \overline{N}}$$

$$= -\frac{1}{kT(\partial \overline{N}/\partial \mu)_{V,T}}$$

Thus we find

$$P(N) = P(\overline{N})\exp\left[\frac{-(N-\overline{N})^2}{2kT(\partial \overline{N}/\partial \mu)_{V,T}}\right] \tag{3–55}$$

which gives the same expression for $\sigma_N{}^2$ as Eq. (3–50). Problems 3–19 through 3–20 involve the determination of fluctuations in the isothermal-isobaric ensemble.

An interesting application of the above fluctuation formulas is to the scattering of light by the atmosphere. It can be shown that if light of intensity I_0 is incident on a region of volume V with a dielectric constant ε, which differs from the average value of ε for the medium $\bar{\varepsilon}$, the intensity of light scattered at an angle θ at a distance R is

$$\frac{I(\theta)}{I_0} = \frac{\pi^2 V^2 \sigma_\varepsilon{}^2}{2\lambda^4}\frac{(1+\cos^2\theta)}{R^2} \tag{3–56}$$

where $\sigma_\varepsilon{}^2$ is the variance of ε, and λ is the wavelength of the incident light in vacuum.

This is called Rayleigh scattering. The dielectric constant ε is related to the density by the so-called Clausius-Mossotti equation

$$\frac{\varepsilon - 1}{\varepsilon + 2} = A\rho \tag{3–57}$$

which is derived and discussed in most physical chemistry texts. The quantity A is a constant, and ρ is the density. We can see from this equation that fluctuations in ρ lead to fluctuations in ε, and hence to Rayleigh scattering by Eq. (3–56). If we calculate $\sigma_\varepsilon{}^2$ in terms of $\sigma_\rho{}^2$ from Eq. (3–57), and use Eq. (3–54) for $\sigma_\rho{}^2$, we find (see Problem 3–21)

$$\frac{I(\theta)}{I_0} = \frac{\pi^2 kT}{18\lambda^4}\kappa(\varepsilon - 1)^2(\varepsilon + 2)^2 V \frac{(1 + \cos^2\theta)}{R^2}$$

where κ is the isothermal compressibility. By integrating this over the surface of a sphere of radius R, we obtain finally

$$\frac{I_{\text{scattered}}}{I_0} = \int \frac{I(\theta)}{I_0} R^2 \sin\theta \, d\theta \, d\phi$$

$$= \frac{8\pi^3}{27\lambda^4} kT\kappa(\varepsilon - 1)^2(\varepsilon + 2)^2 V$$

This equation shows that the blue color of the sky is due to fluctuations in the density of the atmosphere. The λ^4 in the denominator gives rise to a strong dependence on wavelength, so that the short wavelengths (blue) of the sun's light are scattered more than the red, and hence the sky appears blue. Similarly, red sunsets and sunrises are due to the fact that the long wavelengths (red) are not scattered as much as the blue.

There is one result of fluctuation theory which will be very useful to us. We have stated above that the various ensembles and their partition functions are essentially equivalent to each other, and that one can choose to work with a partition function on the basis of mathematical convenience. We now show why this is so.

Consider the canonical partition function:

$$Q(N, V, T) = \sum_E \Omega(N, V, E)e^{-E/kT} \tag{3–58}$$

We have seen in Eq. (3–47) that $P(E) = C\Omega(E)\exp(-E/kT)$ is an extremely narrow Gaussian function of E. In the limit of large N (and it is only in the limit of large N that classical thermodynamics is valid), only one value of E is important, namely, $E = E^* = \bar{E}$. Thus in the summation in Eq. (3–58), only the term with $E = \bar{E}$ contributes, and Eq. (3–58) becomes

$$Q(N, V, T) = \Omega(N, V, \bar{E})e^{-\bar{E}/kT} \tag{3–59}$$

Although the systems of a canonical ensemble can, in principle, assume any value of E (as long as it is an eigenvalue of the N-particle Schrödinger equation), it happens that the energy of the entire ensemble is distributed uniformly throughout the ensemble, and each system is almost certain to be found with the average energy $\bar{E}$. A canonical ensemble degenerates, in a sense, to a microcanonical ensemble.

If we take the logarithm of Eq. (3–59) and use Eq. (2–35), we find that

$$A = \bar{E} - kT \ln \Omega(N, V, \bar{E})$$

or that

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

This is an alternative derivation of the fundamental relation between the entropy and the number of states accessible to the system.

The general results we have obtained here are also obtained for other ensembles. For example, although the systems of a grand canonical ensemble can assume any value of N and E, in practice it turns out that the total energy and the total number of molecules of the entire ensemble are distributed uniformly throughout the ensemble, and each system has the average energy and contains the average number of molecules. This, of course, is exactly what one expects intuitively, as long as the systems are of macroscopic size and the density is not *extremely* low.

These results can be used to write down, by inspection, the characteristic thermodynamic function of any partition function. Equations (3–58) and (3–59) are a good example. Suppose we did not know that $A = -kT \ln Q$. We do know that $S = k \ln \Omega$, however, and so if we take the logarithm of Eq. (3–59), we get that $\ln Q = S/k - \beta E$, which shows that $A = -kT \ln Q$. Since partition functions, in general, are a sum of $\Omega(N, V, E)$ multiplied by exponential factors, this method can always be used to determine the thermodynamic characteristic function. (See Problem 3–15.)

ADDITIONAL READING

General

Hill, T. L. 1960. *Statistical thermodynamics*. Reading, Mass.: Addison-Wesley. Sections 1–5 through 1–7 and Chapter 2.

———. 1956. *Statistical mechanics*. New York: McGraw-Hill. Sections 14, 15, and Chapter 4.

Kestin, J., and Dorfman, J. R. 1971. *A course in statistical thermodynamics*. New York: Academic. Section 5–18.

Knuth, E. 1966. *Statistical thermodynamics*. New York: McGraw-Hill. Chapters 3, 4, and 5.

Kubo, R. 1965. *Statistical mechanics*. Amsterdam: North-Holland Publishing Co. Sections 1–12 to 1–14.

Reif, F. 1965. *Statistical and thermal physics*. New York: McGraw-Hill. Chapter 3.

Rushbrooke, G. S. 1949. *Statistical mechanics*. London: Oxford University Press. Chapters 15 and 17.

Fluctuations

Andrews, F. C. 1963. *Equilibrium statistical mechanics*. New York: Wiley. Chapter 33.

Davidson, N. 1962. *Statistical mechanics*. New York: McGraw-Hill. Chapter 14.

Kestin, J., and Dorfman, J. R. 1971. *A course in statistical thermodynamics*. New York: Academic. Chapter 14.

Landau, L. D., and Lifshitz, E. M. 1958. *Statistical physics*. London: Pergamon Press. Chapter 12.

Münster, A. 1969. *Statistical thermodynamics*, Vol. I. Berlin: Springer-Verlag. Chapter 3.

Schrödinger, E. 1952. *Statistical thermodynamics*. Cambridge: Cambridge University Press. Chapter 5.

Tolman, R. C. 1938. *Statistical mechanics*. London: Oxford University Press. Section 141.

PROBLEMS

3–1. Derive Eq. (3–5).

3–2. Using a grand canonical formalism, show that any two systems at the same temperature have the same value of β.

3–3. For a grand canonical ensemble show that

$$\left(\frac{\partial \bar{E}}{\partial V}\right)_{\gamma, \beta} + \beta\left(\frac{\partial \bar{p}}{\partial \beta}\right)_{\gamma, V} = -\bar{p}$$

Compare this to the thermodynamic equation (see Problem 1–31)

$$\left(\frac{\partial E}{\partial V}\right)_{\mu/T, 1/T} + \frac{1}{T}\left(\frac{\partial p}{\partial(1/T)}\right)_{\mu/T, V} = -p$$

to suggest that $\beta = \text{const}/T$ for a grand canonical ensemble.

3–4. State and use Euler's theorem to show

$$p = kT\left(\frac{\partial \ln \Xi}{\partial V}\right)_{\mu, T} = kT\frac{\ln \Xi}{V}$$

3–5. Show that the entropy given by Eq. (3–12) goes to zero as T goes to zero.

3–6. Derive the principal thermodynamic connection formulas of the grand canonical ensemble starting from

$$pV = kT \ln \Xi$$

and

$$d(pV) - S\,dT + N\,d\mu + p\,dV$$

3–7. Show that for a two-component system

$$\Xi(\mu_1, \mu_2, T, V) = \sum_{N_1}\sum_{N_2} Q(N_1, N_2, V, T)\lambda_1^{N_1}\lambda_2^{N_2}$$

where $\lambda_i = e^{\mu_i/kT}$ $(i = 1, 2)$. From this derive the corresponding thermodynamic connection formulas.

3–8. In the next chapter we shall see that the grand partition function of an ideal monatomic gas is

$$\Xi = e^{q\lambda}$$

where $q = (2\pi mkT/h^2)^{3/2}V$. Derive the thermodynamic properties of an ideal monatomic gas from Ξ.

3–9. Show that the partition function appropriate to an isothermal-isobaric ensemble is

$$\Delta(N, p, T) = \sum_{E}\sum_{V} \Omega(N, V, E)e^{-E/kT}\,e^{-pV/kT}$$

Derive the principal thermodynamic connection formulas for this ensemble.

3–10. In Problem 5–17 we shall show that the isothermal-isobaric partition function of an ideal monatomic gas is

$$\Delta = \left[\frac{(2\pi m)^{3/2}(kT)^{5/2}}{ph^3}\right]^N$$

Derive the thermodynamic properties of an ideal monatomic gas from Δ.

3–11. Derive Eq. (3–23) starting from Eq. (1–37).

3–12. Calculate the probability of observing an energy that differs by 10^{-4} percent from the average energy of 1 mole of an ideal gas.

3–13. Show that for macroscopic ideal systems, ones obtains the same result for the entropy whether one uses $S = k \ln \Phi(E)$, where Φ is the number of quantum states with energy $\leq E$ [Eq. (1–36)], or $S = k \ln \Omega(E, \Delta E)$, where $\Omega(E, \Delta E)$ is the number of quantum states within energy ΔE about E [Eq. (1–37)] as long as $\Delta E/E$ is small, but not zero. Show that S is insensitive to ΔE over a wide range of ΔE. The next problem discusses this remarkable result more generally.

3–14. Let $\Omega(E)\,dE$ be the number of quantum states between E and $E + dE$. In Chapter 1 we showed that $\Omega(E)$ is a monotonically increasing function of E (at least for an ideal gas). We can write two obvious inequalities for $\Omega(E)$:

$$\Phi(E) \equiv \int_0^E \Omega(E')\,dE' \geq \Omega(E)\,\Delta E$$

$$E\Omega(E) \geq \int_0^E \Omega(E')\,dE' = \Phi(E)$$

where ΔE is a small region surrounding E. By multiplying the second inequality by $\Delta E/E$, we get

$$\Omega(E)\,\Delta E \geq \frac{\Delta E}{E}\int_0^E \Omega(E')\,dE'$$

Combining this inequality with the first one above gives

$$\frac{\Delta E}{E}\int_0^E \Omega(E')\,dE' \leq \Omega(E)\,\Delta E \leq \int_0^E \Omega(E')\,dE'$$

Taking logarithms gives

$$\ln \Phi(E) - \ln\left(\frac{E}{\Delta E}\right) \leq \ln\,[\Omega(E)\,\Delta E] \leq \ln \Phi(E)$$

Now unless ΔE is extremely small, $\ln\,(E/\Delta E)$ is completely negligible compared to $\ln \Phi(E)$, since the total number of states with energies equal to or less than E is at least $0(e^N)$. Show that even if the energy could be measured to a millionth of a percent, $\ln\,(E/\Delta E) \approx 18$, which is completely negligible compared to N.

3–15. Fluctuation theory provides a simple method to determine the characteristic function associated with a particular partition function. Consider the canonical partition function

$$Q(N, V, T) = \sum_E \Omega(N, V, E)\, e^{-E/kT}$$

According to the theory of fluctuations, there is effectively only one term in this summation, and so we write

$$Q(N, V, T) = \Omega(N, V, \bar{E})\, e^{-E/kT}$$

Remembering that $S = k \ln \Omega$, we have, upon taking logarithms, that

$$\ln Q = \frac{S}{k} - \frac{E}{kT}$$

or that

$$\ln Q = \frac{-A}{kT}$$

Proceeding in a like manner, determine the characteristic thermodynamic function of the following partition functions:

$$\Xi(V, T, \mu) = \sum_N Q(N, V, T)\, e^{\beta\mu N}$$

$$\Delta(p, T, N) = \sum_V Q(N, V, T)\, e^{-\beta p V}$$

$$\phi(V, E, \beta\mu) = \sum_N \Omega(N, V, E)\, e^{\beta\mu N}$$

$$\Psi(V, T, \mu_1, N_2) = \sum_{N_1} Q(N_1, N_2, T, V)\, e^{\beta\mu N_1}$$

$$W(p, \gamma, T, N) = \sum_V \sum_{\mathscr{A}} Q(N, V, \mathscr{A}, T)\, e^{-\beta p V}\, e^{\beta\gamma\mathscr{A}}$$

where $\mathscr{A}$ is surface area, and γ is the surface tension.

3–16. When we derived the Gaussian expression for $P(E)$ in a canonical ensemble, we expanded $\ln P(E)$ in a Taylor expansion about $E = E^* \approx \bar{E}$, dropping terms after the quadratic term. Show that these terms are negligible.

3–17. Show that

$$\overline{(E-\bar{E})^3} = k^2\left\{T^4\left(\frac{\partial C_V}{\partial T}\right) + 2T^3C_V\right\}$$

and that

$$\frac{\overline{(E-\bar{E})^3}}{\bar{E}^3} = 0(N^{-2})$$

for a canonical ensemble.

3–18. Derive an expression for the fluctuation in the pressure in a canonical ensemble.

3–19. Show that for an isothermal-isobaric ensemble

$$P(V) - P(V^*)\exp\left\{\frac{(V-\bar{V})^2}{2kT\left(\dfrac{\partial V}{\partial p}\right)_{N,T}}\right\}$$

3–20. Derive an equation for the fluctuation in the volume in an isothermal-isobaric ensemble. In other words, derive an equation for $\overline{V^2} - \bar{V}^2$. Express your answer in terms of the isothermal compressibility, defined by

$$\kappa = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_{N,T}$$

Show that $\sigma_v/\bar{V}$ is of the order of $N^{-1/2}$.

3–21. By calculating σ_ε^2 in terms of σ_ρ^2 from Eq. (3–57) and using Eq. (3–54) for σ_ρ^2, show that

$$\frac{I(\theta)}{I_0} = \frac{\pi^2 kT}{18\lambda^4}\kappa(\varepsilon-1)^2(\varepsilon+2)^2 V\frac{(1+\cos^2\theta)}{R^2}$$

3–22. Show that the fluctuation in energy in a grand canonical ensemble is

$$\sigma_E^2 = (kT^2C_V) + \left(\frac{\partial\bar{E}}{\partial\bar{N}}\right)_{T,V}\sigma_N^2$$

3–23. Show that in a two-component open, isothermal ensemble that

$$\begin{aligned}\overline{N_1N_2} - \bar{N}_1\bar{N}_2 &= kT\left(\frac{\partial\bar{N}_1}{\partial\mu_2}\right)_{V,T,\mu_1}\\ &= kT\left(\frac{\partial\bar{N}_2}{\partial\mu_1}\right)_{V,T,\mu_2}\end{aligned}$$

3–24. Show that

$$\overline{H^2} - \bar{H}^2 = kT^2C_p$$

in an N, p, T ensemble.

3–25. Use the formulas in Table 3–1 to derive expressions for any other thermodynamic functions for each of the four ensembles listed there.

3–26. Show that

$$\left(\frac{\partial\mu}{\partial N}\right)_{V,T} = -\frac{V^2}{N^2}\left(\frac{\partial p}{\partial V}\right)_{N,T}$$

CHAPTER 4

BOLTZMANN STATISTICS, FERMI-DIRAC STATISTICS, AND BOSE-EINSTEIN STATISTICS

The results that we have derived up to now are valid for macroscopic systems. In order to apply these equations, it is necessary to have the set of eigenvalues $\{E_j(N, V)\}$ of the N-body Schrödinger equation. In general, this is an impossible task. There are many important systems, however, in which the N-body Hamiltonian operator can be written as a sum of independent individual Hamiltonians. In such cases the total energy of the system can be written as a sum of individual energies. This leads to a great simplification of the partition function, and allows us to apply the results with relative ease.

We shall see that the final equations depend upon whether the individual particles of the system are fermions (that is, the N-body wave function is antisymmetric under the interchange of identical particles) or bosons (the N-body wave function is symmetric under the interchange of identical particles). These two types of particles obey different laws, called Fermi-Dirac or Bose-Einstein statistics. We shall show that under normal conditions (for example, sufficiently high temperatures), both of these distribution laws can be approximately reduced to an even simpler one, called Boltzmann statistics. The Boltzmann distribution law can also be derived from $Q(N, V, T)$ at high temperature without first deriving the Fermi-Dirac and Bose-Einstein distribution laws, and this is done in Section 4–1. We shall discuss in this section just what is meant by "normal" conditions or "sufficiently high" temperatures. Then in Section 4–2 we derive the two fundamental distribution laws, Fermi-Dirac and Bose-Einstein statistics, and show how both of them reduce to Boltzmann statistics in the appropriate limit.

4–1 THE SPECIAL CASE OF BOLTZMANN STATISTICS

In Section 1–3 it was shown that if the Hamiltonian of a many-body system can be written as a sum of one-body Hamiltonians, the energy of the system is the sum of individual energies, and the wave function is a product of the single-particle wave functions. In addition, the wave functions of a system of identical particles must satisfy certain symmetry requirements with respect to the interchange of the particles. All known particles fall into two classes: those whose wave function must be symmetric under the operation of the interchange of two identical particles, and those whose wave function must be antisymmetric under such an exchange. Particles belonging to the first class are called bosons, and the others are called fermions. There is no restriction of the distribution of bosons over their available energy states, but fermions have the very severe restriction that no two identical fermions can occupy the same single-particle energy state. This restriction follows immediately from the requirement that the wave function be antisymmetric (see Problem 1–26). These considerations become important in enumerating the many-body energy states available to the system.

There are many problems in which the Hamiltonian can be written as a sum of simpler Hamiltonians. The most obvious example perhaps is the case of a dilute gas, where the molecules are on the average far apart, and hence their intermolecular interactions can be neglected. Another example, which may be familiar from physical chemistry, is the decomposition of the Hamiltonian of a polyatomic molecule into its various degrees of freedom:

$$\mathscr{H} \approx \mathscr{H}_{\text{translational}} + \mathscr{H}_{\text{rotational}} + \mathscr{H}_{\text{vibrational}} + \mathscr{H}_{\text{electronic}} \tag{4–1}$$

Equation (4–1) is a good first approximation and can be systematically corrected by the introduction of small interaction terms.

There are many other problems in physics in which the Hamiltonian, by a proper and clever selection of variables, can be written as a sum of individual terms. Although these individual terms need not be Hamiltonians for actual individual molecules, they are nevertheless used to define the so-called quasi-particles, which mathematically behave like independent real particles. Some of the names of quasi-particles that are found in the literature are photons, phonons, plasmons, magnons, rotons, and other " ons." In spite of the apparent limitation of this requirement on the Hamiltonian, we can see that it is very useful and can be used to study solids (Chapter 11) and liquids (see Chapter 12 of *Statistical Thermodynamics*), systems in which the decomposition of a many-body Hamiltonian into a sum of independent terms would hardly appear to be justified. First let us consider the canonical partition function for a system of distinguishable particles, in which the Hamiltonian can be written as a sum of individual terms. Denote the individual energy states by $\{\varepsilon_j{}^a\}$, where the superscript denotes the particle (they are distinguishable), and the subscript denotes the state. In this case the canonical partition function becomes

$$\begin{aligned} Q(N, V, T) &= \sum_j e^{-E_j/kT} = \sum_{,j,k,\ldots} e^{-(\varepsilon_i{}^a + \varepsilon_j{}^b + \varepsilon_k{}^c + \cdots)/kT} \\ &= \sum_i e^{-\varepsilon_i{}^a/kT} \sum_j e^{-\varepsilon_j{}^b/kT} \sum_k e^{-\varepsilon_k{}^c/kT} \\ &= q_a q_b q_c \end{aligned} \tag{4–2}$$

where

$$q(V, T) = \sum_i e^{-\varepsilon_i/kT} \tag{4–3}$$

Equation (4–2) is a very important result. It shows that if we can write the N-particle Hamiltonian as a sum of independent terms, and if the particles are *distinguishable*, then the calculation of $Q(N, V, T)$ reduces to a calculation of $q(V, T)$. Since $q(V, T)$

requires a knowledge only of the energy values of an individual particle or quasiparticle, its evaluation is quite feasible. In most cases $\{\varepsilon_i\}$ is a set of molecular energy states; thus $q(V, T)$ is called a molecular partition function.

If the energy states of all the particles are the same, then Eq. (4–2) becomes

$$Q(N, V, T) = [q(V, T)]^N \qquad \text{(distinguishable particles)} \tag{4–4}$$

Equation (4–4) shows that the original N-body problem (the evaluation of $Q(N, V, T)$) can be reduced to a one-body problem (the evaluation of $q(V, T)$) if the particles are independent and distinguishable. Although particles are certainly not distinguishable in general, there are many important cases where they can be treated as such. An excellent example of this is a perfect crystal. In a perfect crystal each atom is confined to one and only one lattice point, which we could, in principle, identify by a set of three numbers. Since each particle, then, is confined to a lattice point and the lattice points are distinguishable, the particles themselves are distinguishable. Furthermore, we shall see in Chapter 11 that although there are strong intermolecular interactions in crystals, we can treat the vibration of each particle about its lattice point as independent to a first approximation.

Another useful application of the separation indicated in Eq. (4–2) is to the molecular partition function itself. Equation (4–1) shows that the molecular Hamiltonian can be approximated by a sum of Hamiltonians for the various degrees of freedom of the molecule. Consequently we get the useful result that

$$q_{\text{molecule}} = q_{\text{translational}} q_{\text{rotational}} q_{\text{vibrational}} q_{\text{electronic}} \cdots \tag{4–5}$$

where, for example,

$$q_{\text{translational}} = \sum_i e^{-\varepsilon_i^{\text{trans}}/kT} \tag{4–6}$$

Thus not only can we reduce an N-body problem to a one-body problem, but it is possible to reduce it further into the individual degrees of freedom of the single particles.

Equation (4–4) is an attractive result, but atoms and molecules are, in general, not distinguishable; thus the utility of Eq. (4–4) is severely limited. The situation becomes more complicated when the inherent indistinguishability of atoms and molecules is considered. In this case, the N-body energy is

$$E_{ijkl\ldots} = \varepsilon_i + \varepsilon_j + \varepsilon_k + \varepsilon_l + \cdots \tag{4–7}$$

and the partition function is

$$Q(N, V, T) = \sum_{i,j,k,l\ldots} e^{-(\varepsilon_i+\varepsilon_j+\varepsilon_k+\varepsilon_l+\cdots)/kT} \tag{4–8}$$

Because the molecules are indistinguishable, one cannot sum over $i, j, k, l, \ldots$ separately as we did to get Eq. (4–2).

Consider, for example, the case of fermions. The antisymmetry of the wave function requires that no two identical fermions can occupy the same single-particle energy state. Thus in Eq. (4–8), terms in which two or more indices are the same cannot be included in the summation. The indices i, j, k, l, and so on, are not independent of one another, and a direct evaluation of $Q(N, V, T)$ for fermions by means of Eq. (4–8) is very difficult.

Bosons do not have the restriction that no two can occupy the same molecular state, but the summation in Eq. (4–8) is still complicated. Consider a term in Eq. (4–8) in which all of the indices are the same except one, that is, a term of the form $\varepsilon_i + \varepsilon_j + \varepsilon_j + \varepsilon_j + \cdots$ with $i \neq j$. Because the particles are indistinguishable, the position of ε_i is unimportant, and so this state is identical with $\varepsilon_j + \varepsilon_i + \varepsilon_j + \varepsilon_j + \cdots$ or $\varepsilon_j + \varepsilon_j + \varepsilon_i + \varepsilon_j + \varepsilon_j + \cdots$, and so on. Such a state should be included only once in Eq. (4–8), but an unrestricted summation over the indices in Eq. (4–8) would produce N terms of this type. Consider the other extreme in which all of the particles are in different molecular states, that is, the state with energy $\varepsilon_i + \varepsilon_j + \varepsilon_k + \cdots$ with $i \neq j \neq k \neq$ Because the particles are indistinguishable, the $N!$ states obtained by permuting the N different subscripts are identical and should occur only once in Eq. (4–8). Such terms will, of course, appear $N!$ times in an unrestricted summation. Consequently, a direct evaluation of Q for bosons by means of Eq. (4–8) also is difficult.

The terms that introduce complications are those in which two or more indices are the same. If it were not for this kind of term, one could carry out the summation in Eq. (4–8) in an unrestricted manner, and then correct the sum by dividing by $N!$ It turns out that this procedure yields an excellent approximation in many (most) cases for the following reason.

We showed in Section 1–3 that for a particle in a box, the number of molecular quantum states with energy $\leq \varepsilon$ is

$$\Phi(\varepsilon) = \frac{\pi}{6}\left(\frac{8ma^2\varepsilon}{h^2}\right)^{3/2} \tag{4–9}$$

For $m = 10^{-22}$ g, $a = 10$ cm, and $T = 300°\text{K}$, $\Phi(\varepsilon) = O(10^{30})$. Although this calculation is done for one particle in a cube (i.e., one molecule of an ideal gas), the order of magnitude of the result is general. Thus we see that the number of molecular quantum states available to a molecule at room temperature, say, is much greater than the number of molecules in the system for all but the most extreme densities. Since each particle has many individual states to choose from, it will be a rare event for two particles to be in the same molecular state. Therefore the vast majority of terms in Eq. (4–8) will have all different indices. This allows us to sum over all the indices unrestrictedly and divide by $N!$ to get

$$Q(N, V, T) = \frac{q^N}{N!} \qquad \text{(indistinguishable particles)} \tag{4–10}$$

with

$$q(V, T) = \sum_j e^{-\varepsilon_j/kT}$$

for a system of identical, indistinguishable particles satisfying the condition that the number of available molecular states is much greater than the number of particles.

Equation (4–10) is an extremely important result, since it reduces a many-body problem to a one-body problem. No longer is there a condition of distinguishability; the indistinguishability of the particles has been included by dividing by $N!$, a valid procedure for most systems under most conditions. We can investigate this condition in more detail using Eq. (4–9) for an ideal gas. Mathematically, we require that

$$\Phi(\varepsilon) \gg N$$

Using Eq. (4–9), we have the condition

$$\frac{\pi}{6}\left(\frac{12mkT}{h^2}\right)^{3/2} \gg \frac{N}{V} \tag{4–11}$$

where we have set $\varepsilon = 3kT/2$. Clearly this condition is favored by large mass, high temperature, and low density. Numerically it turns out that (4–11) is satisfied for all but the very lightest molecules at very low temperatures. Table 4–1 examines this condition for a number of systems. We see that the use of Eq. (4–10) is justified in most cases. We have examined (4–11) for only monatomic systems, but the results are valid for polyatomic molecules as well, since the translational energy states account for almost all of the energy states available to any molecule.

When Eq. (4–10) is valid, that is, when the number of available molecular states is much greater than the number of particles in the system, we say that the particles obey *Boltzmann statistics*. Boltzmann statistics is an approximation that becomes increasingly better at higher temperatures. We shall show in Chapter 7 that at high enough temperatures, one can describe the energy of a system by classical mechanics. Since the limiting case of Boltzmann statistics and the use of classical mechanics both require a high-temperature limit, Boltzmann statistics is also called the classical limit.

Let us examine Eq. (4–10). The total energy of the N-body system is

$$E = N\bar{\varepsilon} = kT^2\left(\frac{\partial \ln Q}{\partial T}\right)_{N,V} = N\sum_j \varepsilon_j \frac{e^{-\varepsilon_j/kT}}{q} \tag{4–12}$$

The first equality is valid, because the molecules are assumed to be independent, and hence their energies are additive. We see from Eq. (4–12) that the average energy of a particle is

$$\bar{\varepsilon} = \sum_j \varepsilon_j \frac{e^{-\varepsilon_j/kT}}{q} \tag{4–13}$$

We can conclude from this equation that the probability that a molecule is in the jth energy state

$$\pi_j = \frac{e^{-\varepsilon_j kT}}{\sum_j e^{-\varepsilon_j kT}} = \frac{e^{-\varepsilon_j kT}}{q} \tag{4–14}$$

It is interesting to note that the fluctuations in ε are of the same order as ε itself (see Problems 4–18 and 4–19), that is, the probability distribution for single molecules is not sharp. A sharp probability distribution is a many-body effect.

Table 4–1. **The quantity $(6N/\pi V)(h^2/12mkT)^{3/2}$ for a number of simple systems***

	T(°K)	$\dfrac{6N}{\pi V}\left(\dfrac{h^2}{12mkT}\right)^{3/2}$
liquid helium	4	1.6
gaseous helium	4	0.11
gaseous helium	20	2.0×10^{-3}
gaseous helium	100	3.5×10^{-5}
liquid neon	27	1.1×10^{-2}
gaseous neon	27	8.2×10^{-5}
gaseous neon	100	3.1×10^{-6}
liquid argon	86	5.1×10^{-4}
gaseous argon	86	1.6×10^{-6}
liquid krypton	127	5.4×10^{-5}
gaseous krypton	127	2.0×10^{-7}
electrons in metals (sodium)	300	1465

* This quantity must be much less than unity for Eq. (4–10) to be valid. The temperatures associated with the liquid states are the normal boiling points [*cf.* Eq. (4–11)].

The similarity between Eq. (4–14) for molecular states and Eq. (2–12) for states of the entire N-body system is not fortuitous. Equation (4–14) can be derived by the same *mathematical* formalism of Chapter 2. The ensemble is considered to be the N actual molecules in thermal contact with each other. The number of molecules n_j in the state with energy ε_j is found by maximizing a combinatorial factor similar to Eq. (1–77). This point of view was the one originally proposed by Boltzmann. It is valid only for systems in which the total energy is a sum of individual molecular energies, that is, only for dilute gases. The conceptual generalization of these ideas by Gibbs was a magnificent achievement, which allowed statistical thermodynamics to be applicable to all physical systems. Furthermore, the derivation given in Chapter 2 is rigorous, since macroscopic systems can be labeled, and the size of the ensemble can be increased arbitrarily. This is not so for the Boltzmann approach, since the molecules cannot be labeled, and the system is finite.

Equation (4–14) can be reduced further if we assume that the energy of the molecule can be written in the form [*cf.* Eq. (4–1)]

$$\varepsilon = \varepsilon_i^{\text{trans}} + \varepsilon_j^{\text{rot}} + \varepsilon_k^{\text{vib}} + \varepsilon_l^{\text{elec}} + \cdots$$

Then Eq. (4–14) and Eq. (4–5) can be combined to give, for example,

$$\pi_j^{\text{vib}} = \frac{e^{-\varepsilon_j^{\text{vib}}}}{q_{\text{vib}}} \tag{4–15}$$

for the probability that a molecule is in the jth vibrational state irrespective of the other degrees of freedom.

Although Eq. (4–10) is applicable to most systems, it is important to complete the development of systems of independent, indistinguishable particles by evaluating Eq. (4–8) for the general case. The exact evaluation of Eq. (4–8) is necessary for several systems that we shall study. We must return, then, to a consideration of the effect of the symmetry requirements of N-body wave functions on the sum over states in Eq. (4–8).

4–2 FERMI-DIRAC AND BOSE-EINSTEIN STATISTICS

There are two cases to consider in the evaluation of Eq. (4–8). The resultant distribution function in the case of fermions is called Fermi-Dirac statistics, and that in the case of bosons is called Bose-Einstein statistics. Since all known particles are either fermions or bosons, these two "statistics" are the only exact distributions. We shall see, however, that in the case of high temperature and/or low density, both of these distributions go over into the Boltzmann or classical distribution.

It is most convenient to treat the general case by means of the grand canonical ensemble for reasons that we shall see shortly. Let $E_j(N, V)$ be the energy states available to a system containing N molecules. Let ε_k be the molecular quantum states. Finally, let $n_k = n_k(E_j)$ be the number of molecules in the kth molecular state when the system itself is in the quantum state with energy E_j. A quantum state of the entire system is specified by the set $\{n_k\}$. The energy of the system is

$$E_j = \sum_k \varepsilon_k n_k \tag{4–16}$$

and, of course,

$$N = \sum_k n_k \tag{4–17}$$

We can write $Q(N, V, T)$ as

$$Q(N, V, T) = \sum_j e^{-\beta E_j} = \sum_{\{n_k\}}{}^{*} e^{-\beta \sum_i \varepsilon_i n_i} \tag{4–18}$$

where the asterisk in the summation signifies the restriction that

$$\sum n_k = N$$

This restriction turns out to be mathematically awkward. We can avoid this restriction by using the grand canonical partition function instead. This will be an excellent example where one partition function is much easier to evaluate than another. Since we have demonstrated the equivalence of ensembles, we are free to make the choice strictly on mathematical convenience. We then use

$$\Xi(V, T, \mu) = \sum_{N=0}^{\infty} e^{\beta\mu N} Q(N, V, T)$$

We use Eq. (4–18) for $Q(N, V, T)$ and the absolute activity $\lambda = e^{\beta\mu}$ to get

$$\begin{aligned}\Xi(V, T, \mu) &= \sum_{N=0}^{\infty} \lambda^N \sum_{\{n_k\}}{}^{*} e^{-\beta \sum_i \varepsilon_i n_i} \\ &= \sum_{N=0}^{\infty} \sum_{\{n_k\}}{}^{*} \lambda^{\sum n_i} e^{-\beta \sum_j \varepsilon_j n_j} \\ &= \sum_{N=0}^{\infty} \sum_{\{n_k\}}{}^{*} \prod_k (\lambda e^{-\beta\varepsilon_k})^{n_k}\end{aligned} \tag{4–19}$$

Now comes the crucial step (which requires some thought). Since we are summing over all values of N, each n_k will range over all possible values, and Eq. (4–19) can be written as (see Problem 4–6)

$$\Xi(V, T, \mu) = \sum_{n_1=0}^{n_1{}^{\max}} \sum_{n_2=0}^{n_2{}^{\max}} \cdots \prod_k (\lambda e^{-\beta\varepsilon_k})^{n_k} \tag{4–20}$$

Equations (4–19) and (4–20) are completely equivalent. Equation (4–20) can be written in a more lucid form:

$$\Xi(V, T, \mu) = \sum_{n_1=0}^{n_1{}^{\max}} (\lambda e^{-\beta\varepsilon_1})^{n_1} \sum_{n_2=0}^{n_2{}^{\max}} (\lambda e^{-\beta\varepsilon_2})^{n_2} \cdots$$

or

$$= \prod_k \sum_{n_k=0}^{n_k{}^{\max}} (\lambda e^{-\beta\varepsilon_k})^{n_k} \tag{4–21}$$

Equation (4–21) is a simple product and is a general result. The crucial step in this series of equations is the step from Eq. (4–19) to Eq. (4–20), from which Eq. (4–21) follows immediately. The step from Eqs. (4–19) to (4–20) is possible only because we are summing over all values of N, or, in other words, since we are using the grand canonical partition function.

We now apply Eq. (4–21) to fermions and bosons. In *Fermi-Dirac* statistics, each of the n_k in Eq. (4–21) can be only either 0 or 1, since no two particles can be in the same quantum state. In this case $n_1{}^{\max} = 1$, and Eq. (4–21) is simply

$$\Xi_{\text{FD}} = \prod_k (1 + \lambda e^{-\beta\varepsilon_k}) \tag{4–22}$$

where FD, of course, signifies Fermi-Dirac.

In *Bose-Einstein* statistics, on the other hand, the n_k can be 0, 1, 2, ..., since there is no restriction on the occupancy of each state. Therefore, $n_k^{\max} = \infty$, and Eq. (4–21) becomes

$$\Xi_{\mathrm{BE}} = \prod_k \sum_{n_k=0}^{\infty} (\lambda e^{-\beta\varepsilon_k})^{n_k} = \prod_k (1 - \lambda e^{-\beta\varepsilon_k})^{-1} \qquad \lambda e^{-\beta\varepsilon_k} < 1 \tag{4–23}$$

To get Eq. (4–23), we have used the fact that

$$\sum_{j=0}^{\infty} x^j = (1 - x)^{-1}$$

for $x < 1$.

Equations (4–22) and (4–23) are the two fundamental distributions of the statistical thermodynamics of systems of independent particles. We can combine these two equations into

$$\Xi_{\substack{\mathrm{FD}\\ \mathrm{BE}}} = \prod_k (1 \pm \lambda e^{-\beta\varepsilon_k})^{\pm 1} \tag{4–24}$$

where as the notation indicates, the upper sign refers to Fermi-Dirac statistics, and the lower sign refers to Bose-Einstein statistics.

Using Eq. (3–33), we see that

$$\bar{N} = N = \sum_k \bar{n}_k = kT\left(\frac{\partial \ln \Xi}{\partial \mu}\right)_{V,T} = \lambda\left(\frac{\partial \ln \Xi}{\partial \lambda}\right)_{V,T} = \sum_k \frac{\lambda e^{-\beta\varepsilon_k}}{1 \pm \lambda e^{-\beta\varepsilon_k}} \tag{4–25}$$

The average number of particles in the kth quantum state is

$$\bar{n}_k = \frac{\lambda e^{-\beta\varepsilon_k}}{1 \pm \lambda e^{-\beta\varepsilon_k}} \tag{4–26}$$

Equation (4–26) is the quantum statistical counterpart of Eq. (4–14). We multiply Eq. (4–26) by ε_k and sum over k to get the quantum statistical version of Eq. (4–13).

$$\bar{E} = N\bar{\varepsilon} = \sum_k \bar{n}_k \varepsilon_k = \sum_k \frac{\lambda \varepsilon_k e^{-\beta\varepsilon_k}}{1 \pm \lambda e^{-\beta\varepsilon_k}} \tag{4–27}$$

Lastly, Eq. (3–16) gives

$$pV = \pm kT \sum_k \ln[1 \pm \lambda e^{-\beta\varepsilon_k}] \tag{4–28}$$

Equations (4–25) through (4–28) are the fundamental formulas of Fermi-Dirac (+) and Bose-Einstein (−) statistics. Note that the molecular partition function q is not a relevant quantity when we are dealing with quantum statistics, that is, Fermi-Dirac or Bose-Einstein statistics. In spite of the fact that we have neglected intermolecular forces, the individual particles of the system are not independent because of the symmetry requirements of the wave functions.

We noted above that both kinds of statistics should go over into Boltzmann or classical statistics in the limit of high temperature or low density, where the number of available molecular quantum states is much greater than the number of particles. This condition implies that the average number of molecules in any state is very small, since most states will be unoccupied and those few states that are occupied will most likely contain only one molecule. This means that $\bar{n}_k \to 0$ in Eq. (4–26). This is achieved by letting $\lambda \to 0$. Thermodynamically, this means the limit of $N/V \to 0$ for fixed T, or $T \to \infty$ for fixed N/V. (See Problem 4–3.) For small λ, Eq. (4–26) becomes

$$\bar{n}_k = \lambda e^{-\beta\varepsilon_k} \qquad (\lambda \text{ small})$$

If we sum both sides of this equation over k to eliminate λ, we have

$$\frac{\bar{n}_k}{N} = \frac{e^{-\beta\varepsilon_k}}{q} \tag{4–29}$$

where

$$q = \sum_j e^{-\beta\varepsilon_j} \tag{4–30}$$

Equation (4–26) then goes over to the Boltzmann or classical limit for both Fermi-Dirac and Bose-Einstein statistics.

Equations (4–27) and (4–28) also reduce to the formulas of Section 4–1 as $\lambda \to 0$. Equation (4–27) becomes

$$\bar{E} \to \sum_j \lambda\varepsilon_j e^{-\beta\varepsilon_j}$$

and since $n_j \to \lambda e^{-\beta\varepsilon_j}$, we have

$$\bar{\varepsilon} = \frac{\bar{E}}{\bar{N}} \to \frac{\sum_j \varepsilon_j e^{-\beta\varepsilon_j}}{\sum_j e^{-\beta\varepsilon_j}} \tag{4–31}$$

This is the same as Eq. (4–13). Similarly, for small λ we can expand the logarithm in Eq. (4–28) to get

$$pV \to (\pm kT)\left(\pm\lambda \sum_j e^{-\beta\varepsilon_j}\right) \tag{4–32}$$

We have used the fact that $\ln(1 + x) \approx x$ for small x. Using Eq. (4–30), this becomes

$$pV = \lambda kT \sum_j e^{-\beta\varepsilon_j} = \lambda kTq \tag{4–33}$$

or

$$\beta pV = \ln \Xi = \lambda q \tag{4–34}$$

Equation (3–33) can be used to show that $\lambda q = N$, and so Eq. (4–34) is the perfect gas law as expected. Thus the formulas of Fermi-Dirac and Bose-Einstein statistics reduce to those of Boltzmann statistics in the classical limits.

We can also derive Eq. (4–10) directly from Eq. (4–34) for Ξ:

$$\Xi = e^{\lambda q} = \sum_{N=0}^{\infty} \frac{(\lambda q)^N}{N!}$$

If we compare this to Eq. (3–15), see that

$$Q(N, V, T) = \frac{q^N}{N!}$$

We shall defer a discussion of the equations of Fermi-Dirac and Bose-Einstein statistics to Chapter 10. There are a few systems such as electrons in metals, liquid helium, electromagnetic radiation, for which one must use quantum statistics. For most systems that we shall study in this book, however, we shall be able to use Boltzmann or classical statistics. In the next chapter we shall apply the limit of Boltzmann statistics to the simplest system, namely, a monatomic ideal gas.

ADDITIONAL READING

General

ANDREWS, F. C. 1963. *Equilibrium statistical mechanics.* New York: Wiley. Chapter 17.

HILL, T. L. 1960. *Statistical thermodynamics.* Reading, Mass.: Addison-Wesley. Chapter 3.

KITTEL, C. 1969. *Thermal physics.* New York: Wiley. Chapter 9.

KNUTH, E. 1966. *Statistical thermodynamics.* New York: McGraw-Hill. Chapter 6.

KUBO, R. 1965. *Statistical mechanics.* Amsterdam: North-Holland Publishing Co. Section 1–15 and Example 1–12.

MANDL, F. 1971. *Statistical physics.* New York: Wiley. Chapter 11.

REIF, F. 1965. *Statistical and thermal physics.* New York: McGraw-Hill. Sections 9–1 through 9–8.

RUSHBROOKE, G. S. 1949. *Statistical mechanics.* London: Oxford University Press. Chapters 2 and 3.

TOLMAN, R. C. 1938. *Statistical mechanics.* London: Oxford University Press. Chapters 10 and 14.

PROBLEMS

4–1. Calculate the temperature below which each of the substances listed below cannot be treated classically at 1 atmosphere. Compare this with the normal boiling temperature for each substance.

$$\text{He, Ne, Ar, Kr, } CO_2,\ N_2,\ H_2,\ Cl_2,\ H_2O$$

4–2. Show that the quantity

$$\frac{6N}{\pi V}\left(\frac{h^2}{12mkT}\right)^{3/2}$$

given in Table 4–1 is indeed very large for electrons in metals at room temperature.

4–3. Show that the condition that $\lambda \to 0$ corresponds thermodynamically to the limit $N/V \to 0$ for fixed T, or $T \to \infty$ for fixed density. Remember that $\lambda = e^{\beta\mu}$.

4–4. In deriving the limiting case of Boltzmann statistics, we claimed that if the number of quantum states M far exceeds the number of particles N, then the terms in the product of the molecular partition functions in which each particle is in a different quantum state constitute the overwhelming number of terms. Show, in fact, that the ratio of this type of term to the total number of terms approaches unity as $N/M \to 0$, N and M both large. Hint: remember that

$$\lim_{x\to\infty}\left(1+\frac{a}{x}\right)^x = e^a$$

4–5. For an ideal gas, show that the relation

$$P = \frac{2}{3}\frac{E_{\text{kin}}}{V}$$

holds irrespective of its statistics, where E_{kin} is the total kinetic energy.

4–6. To convince yourself of the step leading from Eq. (4–19) to Eq. (4–20), consider the summation

$$S = \sum_{N=0}^{\infty} \sum_{\{n_j\}}^{*} x_1^{n_1} x_2^{n_2}$$

where n_1 and $n_2 = 0$, 1, and 2. Show by directly expanding S for this simple case that this is equivalent to [Eq. (4–21)]

$$S = \prod_{K=1}^{2} (1 + x_K + x_K^2)$$

4–7. Recall that the equation of state for an ideal quantum gas is

$$pV = kT \ln \Xi = \pm kT \sum_j \ln\left[1 \pm \lambda e^{-\epsilon_j/kT}\right]$$

where $\lambda = e^{\mu/kT}$. Using the fact that the summation over states can be replaced by an integration over energy levels

$$\omega(\varepsilon)\, d\varepsilon = 2\pi\left(\frac{2m}{h^2}\right)^{3/2} V\varepsilon^{1/2}\, d\varepsilon$$

derive the quantum virial expansion

$$\frac{p}{kT} = \mp\frac{1}{\Lambda^3}\sum_{j=1}^{\infty}\frac{(\mp 1)^j\lambda^j}{j^{5/2}}$$

where $\Lambda = (h^2/2\pi mkT)^{1/2}$.

4–8. Show that the entropy of an ideal quantum gas can be written as

$$S = -k\sum_j [\bar{n}_j \ln \bar{n}_j \pm (1 \mp \bar{n}_j)\ln(1 \mp \bar{n}_j)]$$

where the upper (lower) sign denotes Fermi-Dirac (Bose-Einstein) statistics.

4–9. Show that $pV \geq \langle N\rangle kT$ for fermions, and $pV \leq \langle N\rangle kT$ for bosons.

4–10. Consider a system of independent, distinguishable particles, each of which has only two accessible states; a ground state of energy 0 and an excited state of energy ε. If the system is in equilibrium with a heat bath of temperature T, calculate A, E, S, and C_v. Sketch C_v versus T. Does the choice of the ground-state energy $= 0$ affect P, C_v, or S? How would your results change if ε_0 were added to both energy values?

4–11. Generalize Eq. (4–10) to the case of a mixture of several different species of non-interacting particles.

4–12. Consider a system of N distinguishable independent particles, each of which can be in the state $+\varepsilon_0$ or $-\varepsilon_0$. Let the number of particles with energy $\pm\varepsilon_0$ be $N_\pm$, so that the energy is

$$E = N_+\varepsilon_0 - N_-\varepsilon_0 = 2N_+\varepsilon_0 - N\varepsilon_0$$

Evaluate the partition function Q by summing $\exp(-E/kT)$ over levels and compare your result to $Q = q^N$. Do not forget the degeneracy of the levels, which in this case is the number of ways that N_+ particles out of N can be in the $+$ state. Calculate and plot the heat capacity C_V for this system.

4–13. The vibrational energy levels of a diatomic molecule can be approximated by a quantum mechanical harmonic oscillator. The fundamental vibrational frequency ν is $0(10^{13}\ \text{sec}^{-1})$ for many diatomic molecules. Calculate the fraction of molecules in the first few vibrational levels in an ideal diatomic gas at 25°C. Derive a closed expression for the fraction of molecules in all excited states.

4–14. The rotational energy of diatomic molecules can be well approximated by a quantum mechanical rigid rotor. According to Eq. (1–32), the energy levels depend upon the moment of inertia, which for a diatomic molecule is $0(10^{-40}\ \text{g-cm}^2)$. Calculate and plot the population of rotational levels of a diatomic ideal gas at 25°C. Do not forget to include the degeneracy $2J + 1$.

4–15. Show that $Q(N, V, T) = [q(V, T)]^N/N!$ implies that $q(V, T) = f(T)V$. Do this in both the canonical ensemble and grand canonical ensemble formalisms.

4–16. Show that the most probable distribution of $2N$ molecules of an ideal gas contained in two equal and connected volumes at the same temperature is N molecules in each volume.

4–17. In Fermi-Dirac statistics, the maximum occupancy of any state is 1, while in Bose-Einstein statistics, it is ∞. All particles appear to obey one of these two statistics. In 1940, however, Gentile* investigated the implications of an intermediate statistics, in which the maximum occupancy is m. Derive the distribution law for this case.

---

* G. Gentile, *Nuovo Cimento* **17**, p. 493, 1940.

4–18. Derive the equation

$$C_V = \frac{N}{kT^2}[\overline{\varepsilon^2} - \bar{\varepsilon}^2]$$

for independent particles and show that fluctuations of molecular energies are not at all negligible.

4–19. Starting from Eqs. (4–13) and (4–14), show that the fluctuations in ε, the energy of a single particle, are not small, and in fact, are given by

$$\frac{\sigma_\varepsilon}{\bar{\varepsilon}} = N^{1/2}\frac{\sigma_E}{\bar{E}}$$

4–20. Consider a gas in equilibrium with the surface of a solid. Some of the molecules of the gas will be adsorbed onto the surface, and the number adsorbed will be a function of the pressure of the gas. A simple statistical mechanical model for this system is to picture the solid surface to be a two-dimensional lattice of M sites. Each of these sites can be either unoccupied, or occupied by at most one of the molecules of the gas. Let the partition function of an unoccupied site be 1 and that of an occupied site be $q(T)$. (We do not need to know $q(T)$ here.) Assuming that molecules adsorbed onto the lattice sites do not interact with each other, the partition function of N molecules adsorbed onto M sites is then

$$Q(N, M, T) = \frac{M!}{N!(M-N)!}[q(T)]^N$$

The binomial coefficient accounts for the number of ways of distributing the N molecules over the M sites. By using the fact the adsorbed molecules are in equilibrium with the gas phase molecules (considered to be an ideal gas), derive an expression for the fractional coverage, $\theta \equiv N/M$, as a function of the pressure of the gas. Such an expression, that is, $\theta(p)$, is called an adsorption isotherm, and this model gives the so-called Langmuir adsorption isotherm.

4–21. Consider a lattice of M equivalent noninteracting magnetic dipoles, μ (associated, say, with electron or nuclear spins). When placed in a magnetic field H, each dipole can orient itself either in the same direction, ↑, or opposed to, ↓, the field. The energy of a dipole is $-\mu H$ if oriented with the field, and $+\mu H$ if oriented against the field. Let N be the number of ↓ states and $M - N$ the number of ↑ states. For a given value of N, the total energy is

$$\mu HN - \mu H(M-N) = (2N - M)\mu H$$

The total magnetic moment I is

$$I = (M - 2\bar{N})\mu$$

where $\bar{N}$ is the average value of N for a given M, H, and T. The work necessary to increase H by dH is $-IdH$. Find the specific heat C and the total magnetic moment for this system, and sketch both I versus $\mu H/kT$, that is, the total magnetization versus the applied field, and C/Nk versus $kT/\mu H$.

4–22. (a) Consider a system of M independent and distinguishable macromolecules on which any number from 0 to m small molecules may bind. Let $q(j)$ be the macromolecular partition function when j molecules are bound. If there are N small molecules (or ions) and M macromolecules (say proteins), then

$$Q(N, M, T) = \sum_{\mathbf{a}}{}^{*} \frac{M!\, q(0)^{a_0} q(1)^{a_1} \cdots q(m)^{a_m}}{a_0!\, a_1! \cdots a_m!}$$

where the number of macromolecules having j bound molecules is a_j, and where the asterisk indicates the restrictions

$$\sum_{j=0}^{m} a_j = M \qquad \sum_{j=0}^{m} ja_j = N$$

Show that the grand partition function for this system can be written in the form

$$\Xi(M, T, \mu) = \xi(\mu, T)^M$$

where

$$\xi(\mu, T) = q(0) + q(1)\lambda + \cdots + q(m)\lambda^m$$

Interpret this result.

(b) Extend this result to the case in which the macromolecules are not distinguishable.

CHAPTER 5

IDEAL MONATOMIC GAS

In this chapter, we shall apply the general results of the preceding chapters to an ideal monatomic gas. By ideal, we mean a gas dilute enough that intermolecular interactions can be neglected. The results that we derive here will be applicable to real monatomic gases at pressures and temperatures for which the equation of state is well represented by $pV = NkT$, that is, pressures below 1 atmosphere and temperatures greater than room temperature.

We have shown in Section 4–1 that the number of available quantum states far exceeds the number of particles for an ideal gas. Thus we can write the partition function of the entire system in terms of the individual atomic partition functions:

$$Q(N, V, T) = \frac{[q(V, T)]^N}{N!} \tag{5–1}$$

A monatomic gas has translational, electronic, and nuclear degrees of freedom. The translational Hamiltonian is separable from the electronic and nuclear degrees of freedom, and the electronic and nuclear Hamiltonians are separable to a very good approximation. Thus we have

$$q(V, T) = q_{\text{trans}}\, q_{\text{elect}}\, q_{\text{nucl}} \tag{5–2}$$

We shall study each of these factors separately in the following sections of this chapter.

5–1 THE TRANSLATIONAL PARTITION FUNCTION

In this section we shall evaluate the translational partition function. The energy states are given by

$$\varepsilon_{n_x n_y n_z} = \frac{h^2}{8ma^2}(n_x^{\,2} + n_y^{\,2} + n_z^{\,2}) \qquad n_x, n_y, n_z = 1, 2, \ldots \tag{5–3}$$

We substitute this into q_{trans} to get

$$q_{\text{trans}} = \sum_{n_x, n_y, n_z = 1}^{\infty} e^{-\beta \varepsilon_{n_x n_y n_z}}$$

$$\sum_{n_x=1}^{\infty} \exp\left(-\frac{\beta h^2 n_x^2}{8ma^2}\right) \sum_{n_y=1}^{\infty} \exp\left(-\frac{\beta h^2 n_y^2}{8ma^2}\right) \sum_{n_z=1}^{\infty} \exp\left(-\frac{\beta h^2 n_z^2}{8ma^2}\right)$$

$$= \left(\sum_{n=1}^{\infty} \exp\left(-\frac{\beta h^2 n^2}{8ma^2}\right)\right)^3 \tag{5-4}$$

This summation cannot be evaluated in closed form, that is, it cannot be expressed in terms of any simple analytic function. This does not present any difficulty, however, for the following reason. The successive terms in these summations differ so little from each other that the terms vary essentially continuously, and so the summation can, for all practical purposes, be replaced by an integral. To prove this, we show that the argument of the exponential changes little in going from n_x to $n_x + 1$. This difference, Δ, is given by

$$\Delta = \frac{\beta h^2 (n_x + 1)^2}{8ma^2} - \frac{\beta h^2 n_x^2}{8ma^2} \doteq \frac{\beta h^2 (2n_x + 1)}{8ma^2}$$

At room temperature, for $m = 10^{-22}$ g and $a = 10$ cm, this difference is

$$\Delta \approx (2n_x + 1) \times 10^{-20}$$

A typical value of n_x at room temperature is $O(10^{10})$ (see Problem 5–3), so Δ is indeed very small for all but very large values of n_x. A value of n_x for which Δ is as large as 10^{-5} would correspond to an energy of $(10^{10}kT)$, an extremely improbable energy. Thus we can replace the summation in Eq. (5–4) by an integration:

$$q_{\text{trans}}(V, T) = \left(\int_0^{\infty} e^{-\beta h^2 n^2 / 8ma^2}\, dn\right)^3 = \left(\frac{2\pi mkT}{h^2}\right)^{3/2} V \tag{5-5}$$

where we have written V for a^3.

It is instructive to evaluate q_{trans} in another way. Equation (5–4) is a sum over the states of the system. We could also write q_{trans} as a sum over levels. Recognizing that the levels are very densely distributed, we can write q_{trans} as an integral:

$$q_{\text{trans}} = \int_0^{\infty} \omega(\varepsilon) e^{-\beta\varepsilon}\, d\varepsilon \tag{5-6}$$

The function $\omega(\varepsilon)$ is the number of energy states between ε and $\varepsilon + d\varepsilon$, or, in other words, the effective degeneracy. Equation (5–6) is simply a continuous form of a sum over levels rather than a sum over states. We have already evaluated $\omega(\varepsilon)$ in Section 1–3. It is given by Eq. (1–35)

$$\omega(\varepsilon)\, d\varepsilon = \frac{\pi}{4}\left(\frac{8ma^2}{h^2}\right)^{3/2} \varepsilon^{1/2}\, d\varepsilon \tag{5-7}$$

If we substitute this into Eq. (5–6), we get

$$q_{\text{trans}} = \frac{\pi}{4}\left(\frac{8ma^2}{h^2}\right)^{3/2} \int_0^{\infty} \varepsilon^{1/2} e^{-\beta\varepsilon}\, d\varepsilon$$

$$= \left(\frac{2\pi mkT}{h^2}\right)^{3/2} V \tag{5-8}$$

Of course, we obtain the same result as Eq. (5–5).

The factor $(h^2/2\pi mkT)^{1/2}$ that occurs in the translational partition function has units of length and is usually denoted by Λ. In this notation, Eq. (5–5) or (5–8) read

$$q_{\text{trans}} = \frac{V}{\Lambda^3} \tag{5–9}$$

The quantity Λ can be given the following interpretation. The average translational or kinetic energy of an ideal gas molecule can be calculated immediately from Eq. (5–8) and Eq. (4–13), which in terms of q_{trans} is

$$\bar{\varepsilon}_{\text{trans}} = kT^2 \left(\frac{\partial \ln q_{\text{trans}}}{\partial T}\right)$$

We find that $\bar{\varepsilon}_{\text{trans}} = \frac{3}{2}kT$, and since $\varepsilon_{\text{trans}} = p^2/2m$, where p^2 is the momentum of the particle, we can say that the average momentum is essentially $(mkT)^{1/2}$. Thus Λ is essentially h/p, which is equal to the De Broglie wavelength of the particle. Consequently, Λ is called the thermal De Broglie wavelength. The condition for the applicability of classical or Boltzmann statistics is equivalent to the condition that $\Lambda^3/V \ll 1$, which physically says that the thermal De Broglie wavelength must be small compared to the dimensions of the container. Such a condition is similar to the condition that quantum effects decrease as the De Broglie wavelength becomes small (*cf.* Table 4–1).

5–2 THE ELECTRONIC AND NUCLEAR PARTITION FUNCTIONS

In this section we shall investigate the electronic and nuclear contributions to q.

It is more convenient to write the electronic partition function as a sum of levels rather than a sum over states. We have, then,

$$q_{\text{elect}} = \sum \omega_{ei} e^{-\beta\varepsilon_i} \tag{5–10}$$

where ω_{ei} is the degeneracy, and ε_i the energy of the *i*th electronic level. We first fix the arbitrary zero of energy such that $\varepsilon_1 = 0$, that is, we shall measure all of our electronic energies relative to the ground state. The electronic contribution to q can then be written as

$$q_{\text{elect}} = \omega_{e1} + \omega_{e2} e^{-\beta\Delta\varepsilon_{12}} + \cdots \tag{5–11}$$

where $\Delta\varepsilon_{1j}$ is the energy of the *j*th electronic level relative to the ground state. These $\Delta\varepsilon$'s are typically of the order of electron volts, and so $\beta\Delta\varepsilon$ is typically quite large at ordinary temperatures (see Problem 5–10). Therefore at ordinary temperatures, only the first term in the summation for q_{elect} is significantly different from zero. However, there are some cases, such as the halogen atoms, where the first excited state lies only a fraction of an electron volt above the ground state, so that several terms in q_{elect} are necessary. Even in these cases the sum converges extremely rapidly.

The electronic energies of atoms and ions are determined by atomic spectroscopy and are well tabulated. The standard reference is the tables of Moore* which list the energy levels and energies of many atoms and ions. Table 5–1 lists the first few levels for H, He, Li, O, and F. A look at this table will indicate that electronic states are labeled or characterized by a so-called term symbol, which is briefly explained in Section 5–4. (A knowledge of the meaning of atomic term symbols is not necessary for the calculation of q_{elect}, but they are explained in Section 5–4 for completeness.)

---

* See Table 5–1.

Table 5–1. **Atomic energy states**

atom	electron configuration	term symbol	degeneracy $g=2J+1$	energy (cm^{-1})	energy (eV)
H	1*s*	$^2S_{1/2}$	2	0	0
	2*p*	$^2P_{1/2}$	2	82258.907	10.20
	2*s*	$^2S_{1/2}$	2	82258.942	
	2*p*	$^2P_{3/2}$	4	82259.272	
He	1s^2	$^1S_0$	1	0	
	1*s*2*s*	$^3S_1$	3	159850.318	19.82
		$^1S_0$	1	166271.70	
Li	1$s^2$2*s*	$^2S_{1/2}$	2	0	
	1$s^2$2*p*	$^2P_{1/2}$	2	14903.66	1.85
		$^2P_{3/2}$	4	14904.00	
	1$s^2$3*s*	$^2S_{1/2}$	2	27206.12	
O	1$s^2$2$s^2$2p^4	$^3P_2$	5	0	
		$^3P_1$	3	158.5	0.02
		$^3P_0$	1	226.5	0.03
		$^1D_2$	5	15867.7	1.97
		$^1S_0$	1	33792.4	4.19
F	1$s^2$2$s^2$2p^5	$^2P_{3/2}$	4	0	
		$^2P_{1/2}$	2	404.0	0.05
	1$s^2$2$s^2$2$p^4$3*s*	$^4P_{5/2}$	6	102406.50	12.70
		$^4P_{3/2}$	4	102681.24	
		$^4P_{1/2}$	2	102841.20	
		$^2P_{3/2}$	4	104731.86	
		$^2P_{1/2}$	2	105057.10	

Source: C. E. Moore, "Atomic Energy States," *Natl. Bur. Standards, Circ.*, **1**, p. 467, 1949.

Some general observations about Table 5–1 are: All the rare gases have a ground state $^1S_0$ (called a singlet *S*) with the first excited state O(10 eV) higher; the alkali metals have a $^2S_{1/2}$ (called a doublet *S*) ground state with the next state O(1 eV) higher; the halogen atoms have a $^2P_{3/2}$ (called a doublet *P*) ground state with the next one, a $^2P_{1,2}$ (also a doublet *P*) only O(0.1 eV) higher. Thus at ordinary temperatures the electronic partition function of the rare gases is essentially unity and that of the alkali metals is 2, while those for halogen atoms consist of two terms.

Using the data in Table 5–1, we can now calculate the fraction of He atoms in the lowest triplet state, $^3S_1$. This fraction is given by

$$f_2=\frac{\omega_{e2}e^{-\beta\Delta\varepsilon_{12}}}{\omega_{e1}+\omega_{e2}e^{-\beta\Delta\varepsilon_{12}}+\omega_{e3}e^{-\beta\Delta\varepsilon_{13}}+\cdots}$$

$$=\frac{3e^{-\beta\Delta\varepsilon_{12}}}{1+3e^{-\beta\Delta\varepsilon_{12}}+\omega_{e3}e^{-\beta\Delta\varepsilon_{13}}+\cdots}\qquad(5\text{–}12)$$

At 300°K, $\beta\,\Delta\varepsilon_{12}=770$, and so $f_2\approx 10^{-334}$. Even at 3000°K, $f_2\approx 10^{-33}$. This is typical of the rare gases. The energy separation must be less than a few hundred cm^{-1} or so before any population of that level is significant. Incidentally, it is useful to know that Boltzmann's constant in units of cm^{-1}/deg-molecule is 0.695 (almost ln 2), and 1 eV = 8065.73 cm^{-1}.

Table 5–2 gives the fraction of fluorine atoms in the first excited electronic state as a function of temperature. It can be seen that fluorine is a case where it is necessary to use two terms in q_{elec}.

We shall write the electronic partition function as

$$q_{\text{elec}}(T)\approx\omega_{e1}+\omega_{e2}e^{-\beta\varepsilon_{12}}\qquad(5\text{–}13)$$

Table 5–2. **The fraction of fluorine atoms in the first excited electronic state as a function of temperature**

T(°K)	f_2
200	0.027
400	0.105
600	0.160
800	0.195
1000	0.219
1200	0.236
2000	0.272

but at temperatures at which the second term is not negligible with respect to the first term, we must check the possible contribution of higher terms as well. This will rarely be necessary, however.

We now consider the nuclear partition function. The nuclear partition function has a form similar to that of the electronic partition function. Nuclear energy levels are separated by millions of electron volts, however, which means that it requires temperatures of the order of 10^{10}°K to produce excited nuclei. At terrestrial temperatures then, we need consider only the first term, that is, the degeneracy of the ground nuclear state ω_{n1}. We take our zero of nuclear energy states to be the ground state. Note that we have taken the overall atomic ground state to be the atom in its ground translational, electronic, and nuclear states. The nuclear partition function, $q_n = \omega_{n1}$, then contributes only a multiplicative constant to Q, and hence affects only the entropy and free energies by a constant additive factor. Since the nuclear state is rarely altered in any chemical process, it does not contribute to thermodynamic changes, and so we shall usually not include it in q. We cannot do this in the case of the electronic contribution since there are many chemical processes in which the electronic states change.

This completes the partition function of monatomic ideal gases. In summary then, we have

$$Q = \frac{(q_{\text{trans}} q_{\text{elec}} q_{\text{nucl}})^N}{N!} \tag{5-14}$$

where

$$q_{\text{trans}} = \left(\frac{2\pi mkT}{h^2}\right)^{3/2} V = \frac{V}{\Lambda^3}$$

$$q_{\text{elect}} = \omega_{e1} + \omega_{e2} e^{-\beta \Delta\varepsilon_{12}} + \cdots$$

$$q_{\text{nucl}} = \omega_{n1} + \cdots \tag{5-15}$$

The nuclear partition function, although not always equal to unity, is usually omitted. We can now calculate thermodynamic properties of a monatomic ideal gas.

5–3 THERMODYNAMIC FUNCTIONS

The Helmholtz free energy is given by

$$A(N, V, T) = -kT \ln Q = -NkT \ln\left[\left(\frac{2\pi mkT}{h^2}\right)^{3/2} \frac{Ve}{N}\right] - NkT \ln(\omega_{e1} + \omega_{e2} e^{-\beta\Delta\varepsilon_{12}}) \tag{5-16}$$

The argument of the first logarithm here is much larger than the argument of the second logarithm, and so the electronic contribution to A is quite small.

The thermodynamic energy is

$$E = kT^2\left(\frac{\partial \ln Q}{\partial T}\right)_{N,V} = \tfrac{3}{2}NkT + \frac{N\omega_{e2}\Delta\varepsilon_{12}\,e^{-\beta\Delta\varepsilon_{12}}}{q_{\text{elec}}} + \tag{5–17}$$

The contribution of the electronic degrees of freedom to the energy is small at ordinary temperatures (see Problem 5–12). Since we have neglected the contribution of the intermolecular potential to the total energy of the gas, the first term of Eq. (5–17) represents only kinetic energy. Furthermore, each atom has an average kinetic energy $3kT/2$, or $kT/2$ for each degree of translational freedom. We shall give an interesting interpretation of this result when we study classical statistical mechanics in Chapter 7. If we ignore the very small contribution from the electronic degrees of freedom, the molar heat capacity at constant volume is $3Nk/2$, a well-known experimental result for dilute gases.

The pressure is

$$p = kT\left(\frac{\partial \ln Q}{\partial V}\right)_{N,T} = \frac{NkT}{V} \tag{5–18}$$

Note that Eq. (5–18) results because $q(V, T)$ is of the form $f(T)V$, and the only contribution to the pressure is from the translational energy of the atoms. This is what one expects intuitively, since the pressure is due to bombardment of the walls of the container by the atoms and molecules of the gas.

The entropy is given by

$$S = \tfrac{5}{2}Nk + Nk\ln\left[\left(\frac{2\pi mkT}{h^2}\right)^{3/2}\frac{Ve}{N}\right] + Nk\ln(\omega_{e1} + \omega_{e2}e^{-\beta\Delta\varepsilon_{12}}) + \frac{Nk\omega_{e2}\beta\Delta\varepsilon_{12}\,e^{-\beta\Delta\varepsilon_{12}}}{q_{\text{elec}}} \tag{5–19}$$

$$= Nk\ln\left[\left(\frac{2\pi mkT}{h^2}\right)^{3/2}\frac{Ve^{5/2}}{N}\right] + S_{\text{elec}} \tag{5–20}$$

In Eq. (5–20), S_{elect} denotes the last two terms of Eq. (5–19). Equation (5–20) is called the Sackur-Tetrode equation. Table 5–3 compares the results of this equation with experimental values for several monatomic gases.

Table 5–3. **Comparison of experimental entropies at 1 atm and $T = 298$°K to those calculated from the statistical thermodynamical equation for the entropy of an ideal monatomic gas***

	exp. (e.u.)	calc. (e.u.)
He	30.13	30.11
Ne	34.95	34.94
Ar	36.98	36.97
Kr	39.19	39.18
Xe	40.53	40.52
C	37.76	37.76
Na	36.72	36.70
Al	39.30	39.36
Ag	41.32	41.31
Hg	41.8	41.78

* The experimental values have been corrected for any nonideal gas behavior.

The chemical potential is

$$\begin{aligned}\mu(T, p) &= -kT\left(\frac{\partial \ln Q}{\partial N}\right)_{V,T} = -kT \ln \frac{q}{N} \\ &= -kT \ln\left[\left(\frac{2\pi mkT}{h^2}\right)^{3/2} \frac{V}{N}\right] - kT \ln q_e q_n \\ &= -kT \ln\left[\left(\frac{2\pi mkT}{h^2}\right)^{3/2} \frac{kT}{p}\right] - kT \ln q_e q_n \\ &= -kT \ln\left[\left(\frac{2\pi mkT}{h^2}\right)^{3/2} kT\right] - kT \ln q_e q_n + kT \ln p \\ &= \mu_0(T) + kT \ln p \end{aligned} \qquad (5\text{–}21)$$

where the last line is the thermodynamic equation for $\mu(T, p)$ for an ideal gas [*cf.* Eq. (1–66)]. Thus statistical thermodynamics yields an expression for $\mu_0(T)$:

$$\mu_0(T) = -kT \ln\left[\left(\frac{2\pi mkT}{h^2}\right)^{3/2} kT\right] - kT \ln q_e q_n \qquad (5\text{–}22)$$

The argument of the first logarithm here has units of pressure, but remember that there is a $kT \ln p$ term in Eq. (5–21) so that $\mu(T, p)$ itself has units of energy and is $O(kT)$.

5–4 A DIGRESSION ON ATOMIC TERM SYMBOLS

The electronic state of an atom is designated by a so-called atomic term symbol. Since one encounters atomic term symbols in the calculation of electronic partition functions, we discuss them in this section. (The quantum mechanical level of this section is above that of most of the book, and it is not necessary to read this section on first reading.)

In addition to the usual kinetic energy and electrostatic terms in the Hamiltonian of a many-electron atom, there are a number of magnetic and spin terms. The most important of these is the spin-orbit interaction term, which represents the interaction of the magnetic moment associated with the spin of the electron with the magnetic field generated by the electric current produced by its own orbital motion. There are other terms such as spin–spin and orbit–orbit interaction terms, but these are numerically much less important. The Hamiltonian can then be written as

$$H = -\frac{\hbar^2}{2m}\sum_j \nabla_j^2 - \sum_j \frac{Ze^2}{r_j} + \sum_{i<j} \frac{e^2}{r_{ij}} + \sum_j \xi(r_j)\mathbf{l}_j \cdot \mathbf{s}_j \qquad (5\text{–}23)$$

where $\mathbf{l}_j$ and $\mathbf{s}_j$ are the individual electronic orbital and spin angular momenta, respectively, and $\xi(r_j)$ is a scalar function of r_j, whose form is not necessary here. We can abbreviate this equation by writing

$$H = H_0 + H_{ee} + H_{so} \qquad (5\text{–}24)$$

where H_0 represents the first two terms (no interelectronic interactions), H_{ee} the third, and H_{so} the fourth.

For light atoms ($Z < 40$), H_{so} is small enough to be considered a small perturbation. If H_{so} is neglected altogether, it can be shown that the total orbital angular momentum **L** and the total spin angular momentum **S** are conserved (i.e., have "good" quantum numbers, or commute with $H_0 + H_{ee}$). However, in this case, the *individual* orbital and spin angular momenta are not conserved; hence they are not useful concepts. The eigenvalues of the square of the total orbital angular momentum operator $\hat{L}^2$ and the square of the total spin angular momentum operator $\hat{S}^2$ are $L(L + 1)\hbar^2$ and $S(S + 1)\hbar^2$, respectively. One often interprets these eigenvalues by saying that the orbital angular momentum has the value L or that the total spin is S, but it should always be borne in mind that the orbital or spin angular momentum itself is not an eigenoperator in quantum mechanics.

The quantities **L** and **S** are the vector sums of the individual orbital $\mathbf{l}_j$ and spin $\mathbf{s}_j$, angular momenta. The possible ways of adding these $\mathbf{l}_j$ or $\mathbf{s}_j$ are governed by quantum mechanics with the result being that only certain values of the quantum numbers L and S are allowed. In the case of two electrons, L can take on only the values $l_1 + l_2,\ l_1 + l_2 - 1, \quad |l_1 - l_2|$, with a similar result for S. What we really mean by this, of course, is that the only allowed values of the eigenvalues of $\hat{L}^2$ are $(l_1 + l_2)(l_1 + l_2 + 1)\hbar^2,\ (l_1 + l_2 - 1)(l_1 + l_2)\hbar^2, \quad (|l_1 - l_2|)(|l_1 - l_2| + 1)\hbar^2$. The addition of electronic angular momenta to obtain **L** for cases involving more than two electrons can be accomplished using a scheme which is a straightforward but rather tedious electron-by-electron extension of the above two electron systems. Actually, rather specialized and advanced techniques have been developed to handle this problem, but we need not be concerned with them here.

The electronic energy states are designated by a term symbol, part of which is given by ${}^{2S+1}L$. Terms with $L = 0, 1, 2, \ldots$ are denoted by $S, P, D,$

When the spin–orbit term is taken into account, **L** and **S** are no longer conserved (that is, do not commute with the total H), and only the total angular momentum, $\mathbf{J} = \mathbf{L} + \mathbf{S}$, is conserved. The eigenvalues of $\hat{J}^2 = (\hat{L} + \hat{S})^2$ are $J(J + 1)\hbar^2$, with a degeneracy $2J + 1$, corresponding to the $2J + 1$ eigenvalues of $\hat{J}_z$, namely, $J\hbar, (J - 1)\hbar, \ldots, -J\hbar$. Just as in the addition of $\mathbf{l}_1$ and $\mathbf{l}_2$ above, the allowed values of J are $L + S,\ L + S - 1, \ldots, |L - S|$. The spin–orbit term causes each of these values of J to have a slightly different energy, and so the value of J is included in the term symbol as a subscript to give ${}^{2S+1}L_J$ as a characterization of the electronic state of an atom.

Table 5-1 lists the first few electronic states for some of the first row atoms.

This light atom approximate coupling scheme, in which L and S are almost good quantum numbers (that is, not good quantum numbers because of the small spin–orbit perturbation term) and in which the total angular momentum **J** is found by adding **L** and **S**, is called Russell-Saunders or *L–S* coupling. As the atomic number of the atom becomes larger, the spin–orbit term becomes larger than the interelectronic repulsion term, and H_{ee} can be considered to be a small perturbation on the others. In this case L and S are no longer useful, and the individual total angular momenta, $\mathbf{j}_i = \mathbf{l}_i + \mathbf{s}_i$, become the approximately conserved quantities. One then couples the **j**'s to get the total angular momentum. This scheme is called *j–j* coupling and is applicable to heavier atoms. In spite of the deterioration of *L–S* coupling as Z increases, it is still approximately useful, and so the electronic states of even heavy atoms are designated by term symbols of the form ${}^{2S+1}L_J$.

We are ready to discuss an ideal gas of diatomic molecules. In addition to having translational and electronic degrees of freedom, diatomic molecules have rotational and vibrational degrees of freedom as well. It should be apparent at this point that the additional input into our statistical thermodynamical equations will be the rotation-vibration energy levels.

We could also leave the study of gases for a while and apply our general results to other systems such as solids and liquids. This would involve no more effort than continuing with ideal gases and perhaps would be a change of pace. We could go directly to Chapter 11, for example, but we shall finish gases before treating other systems.

ADDITIONAL READING

General

Andrews, F. C. 1963. *Equilibrium statistical mechanics.* New York: Wiley. Chapter 18.
Hill, T. L. 1960. *Statistical thermodynamics.* Reading, Mass.: Addison-Wesley. Chapter 4.
Kestin, J., and Dorfman, J. R. 1971. *A course in statistical thermodynamics.* New York: Academic. Sections 6–7 through 6–8.
Kittel, C. 1969. *Thermal physics.* New York: Wiley. Chapters 11 and 12.
Knuth, E. 1966. *Statistical thermodynamics.* New York: McGraw-Hill. Chapter 8.
Münster, A. 1969. *Statistical thermodynamics,* Vol. I. Berlin: Springer-Verlag. Sections 6–1 and 6–2.

Term symbols

Strauss, H. L. 1968. *Quantum mechanics.* Englewood Cliffs, N.J.: Prentice-Hall. Chapter 8.

PROBLEMS

5–1. Convert Boltzmann's constant $k = 1.3806 \times 10^{-16}$ ergs-molecule^{-1} deg^{-1} to cm^{-1}-molecule^{-1}-deg^{-1} and to eV-molecule^{-1}-deg^{-1}.

5–2. By considering the special case of an ideal gas, determine the order of magnitude of E, A, G, S, C_V, and μ. Express your answers in terms of N, kT, or Nk, whichever is appropriate.

5–3. Calculate the value of n_x, n_y, and n_z for the case $n_x = n_y = n_z$ for a hydrogen atom (atomic weight 1.00) in a box of dimensions 1 cc if the particle has a kinetic energy $3kT/2$, for $T = 27°\text{C}$. What significant fact does this calculation illustrate?

5–4. Calculate the entropy of Ne at 300°K and 1 atmosphere. Use the entropy, in turn, to estimate the "translational" degeneracy of the gas.

5–5. Calculate the entropy of 1 mole of argon at 298°K and 1 atm and compare this to Table 5–3.

5–6. The quantum mechanical energy of a particle confined to a rectangular parallelepiped of lengths a, b, and c is

$$\varepsilon_{n_x n_y n_z} = \frac{h^2}{8m}\left(\frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2}\right)$$

Show that the translational partition function for this geometry is the same as that of a cube of the same volume.

5–7. Given that the quantum mechanical energy levels of a particle in a two-dimensional box are

$$\varepsilon = \frac{h^2}{8ma^2}(s_x^2 + s_y^2) \qquad s_x, s_y = 1, 2, \ldots$$

First calculate the density of states $\omega(\varepsilon)\,d\varepsilon$, that is, the number of states between ε and $\varepsilon\ d\varepsilon$, and use this to find the translational partition function of a two-dimensional ideal gas. Then find the partition function by another method. And finally find the equation of state, the thermodynamic energy E, the heat capacity C_A, and the entropy. This is a model for a gas adsorbed onto a surface or for long-chain fatty acids on the surface of water, say, as long as the number of molecules per unit area is small enough.

5–8. Calculate the entropy of a mixture of 50 percent neon and 50 percent argon at 500°K and 10 atm, assuming ideal behavior.

5–9. Calculate the De Broglie wavelength of an argon atom at 25°C and compare this with the average interatomic spacing at 1 atm.

5–10. Evaluate $\beta\Delta\varepsilon_{\text{elec}}$ at room temperature, given that electronic energy levels are usually separated by energies of the order of electron volts.

5–11. Using the data in Table 5–1, calculate the population of the first few electronic energy levels of an oxygen atom at room temperature.

5–12. Show that the contribution of the electronic degrees of freedom to the total energy is small at ordinary temperatures [*cf.* Eq. (5–17)].

5–13. Generalize the results of this chapter to an ideal binary mixture. In particular, show that

$$Q = \frac{q_1^{N_1} q_2^{N_2}}{N_1!\, N_2!}$$

$$E = \tfrac{3}{2}(N_1 + N_2)kT$$

and

$$S = N_1 k \ln\left(\frac{Ve^{5/2}}{\Lambda_1{}^3 N_1}\right) + N_2 k \ln\left(\frac{Ve^{5/2}}{\Lambda_2{}^3 N_2}\right)$$

if we ignore q_{elec} and q_{nucl}.

5–14. Derive the standard thermodynamic formula for the entropy of mixing by starting with Eq. (5–20).

5–15. Calculate A, E, μ, C_V, and S for 1 mole of Kr at 25°C and 1 atm (assuming ideal behavior).

5–16. Show that the most probable distribution of $2N$ molecules of an ideal gas contained in two equal and connected volumes at the same temperature is N molecules in each volume.

5–17. Evaluate the isothermal-isobaric partition function of a monatomic ideal gas by converting the summation over V in Eq. (3–17) to an integral. The result is

$$\Delta(N, p, T) = \left(\frac{kT}{p\Lambda^3}\right)^N$$

Using the fact that $G = -kT \ln \Delta$, derive expressions for S and V.

5–18. Consider a monatomic ideal gas of N particles in a volume V. Show that the number n of particles in some small subvolume v is given by the Poisson distribution

$$\begin{aligned} P_n &= (\lambda q)^n \frac{e^{-\lambda q}}{n!} \\ &= (\bar{n})^n \frac{e^{-\bar{n}}}{n!} \end{aligned}$$

Hint: Use the grand canonical ensemble and particularly the result that $\Xi = \exp(\lambda q)$.

5–19. Calculate q_{elec} for a hydrogen atom. The energy levels are given by

$$E_n = -\frac{2\pi^2 m e^4}{n^2 h^2} \qquad n = 1, 2, \ldots$$

and the degeneracy is $2n^2$. How is this seemingly paradoxical result explained? (See S. J. Strickler, *J. Chem. Educ.*, 43, p. 364, 1966.)

CHAPTER 6

IDEAL DIATOMIC GAS

In this chapter we shall treat an ideal gas composed of diatomic molecules. In addition to translational and electronic degrees of freedom, diatomic molecules possess vibrational and rotational degrees of freedom as well. The general procedure would be to set up the Schrödinger equation for two nuclei and *n*-electrons and to solve this equation for the set of eigenvalues of the diatomic molecule. Such a general exact approach is very difficult and has been done only for H_2. Fortunately, a series of very good approximations can be used to reduce this complicated two-nuclei, *n*-electron problem to a set of simpler problems. The simplest of these approximations is the rigid rotor-harmonic oscillator approximation. In Section 6–1 we shall discuss this approximation, and then in Sections 6–2 and 6–3 we discuss the vibrational and rotational partition functions within this approximation. Section 6–4 is a discussion of the symmetry of the wave functions of homonuclear diatomic molecules under the interchange of the two nuclei, and Section 6–5 is an application of these results to the rotational partition function of homonuclear diatomic molecules. This section contains a detailed discussion of ortho- and para-hydrogen. Section 6–6 summarizes the thermodynamic functions under the rigid rotor–harmonic oscillator approximation.

6–1 THE RIGID ROTOR–HARMONIC OSCILLATOR APPROXIMATION

We first make the Born-Oppenheimer approximation. The physical basis of the Born-Oppenheimer approximation is that the nuclei are much more massive than the electrons, and thus move slowly relative to the electrons. Therefore the electrons can be considered to move in a field produced by the nuclei fixed at some internuclear separation. Mathematically, the Schrödinger equation approximately separates into two simpler equations. One equation describes the motion of the electrons in the field of the fixed nuclei. We denote the eigenvalues of this equation by $u_j(r)$, where r is the internuclear separation. The other equation describes the motion of the nuclei in the electronic potential $u_j(r)$, that is, the potential set up by the electrons in the electronic state j. Each electronic state of the molecule creates its own characteristic internuclear potential. As in the atomic case, the first excited electronic state usually lies several electron volts above the ground state, and so only the ground electronic state potential is necessary. The calculation of $u_j(r)$ for even the ground state is a difficult n-electron calculation, and so semiempirical approximations such as the Morse potential are often used. Figure 6–1 illustrates a typical internuclear potential. Given $u_0(r)$, we treat the motion of the two nuclei in this potential.

Problem 1–8 shows that the motion of two masses in a spherically symmetric potential can be rigorously separated into two separate problems by the introduction of center of mass and relative coordinates. The center-of-mass motion is that of a freely

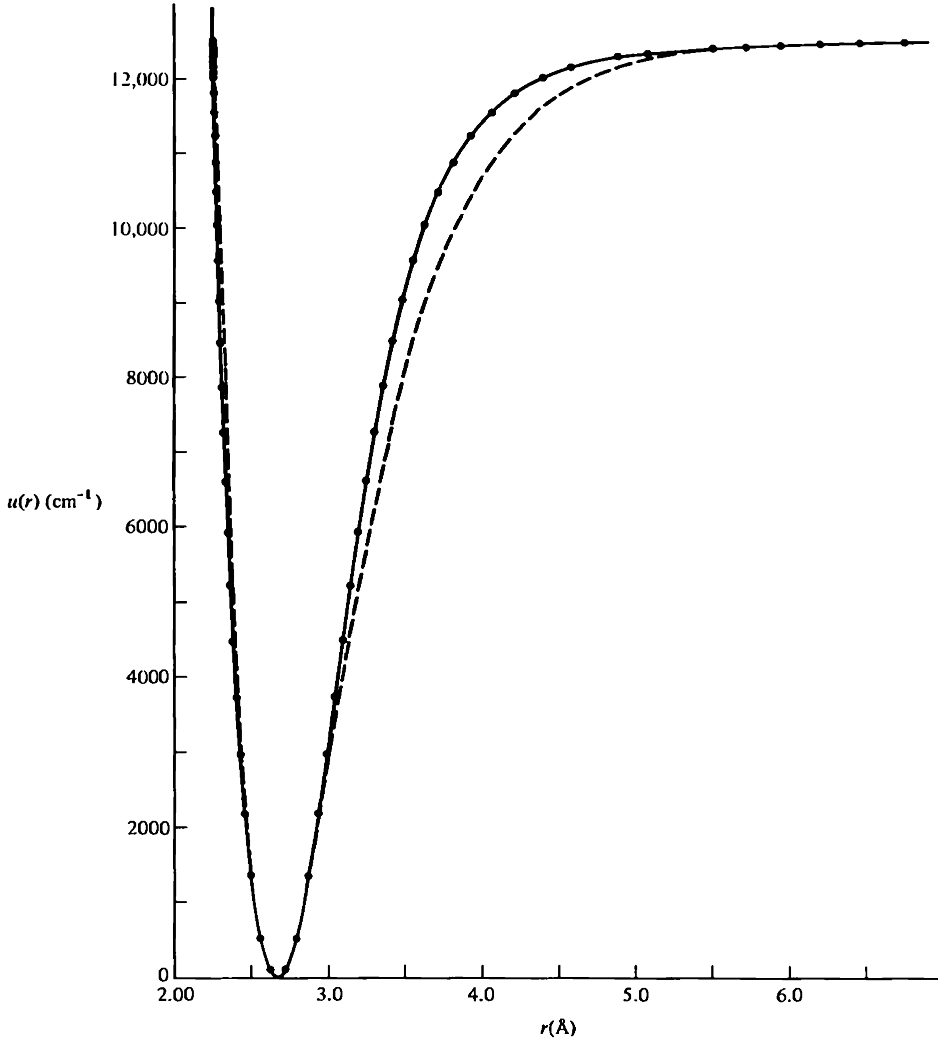

Figure 6–1. **The internuclear potential energy curve for the ground state of I_2 as computed from ultraviolet spectroscopy. The dashed curve is the Morse curve.** (From R. D. Verma, *J. Chem. Phys.*, 32, 738, 1960.)

translating point of mass $m_1 + m_2$ situated at the center of mass. The other problem is that of the *relative* motion of the two bodies, which can be interpreted as one body of reduced mass $\mu = m_1 m_2/(m_1 + m_2)$ moving about the other one fixed at the origin.

The Hamiltonian can be written as

$$\mathscr{H} = \mathscr{H}_{\text{trans}} + \mathscr{H}_{\text{int}}$$

with eigenvalues

$$\varepsilon = \varepsilon_{\text{trans}} + \varepsilon_{\text{int}}$$

The partition function of the diatomic molecule, therefore, is

$$q = q_{\text{trans}} q_{\text{int}}$$

where

$$q_{\text{trans}} = \left[\frac{2\pi(m_1 + m_2)kT}{h^2}\right]^{3/2} V \tag{6–1}$$

The density of translational states alone is so great that we can write

$$Q(N, V, T) = \frac{q_{\text{trans}}^N q_{\text{int}}^N}{N!} \tag{6–2}$$

Thus we must investigate q_{int} to complete our treatment of diatomic molecules.

The relative motion of the two nuclei in the potential $u(r)$ consists of rotary motion about the center of mass and relative vibratory motion of the two nuclei. It turns out that the amplitude of the vibratory motion is very small, and so it is a good approximation to consider the angular motion to be that of a rigid dumbbell of fixed internuclear distance r_e. In addition, the internuclear potential $u(r)$ can be expanded about r_e:

$$\begin{aligned} u(r) &= u(r_e) + (r - r_e)\left(\frac{du}{dr}\right)_{r=r_e} + \tfrac{1}{2}(r - r_e)^2\left(\frac{d^2u}{dr^2}\right)_{r=r_e} + \cdots \\ &= u(r_e) + \tfrac{1}{2}k(r - r_e)^2 + \cdots \end{aligned} \tag{6–3}$$

The linear term vanishes because (du/dr) is zero at the minimum of $u(r)$. The parameter k is a measure of the curvature of the potential at the minimum and is called the force constant. A large value of k implies a stiff bond; a small value implies a loose bond.

The approximation introduced in the previous paragraph is called the rigid rotor–harmonic oscillator approximation. It allows the Hamiltonian of the relative motion of the nuclei to be written as

$$\mathscr{H}_{\text{rot,vib}} = \mathscr{H}_{\text{rot}} + \mathscr{H}_{\text{vib}} \quad \text{(rigid rotor–harmonic oscillator approximation)} \tag{6–4}$$

and

$$\varepsilon_{\text{rot,vib}} = \varepsilon_{\text{rot}} + \varepsilon_{\text{vib}} \tag{6–5}$$

The partition function $q_{\text{rot,vib}}$, then, is

$$q_{\text{rot,vib}} = q_{\text{rot}} q_{\text{vib}} \quad \text{(rigid rotor-harmonic oscillator approximation)} \tag{6–6}$$

The energy eigenvalues and the degeneracy of a rigid rotor are given in Eq. (1–32)

$$\varepsilon_J = \frac{\hbar^2 J(J+1)}{2I} \qquad J = 0, 1, 2,$$
$$\omega_J = 2J + 1 \tag{6–7}$$

where I is the moment of inertia, μr_e^2, of the molecule. The energy and degeneracy of an harmonic oscillator are [*cf.* Eq. (1–31)]

$$\begin{aligned} \varepsilon_{\text{vib}} &= h\nu(n + \tfrac{1}{2}) \\ \omega_n &= 1 \quad \text{for all } n \end{aligned} \qquad n = 0, 1, 2, \ldots \tag{6–8}$$

where

$$\nu = \frac{1}{2\pi}\left(\frac{k}{\mu}\right)^{1/2} \tag{6–9}$$

Transitions from one rotational level to another can be induced by electromagnetic radiation. The selection rules for this are: (1) The molecule must have a permanent dipole moment, and (2) $\Delta J = \pm 1$. The frequency of radiation absorbed in the process of going from a level J to $J + 1$ is given by

$$\nu = \frac{\varepsilon_{J+1} - \varepsilon_J}{h} = \frac{h}{4\pi^2 I}(J+1) \qquad J = 0, 1, 2, \ldots \tag{6–10}$$

We thus expect absorption of radiation at frequencies given by multiples of $h/4\pi^2 I$ and should observe a set of equally spaced spectral lines, which for typical molecular values of μ and r_e^2 will be found in the microwave region. Experimentally one does see a series of almost equally spaced lines in the microwave spectra of linear molecules. The usual units of frequency in this region are wave numbers, or reciprocal wavelengths.

$$\bar{\omega}(\text{cm}^{-1}) = \frac{1}{\lambda} = \frac{\nu}{c} \tag{6–11}$$

Microwave spectroscopists define the rotational constant $\bar{B}$ by $h/8\pi^2 Ic$ (units of cm^{-1}), so that the energy of rigid rotor (in cm^{-1}) becomes

$$\bar{\varepsilon}_J = \bar{B}J(J+1) \tag{6–12}$$

Table 6–1 lists the values of $\bar{B}$ for several diatomic molecules.

For a molecule to change its vibrational state by absorbing radiation it must (1) change its dipole moment when vibrating and (2) obey the selection rule $\Delta n = \pm 1$. The frequency of absorption is, then, seen to be

$$\nu = \frac{\varepsilon_{n+1} - \varepsilon_n}{h} = \frac{1}{2\pi}\left(\frac{k}{\mu}\right)^{1/2} \tag{6–13}$$

Equation (6–13) predicts that the vibrational spectrum of a diatomic molecule will consist of just one line. This line occurs in the infrared, typically around 1000 cm^{-1}, giving force constants k of the order of 10^5 or 10^6 dynes/cm. (See Problem 6–5.) Table 6–1 gives the force constants of a number of diatomic molecules.

Table 6–1. **Molecular constants for several diatomic molecules***

molecule	electronic state	$\bar{\omega}$ (cm^{-1})	Θ_v (°K)	$\bar{B}$ (cm^{-1})	Θ_r (°K)	$k \times 10^{-5}$ (dynes/cm)	D_0 (kcal/mole)
H_2	$^1\Sigma_g^+$	4320	6215	59.3	85.3	5.5	103.2
D_2	$^1\Sigma_g^+$	3054	4394	29.9	42.7	5.5	104.6
Cl_2	$^1\Sigma_g^+$	561	808	0.244	0.351	3.2	57.1
Br_2	$^1\Sigma_g^+$	322	463	0.0809	0.116	2.4	45.4
I_2	$^1\Sigma_g^+$	214	308	0.0373	0.0537	1.7	35.6
O_2	$^3\Sigma_g^-$	1568	2256	1.437	2.07	11.6	118.0
N_2	$^1\Sigma_g^+$	2345	3374	2.001	2.88	22.6	225.1
CO	$^1\Sigma^+$	2157	3103	1.925	2.77	18.7	255.8
NO	$^2\Pi_{1/2}$	1890	2719	1.695	2.45	15.7	150.0
HCl	$^1\Sigma^+$	2938	4227	10.44	15.02	4.9	102.2
HBr	$^1\Sigma^+$	2640	3787	8.36	12.02	3.9	82.4
HI	$^1\Sigma^+$	2270	3266	6.46	9.06	3.0	70.5
Na_2	$^1\Sigma_g^+$	159	229	0.154	0.221	0.17	17.3
K_2	$^1\Sigma_g^+$	92.3	133	0.0561	0.081	0.10	15.8

* These parameters were obtained from a variety of sources and do not necessarily represent the most accurate values since they are obtained under the rigid rotor–harmonic oscillator approximation.

We furthermore assume that the electronic and nuclear degrees of freedom can be written separately, and thus we have

$$\mathscr{H} = \mathscr{H}_{\text{trans}} + \mathscr{H}_{\text{rot}} + \mathscr{H}_{\text{vib}} + \mathscr{H}_{\text{elec}} + \mathscr{H}_{\text{nucl}} \tag{6–14}$$

which implies that

$$\varepsilon = \varepsilon_{\text{trans}} + \varepsilon_{\text{rot}} + \varepsilon_{\text{vib}} + \varepsilon_{\text{elec}} + \varepsilon_{\text{nucl}} \tag{6–15}$$

and

$$q = q_{\text{trans}} q_{\text{rot}} q_{\text{vib}} q_{\text{elec}} q_{\text{nucl}} \tag{6–16}$$

The translational partition function is given by Eq. (6–1); the electronic partition function will be similar to Eq. (5–11) for a monatomic gas; and we shall usually adopt the convention that $q_{\text{nucl}} = 1$. Although Eq. (6–14) is not exact, it serves as a useful approximation. Within this approximation, the partition function of the gas itself is given by

$$Q(N, V, T) = \frac{(q_{\text{trans}} q_{\text{rot}} q_{\text{vib}} q_{\text{elec}} q_{\text{nucl}})^N}{N!} \tag{6–17}$$

We shall introduce several corrections to Eqs. (6–14) through (6–17) in Problems 6–23 and 6–24.

Only the vibrational and rotational contributions to the partition function are not known in Eq. (6–17), and we shall discuss these contributions in the next few sections. Before discussing these, however, we must choose a zero of energy for the rotational and vibrational states. The zero of rotational energy will usually be taken to be the $J = 0$ state. In the vibrational case we have two choices. One is to take the zero of vibrational energy to be that of the ground state, and the other is to take the zero to be the bottom of the internuclear potential well. In the first case, the energy of the ground vibrational state is zero, and in the second case it is $h\nu/2$. We shall choose the zero of

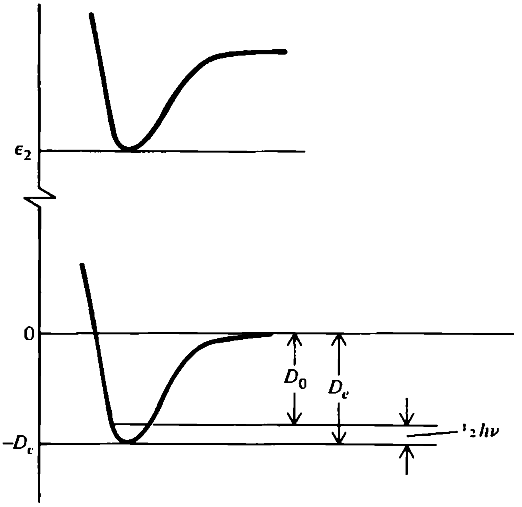

Figure 6–2. **The ground and first excited electronic states as a function of the internuclear separation r, illustrating the quantities D_0, D_e, and ε_2.**

vibrational energy to be the bottom of the internuclear potential well of the lowest electronic state. Lastly, we take the zero of the electronic energy to be the separated, electronically unexcited atoms at rest. If we denote the depth of the ground electronic state potential well by D_e, the energy of the ground electronic state is $-D_e$, and the electronic partition function is

$$q_{\text{elec}} = \omega_{e1} e^{D_e/kT} + \omega_{e2} e^{-\varepsilon_2/kT} + \cdots \tag{6–18}$$

where D_e and ε_2 are shown in Fig. 6–2. We also define a quantity D_0 by $D_e - \frac{1}{2}h\nu$. As Fig. 6–2 shows, D_0 is the energy difference between the lowest vibrational state and the dissociated molecule. The quantity D_0 can be measured spectroscopically (by predissociation spectra, for example) or calorimetrically from the heat of reaction at any one temperature and the heat capacities from 0°K to that temperature. The values of D_0 for a number of diatomic molecules are given in Table 6–1.

6–2 THE VIBRATIONAL PARTITION FUNCTION

Since we are measuring the vibrational energy levels relative to the bottom of the internuclear potential well, we have

$$\varepsilon_n = (n + \tfrac{1}{2})h\nu \qquad n = 0, 1, 2, \ldots \tag{6–19}$$

with $\nu = (k/\mu)^{1/2}/2\pi$, where k is the force constant of the molecule, and μ is its reduced mass [*cf.* Eq. (6–9)]. The vibrational partition function q_{vib}, then, becomes

$$\begin{aligned} q_{\text{vib}}(T) &= \sum_n e^{-\beta\varepsilon_n} \\ &= e^{-\beta h\nu/2} \sum_{n=0}^{\infty} e^{-\beta h\nu n} \\ &= \frac{e^{-\beta h\nu/2}}{1 - e^{-\beta h\nu}} \end{aligned} \tag{6–20}$$

where we have recognized the summation above as a geometric series. This is one of the rare cases in which q can be summed directly without having to approximate it by an integral, as we did in the translational case in Chapter 5 and shall do shortly in the rotational case. The quantity $\beta h\nu$ is ordinarily larger than 1, but if the temperature is high enough, $\beta h\nu \ll 1$, and we can replace the sum in Eq. (6–20) by an integral to get

$$q_{\text{vib}}(T) = e^{-\beta h\nu/2} \int_0^\infty e^{-\beta h\nu n}\, dn = \frac{kT}{h\nu} \qquad (kT \gg h\nu) \tag{6–21}$$

which we see is what results from Eq. (6–20) if $\beta h\nu \ll 1$. Although we shall rarely use this approximation since we have $q_{\text{vib}}(T)$ exactly, it will be interesting to compare this limit to some others which we shall derive later on. From $q_{\text{vib}}(T)$ we can calculate the vibrational contribution to the thermodynamic energy

$$E_v = NkT^2 \frac{d \ln q_v}{dT} = Nk\left(\frac{\Theta_v}{2} + \frac{\Theta_v}{e^{\Theta_v/T} - 1}\right) \tag{6–22}$$

where $\Theta_v \equiv h\nu/k$ and is called the vibrational temperature. Table 6–1 gives Θ_v for a number of diatomic molecules. The vibrational contribution to the heat capacity is

$$\left(\frac{\partial E_v}{\partial T}\right)_N = Nk\left(\frac{\Theta_v}{T}\right)^2 \frac{e^{\Theta_v/T}}{(e^{\Theta_v/T} - 1)^2} \tag{6–23}$$

Notice that as $T \to \infty$, $E_v \to NkT$ and $C_v \to Nk$, a result given in many physical chemistry courses and one whose significance we shall understand more fully when we discuss equipartition of energy.

Figure 6–3 shows the vibrational contribution of an ideal diatomic gas to the molar heat capacity as a function of temperature.

An interesting quantity to calculate is the fraction of molecules in excited vibrational states. The fraction of molecules in the vibrational state designated by n is

$$f_n = \frac{e^{-\beta h\nu(n+1/2)}}{q_{\text{vib}}} \tag{6–24}$$

This equation is shown in Fig. 6–4 for Br_2 at 300°K. Notice that most molecules are in the ground vibrational state and that the population of the higher vibrational states

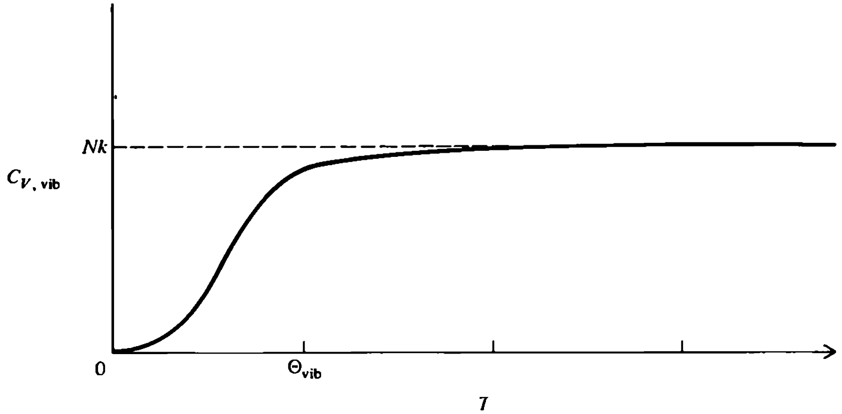

Figure 6–3. **The vibrational contribution of an ideal diatomic gas to the molar heat capacity as a function of temperature. Room temperature is typically $O(0.1\Theta_v)$(*Cf.* Table 6–1).**

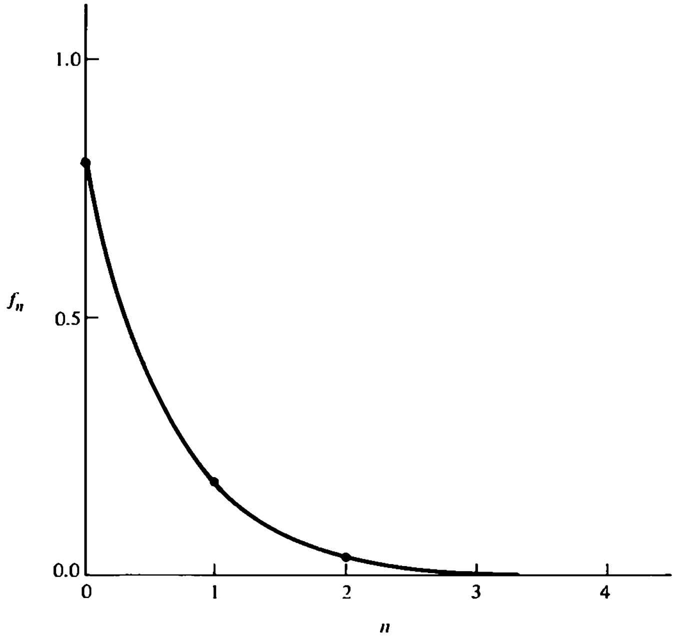

Figure 6–4. **The population of the vibrational levels of Br_2 at 300°K.**

decreases exponentially. Bromine has a force constant smaller than most molecules, however (*cf.* Table 6–1), and so the population of excited vibrational levels of Br_2 is greater than most other molecules. Table 6–2 gives the fraction of molecules in all excited states for a number of molecules. This fraction is given by

$$f_{n>0} = \sum_{n=1}^{\infty} \frac{e^{-\beta h\nu(n+1/2)}}{q_{\text{vib}}} = 1 - f_0 = e^{-\beta h\nu} = e^{-\Theta_v/T} \tag{6–25}$$

Table 6–2. **The fraction of molecules in excited vibrational states at 300°K and 1000°K**

gas	Θ_v, °K	$e^{-\Theta_v/T}$ 300°K	 1000°K
H_2	6215	1.04×10^{-9}	2.03×10^{-3}
HCl	4227	1.02×10^{-6}	1.59×10^{-2}
N_2	3374	1.51×10^{-5}	3.55×10^{-2}
CO	3100	3.71×10^{-5}	4.65×10^{-2}
Cl_2	810	6.72×10^{-2}	4.45×10^{-1}
I_2	310	3.56×10^{-1}	7.33×10^{-1}

6–3 THE ROTATIONAL PARTITION FUNCTION OF A HETERONUCLEAR DIATOMIC MOLECULE

For heteronuclear diatomic molecules, the calculation of the rotational partition function is straightforward. The rotational partition function is given by

$$q_{\text{rot}}(T) = \sum_{J=0}^{\infty} (2J+1)e^{-\beta BJ(J+1)} \tag{6–26}$$

In the nomenclature of Chapter 2, Eq. (6–26) is a summation over levels rather than over states.

The ratio $\bar{B}/k$ is denoted by Θ_r and is called the characteristic temperature of rotation. This is given in Table 6–1 for a number of molecules. Unlike the vibrational case, this sum cannot be written in closed form. However, because Θ_r/T is quite small at ordinary temperatures for most molecules, we can approximate this sum by an integral. (It is really $\Delta\varepsilon/kT = 2\Theta_r(J+1)/T$ that must be small compared to one, and this of course cannot be true as J increases. However, by the time J is large enough to contradict this, the terms are so small that it makes no difference.)

At high enough temperatures, then,

$$q_{\mathrm{rot}}(T) = \int_0^\infty (2J+1)e^{-\Theta_r J(J+1)/T}\,dJ \tag{6–27}$$

$$= \int_0^\infty e^{-\Theta_r J(J+1)/T}\,d\{J(J+1)\} = \frac{T}{\Theta_r} \tag{6–28}$$

$$= \frac{8\pi^2 IkT}{h^2} \qquad \Theta_r \ll T \tag{6–29}$$

This result improves as the temperature increases and is called the high-temperature limit. For low temperatures or for molecules with large values of Θ_r, say HD with $\Theta_r = 42.7°\text{K}$, one can use the sum directly. For example,

$$q_{\mathrm{rot}}(T) = 1 + 3e^{-2\Theta_r/T} + 5e^{-6\Theta_r/T} + 7e^{-12\Theta_r/T} \tag{6–30}$$

is sufficient to give the sum to within 0.1 percent for $\Theta_r > 0.7T$. For Θ_r less than $0.7T$ but not small enough for the integral to give a good approximation, we need some intermediate approximation.

The replacement of a sum by an integral can be viewed as the first of a sequence of approximations. The full scheme is a standard result of the field of the calculus of finite differences and is called the Euler-MacLaurin summation formula. It states that if $f(n)$ is a function defined on the integers and continuous in between, then

$$\sum_{n=a}^{b} f(n) = \int_a^b f(n)\,dn + \tfrac{1}{2}\{f(b) + f(a)\}$$
$$+ \sum_{j=1}^{\infty} (-)^j \frac{B_j}{(2j)!}\{f^{(2j-1)}(a) - f^{(2j-1)}(b)\} \tag{6–31}$$

where $f^{(k)}(a)$ is the kth derivative of f evaluated at a. The B_j's are the Bernoulli numbers, $B_1 = \frac{1}{6}$, $B_2 = \frac{1}{30}$, $B_3 = \frac{1}{42}$, Before applying this to $q_{\mathrm{rot}}(T)$, let us apply it first to a case we can do exactly. Consider the sum [*cf.* Eq. (6–20)]

$$\sum_{j=0}^{\infty} e^{-\alpha j} = \frac{1}{1 - e^{-\alpha}} \tag{6–32}$$

Applying the Euler-MacLaurin summation formula, we get

$$\sum_{j=0}^{\infty} e^{-\alpha j} = \frac{1}{\alpha} + \frac{1}{2} + \frac{\alpha}{12} - \frac{\alpha^3}{720} + \cdots \tag{6–33}$$

The expansion of $(1 - e^{-\alpha})^{-1}$ is

$$\frac{1}{1 - e^{-\alpha}} = \frac{1}{\alpha - \dfrac{\alpha^2}{2} + \dfrac{\alpha^3}{6} - \cdots} = \frac{1}{\alpha} + \frac{1}{2} + \frac{\alpha}{12} - \frac{\alpha^3}{720} + \cdots \tag{6–34}$$

We see that these two expansions are the same. If α is large, we can use the first few terms of Eq. (6–33); otherwise, we use the Euler-MacLaurin expansion in α.

Applying this formula to $q_{\text{rot}}(T)$ gives (see Problem 6–9):

$$q_{\text{rot}}(T) = \frac{T}{\Theta_r}\left\{1 + \frac{1}{3}\left(\frac{\Theta_r}{T}\right) + \frac{1}{15}\left(\frac{\Theta_r}{T}\right)^2 + \frac{4}{315}\left(\frac{\Theta_r}{T}\right)^3 + \cdots\right\} \tag{6–35}$$

which is good to within one percent for $\Theta_r < T$. For simplicity we shall use only the high-temperature limit in what we do here since Θ_r is $\ll T$ for most molecules at room temperature (*cf.* Table 6–1).

The rotational contribution to the thermodynamic energy is

$$E_{\text{rot}} = NkT^2\left(\frac{\partial \ln q_{\text{rot}}}{\partial T}\right) = NkT + \cdots \tag{6–36}$$

and the contribution to the heat capacity is

$$C_{V,\text{rot}} = Nk + \cdots \tag{6–37}$$

The fraction of molecules in the Jth rotational state is

$$\frac{N_J}{N} = \frac{(2J+1)e^{-\Theta_r J(J+1)/T}}{q_{\text{rot}}(T)} \tag{6–38}$$

Figure 6–5 shows this fraction for HCl at 300°K. Contrary to the vibrational case, most molecules are in excited rotational levels at ordinary temperatures. We can find the maximum of this curve by differentiating Eq. (6–38) with respect to J to get

$$J_{\max} = \left(\frac{kT}{2\bar{B}}\right)^{1/2} - \frac{1}{2} \approx \left(\frac{T}{2\Theta_r}\right)^{1/2} = \left(\frac{kT}{2\bar{B}}\right)^{1/2}$$

We see then that $J_{\max}$ increases with T and is inversely related to $\bar{B}$, and so increases with the moment of inertia of the molecule since $\bar{B} \propto 1\ I$.

The next two sections deal with the rotational partition function of homonuclear diatomic molecules. The wave function of a homonuclear diatomic molecule must possess a certain symmetry with respect to the interchange of the two identical nuclei in the molecule. In particular, if the two nuclei have integral spins, the wave function must be symmetric with respect to an interchange; if the nuclei have half odd integer

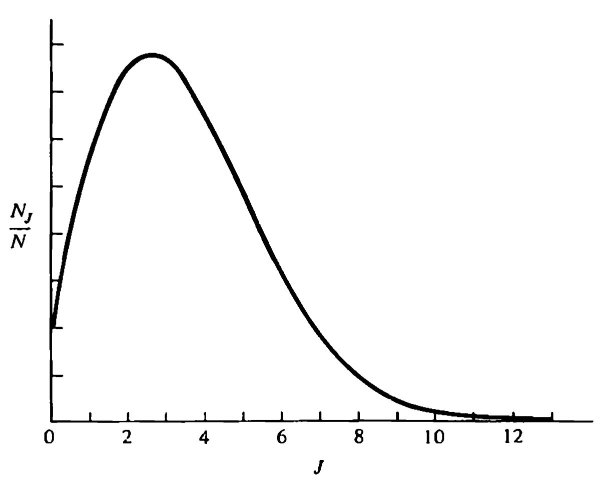

Figure 6–5. **The population of the rotational levels of hydrogen chloride at 300°K.**

spin, the wave function must be antisymmetric. This symmetry requirement has a profound effect on the rotational energy levels of a homonuclear diatomic molecule, which can be understood only by understanding the general symmetry properties of a homonuclear diatomic molecule. This is discussed in the next section. Then in Section 6–5 we apply these results to the rotational partition function. We shall see there that at low temperatures, the symmetry properties have an important effect on the thermodynamic properties of certain molecules, and in particular we shall discuss ortho- and para-hydrogen.

The discussion of the symmetry requirement is somewhat involved, however, and so we present here a summary of the high-temperature limit, which for most systems is completely adequate. At temperatures such that the summation in Eq. (6–26) can be approximated by an integral or even the Euler-MacLaurin expansion, the result for a homonuclear diatomic molecule is

$$q_{\text{rot}} = \frac{T}{2\Theta_r}\left\{1 + \frac{\Theta_r}{3T} + \frac{1}{15}\left(\frac{\Theta_r}{T}\right)^2 + \cdots\right\} \tag{6–39}$$

Note that this equation is the same as Eq. (6–35) except for the factor of 2 in the denominator. This factor is due to the symmetry of the homonuclear diatomic molecule and, in particular, is due to the fact that there are two indistinguishable orientations of a homonuclear diatomic molecule. There is a two-fold axis of symmetry perpendicular to the internuclear axis.

Equations (6–35) and (6–39) can be written as one equation by introducing a factor σ into the denominator of Eq. (6–35). If $\sigma = 1$, we have Eq. (6–35), and if $\sigma = 2$, we have Eq. (6–39). The factor σ is called the *symmetry number* of the molecule and represents the number of indistinguishable orientations that the molecule can have. The exact origin of σ can only be completely understood from the arguments presented in Sections 6–4 and 6–5, but on first reading it is possible to accept the factor of σ and proceed directly to Eq. (6–47) from here.

6–4 THE SYMMETRY REQUIREMENT OF THE TOTAL WAVE FUNCTION OF A HOMONUCLEAR DIATOMIC MOLECULE

The calculation of the rotational partition function is not quite so straightforward for homonuclear diatomic molecules. The total wave function of the molecule, that is, the electronic, vibrational, rotational, translational, and nuclear wave function, must be either symmetric or antisymmetric under the interchange of the two identical nuclei. It must be symmetric if the nuclei have integral spins (bosons), or antisymmetric if they have half-integral spins (fermions). This symmetry requirement has profound consequences on the thermodynamic properties of homonuclear diatomic molecules at low temperatures. We shall discuss the interchange of the two identical nuclei of a homonuclear diatomic molecule in this section, and then apply the results to the calculation of q_{rot} in the next section.

It is convenient to imagine this interchange as a result of (1) an inversion of all the particles, electrons and nuclei, through the origin, and then (2) an inversion of just the electrons back through the origin. This two-step process is equivalent to an exchange of the nuclei. Let us write ψ_{total} *exclusive* of the nuclear part as

$$\psi'_{\text{total}} = \psi_{\text{trans}}\psi_{\text{rot}}\psi_{\text{vib}}\psi_{\text{elec}}$$

where the prime on ψ_{total} indicates that we are ignoring the nuclear contribution for now. The translational wave function depends only upon the coordinates of the center of mass of the molecule, and so this factor is not affected by inversion. Furthermore, ψ_{vib} depends only upon the magnitude of $(r - r_e)$, and so this part of the total wave function is unaffected by any inversion operation. Therefore, we concentrate on ψ_{elec} and ψ_{rot}.

The property of ψ_{elec} under the inversions in both Steps (1) and (2) above depends upon the symmetry of the ground electronic state of the molecule. The ground electronic state of most molecules is symmetric under both of these operations. Such a state is designated by the term symbol $\sum_g^+$. Thus it is ψ_{rot} that controls the symmetry of ψ_{total}.

Only Step (1) above, the inversion of both electrons and nuclei through the origin, affects ψ_{rot}. The effect of this inversion is to change the coordinates (r, θ, ϕ) that describe the orientation of the diatomic molecule into $(r, \pi - \theta, \phi + \pi)$. One can see this either analytically from the eigenfunctions themselves or pictorially from the rotational wave functions shown in Fig. 6–6. Notice that the rigid rotor wave functions are the same functions as the angular functions of the hydrogen atom.

The net result then, when the ground electronic state is symmetric, that is, $\sum_g^+$ is that ψ_{total} remains unchanged for even J and changes sign for odd J. This result applies to the total wave function, exclusive of nuclear spin.

Now consider a molecule such as H_2, whose nuclei have a spin of $\frac{1}{2}$. Just as in the case of the two electrons in the helium atom, the two nuclei of spin $\frac{1}{2}$ have three symmetric spin functions $\alpha\alpha$, $\beta\beta$, and $2^{-1/2}(\alpha\beta + \beta\alpha)$, and one antisymmetric spin function $2^{-1/2}(\alpha\beta - \beta\alpha)$. Since nuclei with spin $\frac{1}{2}$ act as fermions, the total wave function must be antisymmetric in the exchange of these two nuclei. Now states with both even and odd values of J can be brought to the required antisymmetry by coupling them with the right spin functions. Since three symmetric nuclear spin functions can be combined with the odd J levels to achieve the correct overall antisymmetry for ${}^1\sum_g^+$ electronic states, we see that the odd J levels have a statistical weight of 3, compared to a weight of 1 for even J levels. This leads to the existence of ortho- (parallel nuclear spins) states and para- (opposed nuclear spins) states in H_2. This weighting of the rotational states will be seen shortly to have a profound effect on the low-temperature thermodynamics of H_2.

More generally, for nuclei of spin I, there are $2I + 1$ spin states for each nucleus. Let the eigenfunctions of these spin states be denoted by $\alpha_1, \alpha_2, \ldots, \alpha_{2I+1}$. There are $(2I + 1)^2$ nuclear wave functions to include in ψ_{total}. (In the case of H_2, $I = \frac{1}{2}$, there are two spin states α and β, and there are four nuclear spin functions, three of which are symmetric and one of which is antisymmetric.) The antisymmetric nuclear spin functions are of the form $\alpha_i(1)\alpha_j(2) - \alpha_i(2)\alpha_j(1)$, $1 \leq i, j \leq 2I + 1$. There are $(2I + 1)(2I)/2$ such combinations, and so this is the number of antisymmetric nuclear spin functions. (For H_2, we find that there is only one antisymmetric choice in agreement with the above paragraph.) All the remaining $(2I + 1)^2$ total nuclear spin functions are symmetric, and so their number is $(2I + 1)^2 - I(2I + 1) = (I + 1)(2I + 1)$. Thus we can write the following summary for $\sum_g^+$ states;

integral spin

$I(2I + 1)$ antisymmetric nuclear spin functions couple with odd J

$(I + 1)(2I + 1)$ symmetric nuclear spin functions couple with even J

half-integral spin

$I(2I + 1)$ antisymmetric nuclear spin functions couple with even J

$(I + 1)(2I + 1)$ symmetric nuclear spin functions couple with odd J

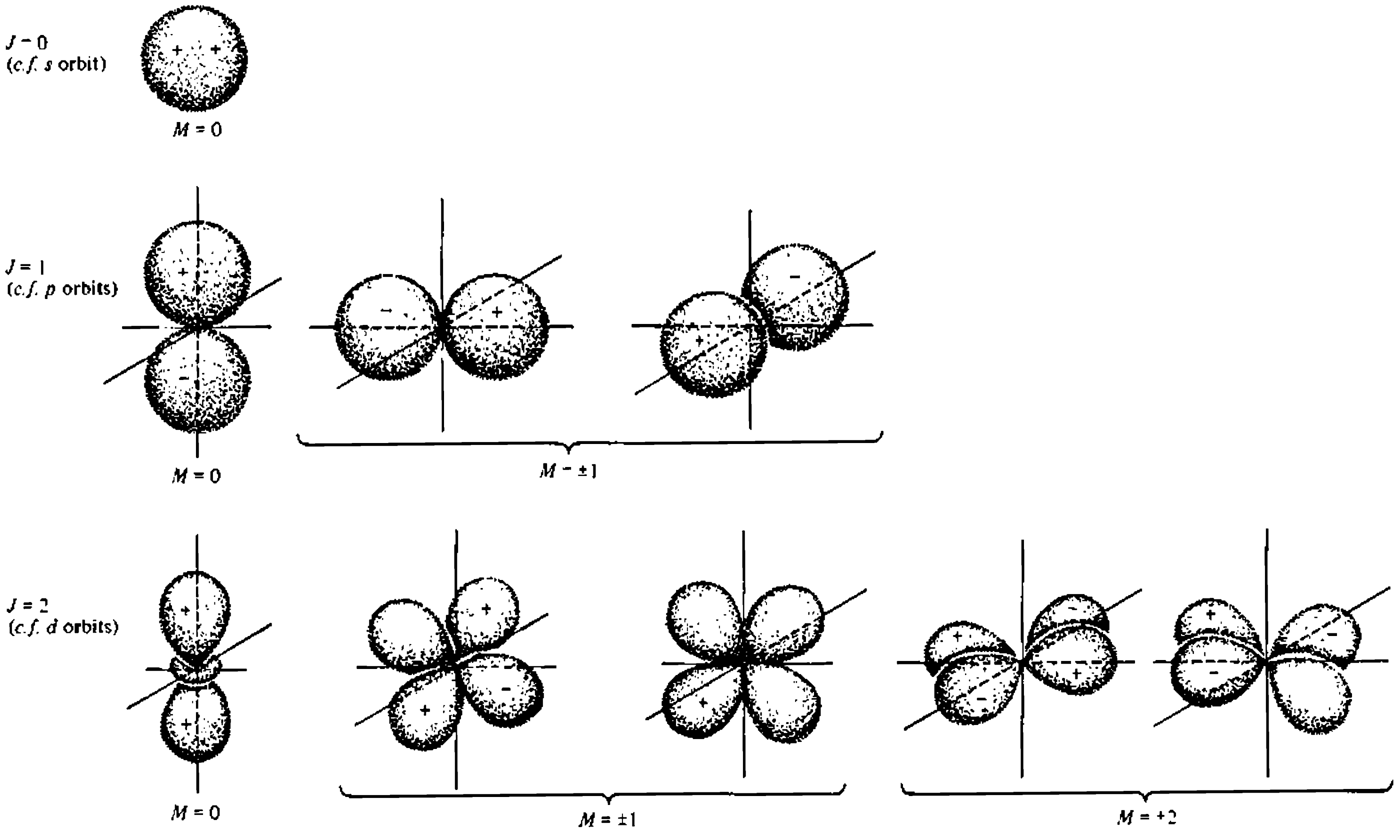

Figure 6-6. **Diagrammatic representations of the rotational wave functions for $J = 0, 1$, and 2.**

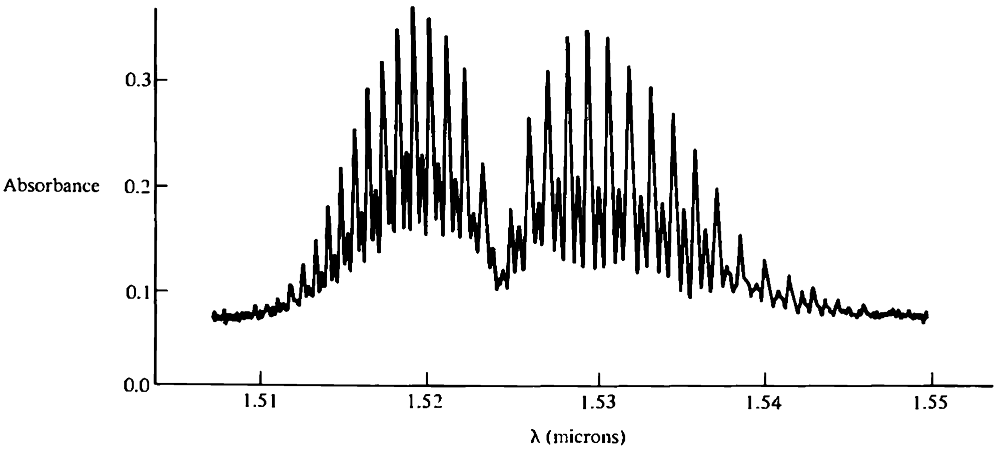

Figure 6–7. **The vibration–rotation spectrum of acetylene. This represents one vibrational line. The alternation in the intensity of the lines is due to the statistical weights of the rotational levels. (From L. W. Richards, *J. Chem. Ed.*, 43, p. 645, 1966.)**

These combinations of nuclear and rotational wave functions produce the correct symmetry required of the total wave function under interchange of identical nuclei. Remember that all of these conclusions are for Σ_g^+ electronic states, the most commonly occurring ground state. (See Problem 6–26 for a discussion of O_2.)

Even though we have considered only diatomic molecules here, the results of this section apply also to linear polyatomic molecules such as CO_2, H_2C_2. For example, the molecules $HC^{12}C^{12}H$ and $DC^{12}C^{12}D$ have their rotational states weighted in a similar way as H_2 and D_2. Figure 6–7 shows the vibration-rotation spectrum of H_2C_2. The alternation in the intensity of these rotational lines due to the statistical weights is very apparent.

6–5 THE ROTATIONAL PARTITION FUNCTION OF A HOMONUCLEAR DIATOMIC MOLECULE

The results of the previous section show that for homonuclear diatomic molecules with nuclei having integral spin, rotational levels with odd values of J must be coupled with the $I(2I + 1)$ antisymmetric nuclear spin functions, and that rotational levels with even values of J must be coupled with the $(I + 1)(2I + 1)$ symmetric nuclear spin functions. Thus we write

$$\begin{aligned} q_{\text{rot, nucl}}(T) = (I + 1)(2I + 1) \sum_{J \text{ even}} (2J + 1)e^{-\Theta_r J(J+1)/T} \\ + I(2I + 1) \sum_{J \text{ odd}} (2J + 1)e^{-\Theta_r J(J+1)/T} \end{aligned} \tag{6–40}$$

Likewise, for molecules with nuclei with half-integer spins, we have

$$\begin{aligned} q_{\text{rot, nucl}}(T) = I(2I + 1) \sum_{J \text{ even}} (2J + 1)e^{-\Theta_r J(J+1)/T} \\ + (I + 1)(2I + 1) \sum_{J \text{ odd}} (2J + 1)e^{-\Theta_r J(J+1)/T} \end{aligned} \tag{6–41}$$

Notice that in this case the combined rotational and nuclear partition function *does not factor* into $q_{\text{rot}} q_{\text{nucl}}$. This is a situation in which we cannot ignore q_{nucl}. For most molecules at ordinary temperatures, $\Theta_r \ll T$, and we can replace the sum by an integral. We see then that

$$\sum_{J\,\text{even}} \approx \sum_{J\,\text{odd}} \approx \frac{1}{2}\sum_{\text{all}\,J} \approx \frac{1}{2}\int_0^\infty (2J+1)e^{-\Theta_r J(J+1)/T}\,dJ = \frac{T}{2\Theta_r} \tag{6–42}$$

and so both Eqs. (6–40) and (6–41) become

$$q_{\text{rot, nucl}}(T) = \frac{(2I+1)^2 T}{2\Theta_r} \tag{6–43}$$

which can be written as $q_{\text{rot}}(T)q_{\text{nucl}}$ where

$$q_{\text{rot}}(T) = \frac{T}{2\Theta_r} \quad \text{and} \quad q_{\text{nucl}} = (2I+1)^2 \tag{6–44}$$

For Eq. (6–42) to be valid, Θ_r/T must be less than about 0.20.

This result is to be compared to the result for a heteronuclear diatomic molecule, namely, $q_{\text{rot}}(T) = T/\Theta_r$. The factor of 2 that appears above in the high-temperature limit takes into account that the molecule is homonuclear, and so its rotational partition function is given by Eq. (6–40) or (6–41) instead of (6–26). This factor of 2 is called the symmetry number and is denoted by σ. It legitimately appears only when Θ_r is less than approximately $0.2T$, since only then can we use Eq. (6–42). Understanding the origin of this fact then, we can write

$$q_{\text{rot}}(T) \approx \frac{8\pi^2 IkT}{\sigma h^2} \approx \frac{1}{\sigma}\sum_{J=0}^{\infty}(2J+1)e^{-\Theta_r J(J+1)/T} \qquad \Theta_r \ll T \tag{6–45}$$

where $\sigma = 1$ for heteronuclear molecules, and $\sigma = 2$ for homonuclear diatomic molecules. Remember that this is applicable only to the high-temperature limit or its Euler-MacLaurin correction. A similar factor will appear for polyatomic molecules also.

There are some interesting systems in which Θ_r/T is not small. Hydrogen is one of the most important such cases. Each nucleus in H_2 has nuclear spin $\frac{1}{2}$, and so

$$q_{\text{rot, nucl}} = \sum_{J\,\text{even}} (2J+1)e^{-\Theta_r J(J+1)/T} + 3\sum_{J\,\text{odd}}(2J+1)e^{-\Theta_r J(J+1)/T} \tag{6–46}$$

The hydrogen with only even rotational levels allowed (antisymmetric nuclear spin function or "opposite" nuclear spins) is called para-hydrogen; that with only odd rotational levels allowed (symmetric nuclear spin function or "parallel" nuclear spins) is called ortho-hydrogen. The ratio of the number of ortho-H_2 molecules to the number of para-H_2 molecules is

$$\frac{N_{\text{ortho}}}{N_{\text{para}}} = \frac{3\sum\limits_{J\,\text{odd}}(2J+1)e^{-\Theta_r J(J+1)/T}}{\sum\limits_{J\,\text{even}}(2J+1)e^{-\Theta_r J(J+1)/T}}$$

Figure 6–8 shows the percentage of p–H_2 versus temperature in an equilibrium mixture of ortho- and para-hydrogen. Note that the system is all para- at 0°K and 25 percent para- at high temperatures.

Figure 6–9 illustrates an interesting situation that occurs with low-temperature heat capacity measurements on H_2. Equation (6–46) can be used to calculate the heat capacity of H_2, and this is plotted in Fig. 6–9, along with the experimental results. It can be seen that the two curves are in great disagreement. These calculations and

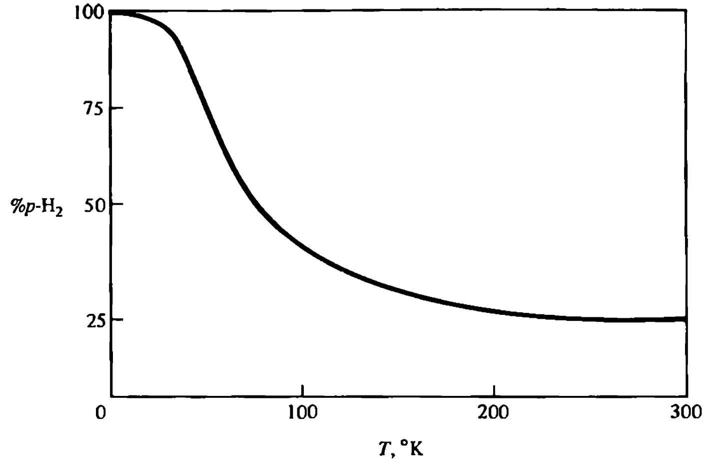

Figure 6–8. **The percentage of para-hydrogen in an equilibrium mixture as a function of temperature.**

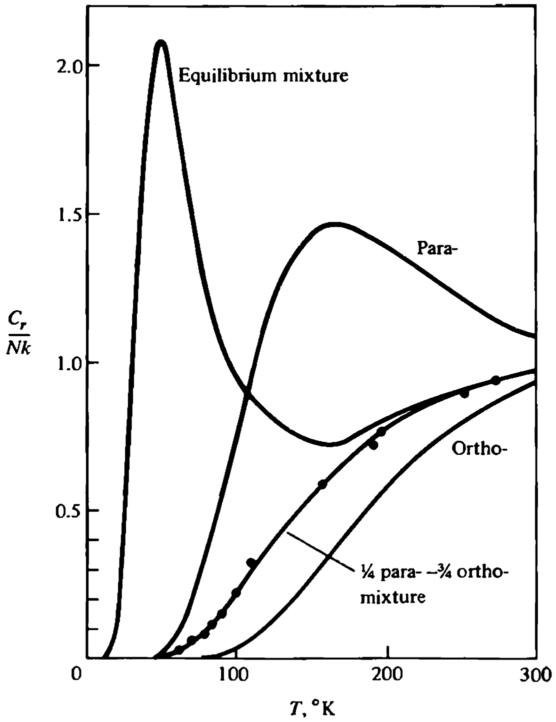

Figure 6–9. **The rotational and nuclear contribution to the molar heat capacity for ortho-hydrogen, para-hydrogen, an equilibrium mixture of ortho- and para-hydrogen, a metastable 75 percent ortho- and 25 percent para- mixture, and the experimental data. (From K. F. Bonhoeffer and P. Harteck, *Z. Physikal. Chem.*, 4B, p. 113, 1929.)**

measurements were made at a time when quantum mechanics was being developed, and was not accepted by all scientists. For a while, the disagreement illustrated in Fig. 6–9 was a blow to the proponents of the new quantum mechanics. It was Dennison* who finally realized that the conversion between ortho- and para-hydrogen is extremely slow in the absence of a catalyst, and so when hydrogen is prepared in the laboratory at room temperature and then cooled down for the low-temperature heat capacity measurements, the room-temperature composition persists instead of the equilibrium composition. Thus the experimental data illustrated in Fig. 6–9 are not for an equilibrium system of ortho- and para-hydrogen, but for a metastable system whose ortho-para composition is that of equilibrium room-temperature hydrogen, namely, 75 percent ortho- and 25 percent para-. If one calculates the heat capacity of such a system, according to

$$C_V = \tfrac{3}{4}C_V(\text{ortho-}) + \tfrac{1}{4}C_V(\text{para-})$$

where C_V(ortho-) is obtained from just the second term of Eq. (6–46), and C_V(para-) is obtained from the first term of Eq. (6–46), one obtains excellent agreement with the experimental curve. A clever confirmation of this explanation was shortly after obtained by Bonhoeffer and Harteck,† who performed heat capacity measurements on hydrogen in the presence of activated charcoal, a catalyst for the ortho–para conversion. This produces an equilibrium system at each temperature. The experimental data are in excellent agreement with the equilibrium calculation in Fig. 6–9.

The explanation of the heat capacity of H_2 was one of the great triumphs of post-quantum mechanical statistical mechanics. You should be able to go through a similar argument for D_2, sketching the equilibrium heat capacity, the pure ortho- and para-heat capacity, and finally what you should expect the experimental curve to be for D_2 prepared at room temperature and at some other temperature, say 20°K. (See Problem 6–17.)

In principle, such nuclear spin effects should be observable in other homonuclear molecules, but a glance at Table 6–1 shows that the characteristic rotational temperatures for all the other molecules are so small that these molecules reach the "high-temperature limit" while still in the solid state. Hydrogen is somewhat unusual in that its rotational constant is so much greater than its boiling point.

For most cases then, we can use Eq. (6–45) which, when we use the Euler-MacLaurin expansion, becomes

$$q_{\text{rot}}(T) = \frac{T}{\sigma\Theta_r}\left\{1 + \frac{\Theta_r}{3T} + \frac{1}{15}\left(\frac{\Theta_r}{T}\right)^2 + \frac{4}{315}\left(\frac{\Theta_r}{T}\right)^3 + \cdots\right\} \tag{6–47}$$

Usually only the first term of this is necessary. Some of the thermodynamic functions are

$$E_{\text{rot}} = NkT\left\{1 - \frac{\Theta_r}{3T} - \frac{1}{45}\left(\frac{\Theta_r}{T}\right)^2 + \cdots\right\} \tag{6–48}$$

$$C_{\text{rot}} = Nk\left\{1 + \frac{1}{45}\left(\frac{\Theta_r}{T}\right)^2 + \cdots\right\} \tag{6–49}$$

---

* D. M. Dennison, *Proc. Roy. Soc.* **A115**, 483, 1927.

† K. F. Bonhoeffer and P. Harteck, *Z. Phys. Chem.*, **4B**, 113, 1926.

$$S_{\text{rot}} = Nk\left\{1 - \ln\left(\frac{\sigma\Theta_r}{T}\right) - \frac{1}{90}\left(\frac{\Theta_r}{T}\right)^2 + \cdots\right\} \tag{6–50}$$

where all of these formulas are valid in the same region, in which σ itself is a meaningful concept, that is, $\Theta_r < 0.2T$. The terms in Θ_r/T and its higher powers are usually not necessary. Note that Eq. (6–47) is identical to Eq. (6–35) except for the occurrence of the symmetry number in Eq. (6–47).

6-6 THERMODYNAMIC FUNCTIONS

Having studied each contribution to the total partition function q in Eq. (6–17), we can write in the harmonic oscillator–rigid rotor approximation

$$q(V, T) = \left(\frac{2\pi mkT}{h^2}\right)^{3/2} V \frac{8\pi^2 IkT}{\sigma h^2} e^{-\beta h\nu/2}(1 - e^{-\beta h\nu})^{-1}\omega_{e_1} e^{D_e/kT} \tag{6–51}$$

Remember that this requires that $\Theta_r \ll T$, that only the ground electronic state is important, and that the zero of energy is taken to be the separated states at rest in their ground electronic states. Note that only q_{trans} is a function of V, and this is of the form $f(T)V$ which, we have seen before, is responsible for the ideal gas equation of state. The thermodynamic functions associated with Eq. (6–51) are

$$\frac{E}{NkT} = \frac{5}{2} + \frac{h\nu}{2kT} + \frac{h\nu/kT}{e^{h\nu/kT} - 1} - \frac{D_e}{kT} \tag{6–52}$$

$$\frac{C_V}{Nk} = \frac{5}{2} + \left(\frac{h\nu}{kT}\right)^2 \frac{e^{h\nu/kT}}{(e^{h\nu/kT} - 1)^2} \tag{6–53}$$

$$\frac{S}{Nk} = \ln\left[\frac{2\pi(m_1 + m_2)kT}{h^2}\right]^{3/2} \frac{Ve^{5/2}}{N} + \ln\frac{8\pi^2 IkTe}{\sigma h^2} + \frac{h\nu/kT}{e^{h\nu/kT} - 1} - \ln(1 - e^{-h\nu/kT}) + \ln\omega_{e1} \tag{6–54}$$

$$pV = NkT \tag{6–55}$$

$$\frac{\mu^0(T)}{kT} = -\ln\left[\frac{2\pi(m_1 + m_2)kT}{h^2}\right]^{3/2} kT - \ln\frac{8\pi^2 IkT}{\sigma h^2} + \frac{h\nu}{2kT} + \ln(1 - e^{-h\nu/kT}) - \frac{D_e}{kT} - \ln\omega_{e1} \tag{6–56}$$

Table 6–1 contains the characteristic rotational temperatures, the characteristic vibrational temperatures, and $D_0 = D_e - \frac{1}{2}h\nu$ for a number of diatomic molecules.

Table 6-3 presents a comparison of Eq. (6–54) with experimental data. It can be seen that the agreement is quite good and is typical of that found for the other thermodynamic functions. It is possible to improve the agreement considerably by including the first corrections to the rigid rotor–harmonic oscillator model. These include centrifugal distortion effects, anharmonic effects, and other extensions. The consideration of these effects introduces a new set of molecular constants, all of which are determined spectroscopically and are well tabulated. (See Problem 6–24.) The use of such

Table 6–3. **The entropies of some diatomic molecules calculated according to Eq. (6–54) compared to the experimental values at 1-atm pressure and 25°C***

	S(calc.) (e.u.)	S(exp.) (e.u.)
H_2	31.1	31.2
O_2	49.0	49.0
N_2	45.7	45.7
Cl_2	53.2	53.3
HCl	44.6	44.6
HBr	47.4	47.4
HI	49.4	49.3
CO	47.2	46.2

* The experimental values have been corrected for any nonideal gas behavior.

additional parameters from spectroscopic data can give calculated values of the entropy and heat capacity that are actually more accurate than experimental ones.

It should be pointed out, however, that extremely accurate calculations can require a sophisticated knowledge of molecular spectroscopy. For example, we said above that the electronic partition function was similar to that in the atomic case. This, however, is not entirely true. For molecules in states other than a $\sum$ state (which has zero total angular momentum), the total electronic angular momentum must be coupled with the overall rotational angular momentum, and this coupling must be treated in a detailed quantum mechanical way. This is too specialized to discuss here, but the result of this coupling is that the electronic and rotational partition functions do not separate. When $T \gg \Theta_r$, however, the molecules are in states with large enough rotational quantum numbers [*cf.* Eq. (6–39)] that the angular momentum coupling is no longer important, and the rotational-electronic partition function separates into a rotational part and an electronic part. Since we have chosen the zero of energy to be the separated electronically unexcited atoms at rest, the electronic partition function is

$$q_e(T) = \omega_{e1} e^{D_e/kT} + \omega_{e2} e^{-\varepsilon_2/kT} + \cdots$$

where the ω_{ej} are the degeneracies, and the ε_j's are measured relative to the ground electronic state of the molecule. Keep in mind, however, that for some molecules, such as NO, this equation is valid only at high temperatures, and that the low-temperature partition function requires a fairly specialized knowledge of the coupling of electronic and rotational angular momenta. See Herzberg under "Additional Reading" for a thorough discussion of this complication.

It seems logical at this time to go on to a discussion of polyatomic molecules in much the same manner as we have for diatomics. We would see very quickly, however, that unless the molecule possesses a certain degree of symmetry, it is impossible to write down any closed-form expression for its rotational energy levels. This means that a calculation of $q_{\mathrm{rot}}(T)$ is at best a complicated numerical problem. This would appear to imply that we have come to the end of the line for statistical thermodynamical applications, and we have not even begun to consider *interactions* between molecules! Even just *two* simple molecules, to say nothing of N particles, interacting through any kind of realistic interatomic potential becomes an extremely complicated quantum mechanical problem. At this point we must go back and reconsider some of the results we have derived up to now.

ADDITIONAL READING

General

Davidson, N. 1962. *Statistical mechanics.* New York: McGraw-Hill. Chapters 8 and 9.

Fowler, R. H., and Guggenheim, E. A. 1956. *Statistical thermodynamics.* Cambridge: Cambridge University Press. Sections 312 through 325.

Hill, T. L. 1960. *Statistical thermodynamics.* Reading, Mass.: Addison-Wesley. Chapter 8.

Kestin, J., and Dorfman, J. R. 1971. *A course in statistical thermodynamics.* New York: Academic. Sections 6–9, 6–10, and 6–13.

Knuth, E. 1966. *Statistical and thermodynamics.* New York: McGraw-Hill. Chapter 13.

Kubo, R. 1965. *Statistical mechanics.* Amsterdam: North-Holland Publishing Co. Chapter 3 and Example 3–1.

Mayer, J. E., and Mayer, M. G. 1940. *Statistical mechanics.* New York: Wiley. Chapter 7.

Münster, A. 1969. *Statistical thermodynamics.* Vol. I. Berlin: Springer-Verlag. Sections 6–3 through 6–7.

Rushbrooke, G. S. 1949. *Statistical mechanics.* London: Oxford University Press. Chapters 6, 7, and 8.

Spectroscopy

Barrow, G. 1962. *Introduction to molecular spectroscopy.* New York: McGraw-Hill.

Dunford, H. B. 1968. *Elements of diatomic molecular spectra.* Reading, Mass.: Addison-Wesley.

Herzberg, G. 1950. *Spectra of diatomic molecules,* 2nd ed. New York: Van Nostrand.

King, G. W. 1964. *Spectroscopy and molecular structure.* New York: Holt, Rinehart and Winston.

PROBLEMS

6–1. The Morse potential is

$$U(r) = D_e(1 - e^{-\beta(r-r_e)})^2$$

Show that $\beta = \nu(2\pi^2\mu/D_e)^{1/2}$.

6–2. The dissociation energy D_0 of H_2 is 103.2 kcal/mole, and its fundamental vibrational frequency $\bar{\omega}$ is 4320 cm^{-1}. From this information, calculate D_0 and $\bar{\omega}$ for D_2, T_2, and HD, assuming the Born-Oppenheimer approximation.

6–3. Given that D_0 for H_2 is 103.2 kcal/mole and that Θ_v is 6215°K, calculate D_0 for both D_2 and T_2.

6–4. Show that the moment of inertia of a diatomic molecule is μr_e^2, where μ is the reduced mass, and r_e is the equilibrium separation.

6–5. Show that the force constants in Table 6–1 are consistent with the frequencies given there.

6–6. Using the data in Table 6–1, calculate the frequencies that are expected to be found in the rotational spectrum of HCl.

6–7. In the far infrared spectrum of HBr, there is a series of lines separated by 16.72 cm^{-1}. Calculate the moment of inertia and internuclear separation in HBr.

6–8. Show that the vibrational contribution to the heat capacity C_V of a diatomic molecule is Nk as $T \to \infty$.

6–9. Derive Eq. (6–35) from the Euler–MacLaurin summation formula.

6–10. Show that the rotational level that is most populated is given by $J_{max} = (kT/2\bar{B})^{1/2}$. Calculate J_{max} for CO_2 and H_2 at room temperature.

6–11. The rotational constant $\bar{B}$ for $HC^{12}N^{14}$ is 44,315.97 MHz (megahertz) and $DC^{12}N^{14}$ for 36,207.40 MHz. Deduce the moments of inertia for these molecules. Assuming that the bond lengths are independent of isotopic substitution, calculate the H–C and C–N bond length.

6–12. Given that the values of Θ_r and Θ_v for H_2 are 85.3 K and 6215°K, respectively, calculate these quantities for HD and D_2.

6–13. What is the most probable value of the rotational quantum number J of a gas phase N_2 molecule at 300°K? What is the most probable vibrational quantum number n for this same situation?

6–14. Using the Euler-MacLaurin expansion, derive the second- and third-order corrections to the (first-order) high-temperature limit of E_r and C_{V_r}. Express your result in terms of a power series of Θ_r/T.

6–15. Calculate the rotational contribution to the entropy of HD at 20°K, 100°K, and 300°K, using the formulas appropriate for each particular temperature, and estimate the error involved in each.

6–16. Discuss the statistical weights of a hypothetical diatomic molecule X_2 with a ground electronic state $\sum_u^{-}$, supposing the X nuclei have integral spin (bosons) and half-integral spin (fermions). Derive the rotational partition function and the rotational contribution to the heat capacity C_V for each case.

6–17. Calculate the percent of para-D_2 as a function of temperature (assuming equilibrium) and also calculate the heat capacity of the equilibrium mixture, para-D_2, ortho-D_2, and finally what you expect would be the experimental heat capacity.

6–18. Why does one not see discussions in the literature concerning the ortho-para forms of fluorine?

6–19. Show that the thermodynamic quantities p and C_V are independent of the choice of a zero of energy.

6–20. In the far infrared spectrum of HCl, there is a series of lines with an almost constant spacing of 20.7 cm^{-1}. In the near infrared spectrum, there is one intense band at 3.46 microns. Use these data to calculate the entropy of HCl at 300°K and 1 atm (assuming ideal behavior).

6–21. Molecular nitrogen is heated in an electric arc, and it is found spectroscopically that the relative populations of excited vibrational levels is

n	0	1	2	3	4
$\dfrac{f_n}{f_0}$	1.000	0.200	0.040	0.008	0.002

Is the nitrogen in thermodynamic equilibrium with respect to vibrational energy? What is the vibrational temperature of the gas? Is this necessarily the same as the translational temperature?

6–22. Without looking anything up, put in order of decreasing magnitudes the following "temperatures":

$$\Theta_v{}^{\mathrm{H2}},\ \Theta_r{}^{\mathrm{H2}},\ \Theta_v{}^{\mathrm{Cl2}},\ \Theta_r{}^{\mathrm{Cl2}},\ \Theta_v{}^{\mathrm{HCl}},\ \Theta_r{}^{\mathrm{HCl}}$$

6–23. A more accurate expression for the vibrational energy of a diatomic molecule is

$$\varepsilon_n = (n + \tfrac{1}{2})h\nu - x_e(n + \tfrac{1}{2})^2 h\nu$$

where x_e is called the anharmonicity constant. The additional term here represents the first deviations from strictly harmonic behavior. Treating x_e as a small parameter, calculate the anharmonic effect on the various thermodynamic functions at least to first order in x_e.

6–24. The model of a diatomic molecule presented in this chapter is called the rigid rotor–harmonic oscillator model. The rotational-vibrational energy in this approximation is

$$\varepsilon_{vr} = (n + \tfrac{1}{2})h\nu + \bar{B}J(J + 1)$$

This expression can be improved in a number of ways. The harmonic oscillator approximation can be modified to include terms that reflect the deviations from harmonic behavior (anharmonicity) as the vibrational energy of the molecule increases. This is done by quantum mechanical perturbation theory, which gives

$$\varepsilon_v = (n + \tfrac{1}{2})h\nu - x_e(n + \tfrac{1}{2})^2 h\nu + \cdots$$

where x_e is a small constant called the anharmonicity constant. In addition to this, there is a correction due to the fact that the molecule is not a rigid rotor and, in fact, stretches some as the molecule rotates with greater energy. This is also handled by perturbation theory and gives

$$\varepsilon_r = \bar{B}J(J+1) - \bar{D}J^2(J+1)^2 + \cdots$$

where $\bar{D}$ is a small constant called the centrifugal distortion constant. Lastly, there exists a coupling between the rotational and vibrational modes of the molecule, since its moment of inertia changes as the molecule vibrates. Putting all this together gives

$$\varepsilon_{vr} = (n+\tfrac{1}{2})h\nu + \bar{B}J(J+1) - x_e(n+\tfrac{1}{2})^2h\nu - \bar{D}J^2(J+1)^2 - \alpha(n+\tfrac{1}{2})J(J+1)$$

where α is the rotation-vibration coupling constant. These terms, which correct the rigid rotor–harmonic oscillator approximation, are usually quite small. Using this more rigorous expression for ε_{vr}, show that the molecular partition function can be written in the form

$$q(V, T) = q_{\text{rr-ho}}q_{\text{corr}}$$

where

$$q_{\text{corr}} = 1 + \frac{2kT}{\bar{B}}\left(\frac{\bar{D}}{\bar{B}}\right) + \frac{1}{e^{\beta h\nu}-1}\left(\frac{\alpha}{\bar{B}}\right)$$

$$+ \frac{2\beta h\nu}{(e^{\beta h\nu}-1)^2}\,x_e + \text{higher-order terms in } \bar{D},\ \alpha, \text{ and } x_e$$

Calculate the effect of q_{corr} on E and C_V for O_2 at 300°K, given the following values of the spectroscopic parameters: $x_e = 0.0076$, $\bar{D} = 4.8 \times 10^{-6}\ \text{cm}^{-1}$, and $\alpha = 0.016\ \text{cm}^{-1}$.

6–25. Consider a system of independent diatomic molecules constrained to move in a plane, that is, a two-dimensional ideal diatomic gas. How many degrees of freedom does a two-dimensional diatomic molecule have? Given that the energy eigenvalues of a two-dimensional rigid rotor are

$$\varepsilon_J = \frac{\hbar^2J^2}{2I} \qquad J = 0, 1, 2, \ldots$$

with a degeneracy $\omega_J = 2$ for all J except J 0, calculate the rotational partition function. I is the moment of inertia of the molecule. The vibrational partition function is the same as for a three-dimensional diatomic gas. Write out

$$q(T) = q_{\text{trans}}(T)q_{\text{rot}}(T)q_{\text{vib}}(T)$$

and derive an expression for the average energy of this two-dimensional ideal diatomic gas.

6–26. Show that the molecule $O^{16}O^{16}$ has only odd rotational levels in its ground electronic state.

CHAPTER 7

CLASSICAL STATISTICAL MECHANICS

So far we have been able to derive translational, rotational, and vibrational partition functions for linear molecules. In each case we saw that if the temperature were high enough we could replace sums by integrals and obtain high-temperature limits. These prove to be numerically satisfactory for most gases at ordinary temperatures (with the exception perhaps of the vibrational case).

As the temperature increases, the average energy per molecule increases, and so in a quantum mechanical sense, the quantum numbers describing this motion (n_x for translational, J for rotational, etc.) also increase, meaning that the molecules are in the high quantum number limit. For example at room temperature translational quantum numbers are typically 10^8. (See Problem 5–3.) It is the recognition of this fact that will point the way to a solution to the problem discussed in the last paragraph of Chapter 6.

It is one of the fundamental principles of quantum mechanics that classical behavior is obtained in the limit of large quantum numbers. So we see that up to now our procedure has been to solve a particular quantum mechanical problem, use this result in the molecular partition function, use a high-temperature approximation, and then find that this high-temperature limit is satisfactory. In other words, we were starting with a quantum mechanical solution and then taking the classical limit at a later stage. It is natural to seek a procedure in which we can use classical mechanics throughout, and such an approach is developed in this chapter. For simplicity we shall first consider only molecular partition functions, although we shall generalize our results afterward.

7–1 THE CLASSICAL PARTITION FUNCTION

Consider the molecular partition function

$$q = \sum_j e^{-\beta \varepsilon_j} \tag{7–1}$$

This is of the form of a sum of $e^{-\beta(\text{energy})}$ over all possible quantum states. It is natural to assume that the corresponding classical expression is a similar sum, or since the energy in the classical sense is a continuous function of the momenta p_j and coordinates q_j, this sum would become an *integral* over all the possible classical "states" of the system. Since the classical energy is the Hamiltonian function $H(p, q)$, the molecular partition function $q(V, T)$ becomes

$$q_{\text{class}} \sim \int \cdots \int e^{-\beta H(p,q)}\, dp\, dq \tag{7-2}$$

In Eq. (7–2) the notation (p, q) denotes all the momenta and coordinates on which H depends; dp stands for $dp_1\, dp_2 \cdots dp_s$ and dq for $dq_1 \cdots dq_s$, where s is the number of momenta or coordinates necessary to completely specify the motion or position of the molecule. The quantity s represents the number of degrees of freedom of the molecule. The set of coordinates $\{q_j\}$ does not necessarily have to be a set of Cartesian coordinates, and more usually represents a set of generalized coordinates, that is, any set of coordinates that conveniently specifies the position of the molecule. For a mass point, for example, the generalized coordinates might be simply x, y, and z; for a rigid rotor, we might choose the two angles θ and ϕ needed to specify the orientation of the molecule. Usually the choice of generalized coordinates is obvious. The momenta $\{p_j\}$ in Eq. (7–2) are the generalized momenta conjugate to the $\{q_j\}$ [*cf.* Eq. (1–19)].

At this stage Eq. (7–2) is just a plausible conjecture. Let us now pursue this idea by considering a monatomic ideal gas once again. From Eq. (5–8), we have

$$q_{\text{trans}}(V, T) = \left(\frac{2\pi mkT}{h^2}\right)^{3/2} V$$

The classical Hamiltonian of one atom of a monatomic ideal gas is simply the kinetic energy:

$$H = \frac{1}{2m}(p_x^{\,2} + p_y^{\,2} + p_z^{\,2})$$

According to Eq. (7–2), then,

$$q_{\text{class}} \sim \int \cdots \int \exp\left\{-\frac{\beta(p_x^{\,2} + p_y^{\,2} + p_z^{\,2})}{2m}\right\} dp_x\, dp_y\, dp_z\, dx\, dy\, dz \tag{7-3}$$

Notice here that since it takes three coordinates to specify the position of a point particle, q_{class} is a six-fold integral. The integral over $dx\, dy\, dz$ simply yields the volume of the container V, and so we have

$$q_{\text{class}} \sim V\left\{\int_{-\infty}^{\infty} e^{-\beta p^2/2m}\, dp\right\}^3 = (2\pi mkT)^{3/2} V \tag{7-4}$$

We see that except for a factor of Planck's constant cubed, this is just the translational partition function that we obtained before. Of course, we cannot expect to derive a purely classical expression that contains h, and so although our conjecture may be incomplete, there seems to be some element of truth to it.

Let us see how this procedure works for the other partition functions that we have evaluated. For the rigid rotor, the Hamiltonian is

$$H = \frac{1}{2I}\left(p_\theta^{\,2} + \frac{p_\phi^{\,2}}{\sin^2\theta}\right)$$

where I is the moment of inertia of the molecule. The generalized coordinates and momenta in this case are θ, ϕ, p_θ, and p_ϕ, and so Eq. (7–2) is

$$q_{\text{rot}} \sim \int_{-\infty}^{\infty} \int dp_\theta \, dp_\phi \int_0^{2\pi} d\phi \int_0^{\pi} d\theta e^{-\beta H} = (8\pi^2 IkT) \tag{7–5}$$

For the classical harmonic oscillator,

$$H = \frac{p^2}{2\mu} + \frac{k}{2} x^2 \tag{7–6}$$

and

$$q_{\text{vib}} \sim \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dx e^{-\beta H} = \frac{kT}{\nu} \tag{7–7}$$

where

$$\nu = \frac{1}{2\pi}\left(\frac{k}{\mu}\right)^{1/2}$$

We can see from these three examples that the translational partition function is incorrect by a factor of h^3; the rotational partition function is incorrect by a factor of h^2; and the vibrational partition function is incorrect by a factor of h. It appears that a factor of h results for each product $dp_j \, dq_j$ occurring in q_{class}. Since partition functions are dimensionless, and h has units of momentum times length, we see that this at least automatically satisfies a dimensional requirement. We shall therefore *assume* that

$$q = \sum_j e^{-\beta \varepsilon_j} \to \frac{1}{h^s} \int \cdots \int e^{-\beta H} \prod_{j=1}^{s} dp_j \, dq_j \tag{7–8}$$

We now shall extend this assumption to systems of molecules. Equation (4–10) says that at high enough temperatures, we can write for a system of N independent indistinguishable particles

$$\begin{aligned} Q &= \frac{q^N}{N!} \\ &= \frac{1}{N!} \prod_{j=1}^{N} \left\{ \frac{1}{h^s} \int \cdots \int e^{-\beta H_j} \prod_{i=1}^{s} dp_{ji} \, dq_{ji} \right\} \end{aligned}$$

where H_j is the Hamiltonian of the jth molecule and is a function of $p_{j1}, \ldots, p_{js}$, $q_{j1}, \ldots, q_{js}$. We now simply relabel the momenta and coordinates such that p_1 through p_s represent p_{11} through p_{1s}; p_{s+1} through p_{s+s} represent p_{21} through p_{2s}, and so on, and write

$$\begin{aligned} Q &= \frac{1}{N! \, h^{sN}} \int \cdots \int e^{-\beta \Sigma_j H_j} \prod_{i=1}^{sN} dp_i \, dq_i \\ &= \frac{1}{N! \, h^{sN}} \int \cdots \int e^{-\beta H} \prod_{i=1}^{sN} dp_i \, dq_i \end{aligned}$$

where H is the Hamiltonian of the N-body system. This form suggests the classical limit of Q for systems of interacting particles. We *conjecture* that

$$Q = \frac{1}{N! \, h^{sN}} \int \cdots \int e^{-\beta H(p,q)} \, dp \, dq \tag{7–9}$$

where $H(p, q)$ is the classical N-body Hamiltonian for *interacting* particles. The notation (p, q) represents the set of p_j's and q_j's that describes the entire system, and $dp\, dq$ represents

$$\prod_{j=1}^{sN} dp_j\, dq_j$$

We have assumed then that the classical limit of $Q(N, V, T)$ is given by

$$Q = \sum_j e^{-\beta E_j} \rightarrow \frac{1}{N!\,h^{sN}} \int \cdots \int e^{-\beta H(p,q)}\, dp\, dq \tag{7–10}$$

For a monatomic gas, for example,

$$H(p, q) = \frac{1}{2m} \sum_{j=1}^{N} (p_{xj}{}^2 + p_{yj}{}^2 + p_{zj}{}^2) + U(x_1, y_1, z_1, \ldots, x_N, y_N, z_N) \tag{7–11}$$

Equation (7–10) is, in fact, the correct classical limit of Q, although we have not proved it here. It is actually possible to start with the quantum mechanical sum in Eq. (7–10) and to derive the integral as the classical limit, that is, the limiting result as $h \rightarrow 0$ (*cf.* Section 10–7).

If we substitute Eq. (7–11) into Eq. (7–10), the momentum integrations can be done easily, and we get

$$Q_{\text{class}} = \frac{1}{N!} \left(\frac{2\pi mkT}{h^2}\right)^{3N/2} Z_N \tag{7–12}$$

where

$$Z_N = \int_V e^{-U(x_1, \ldots, z_N)/kT}\, dx_1 \cdots dz_N \tag{7–13}$$

In Eq. (7–13), Z_N is called the *classical configuration integral.* Since the intermolecular forces depend upon the relative distances between molecules, this integral is, in general, extremely difficult and is essentially responsible for the research in equilibrium statistical mechanics. In the absence of intermolecular forces, $U = 0$ and $Z_N = V^N$ Equations (7–12) and (7–13) are fundamental equations in the study of monatomic, classical, imperfect gases and liquids.

It often happens that not all of the degrees of freedom of a molecule can be treated classically. For example, we have seen that the spacing between translational and rotational levels is small enough that the sum over states or levels can be replaced by an integral, that is, these degrees of freedom can be treated classically. This is not the case, however, with the vibrational degrees of freedom, and these degrees of freedom must be treated quantum mechanically.

Suppose, then, that the Hamiltonian of a molecule can be written as

$$H = H_{\text{class}} + H_{\text{quant}} \tag{7–14}$$

where H_{class} refers to the s degrees of freedom that can be treated classically, and H_{quant} refers to the degrees of freedom that cannot be treated classically. Then

$$q = q_{\text{class}}\, q_{\text{quant}} \tag{7–15}$$

where

$$q_{\text{class}} = \frac{1}{h^s} \int e^{-H_{\text{class}}(p,q)/kT}\, dp_1\, dq_1 \cdots dp_s\, dq_s \tag{7–16}$$

Note that Eq. (6–51) is of the form of Eq. (7–15), where the translational and rotational degrees of freedom are treated classically, and the vibrational and electronic degrees of freedom are treated quantum mechanically.

Equations (7–14) to (7–16) are immediately generalizable to a system of interacting molecules. If the Hamiltonian of the entire system is separable into a classical part and a quantum part, then

$$H = H_{\text{class}} + H_{\text{quant}} \tag{7–17}$$

$$Q = Q_{\text{class}} Q_{\text{quant}} \tag{7–18}$$

$$= \frac{Q_{\text{quant}}}{N!\, h^{sN}} \int e^{-H_{\text{class}}/kT}\, dp_{\text{class}}\, dq_{\text{class}} \tag{7–19}$$

7-2 PHASE SPACE AND THE LIOUVILLE EQUATION

Until now our approach has been to go to the classical limit only when it was necessary. Historically, however, statistical mechanics was originally formulated by Boltzmann, Maxwell, and Gibbs in the nineteenth century before the evolution of quantum mechanics. Their formulation, therefore, was based on classical mechanics, and since this still is a most useful limit, we shall now discuss the classical mechanical formulation of statistical mechanics. This formalism forms the basis of most of the work involving interacting systems in equilibrium and nonequilibrium statistical mechanics that is done today.

Consider any classical system containing N (interacting) molecules. Let each molecule have s degrees of freedom, that is, each molecule requires s coordinates to completely describe its position. Let the number of coordinates necessary to describe the positions of all N molecules be $l = sN$. The l coordinates, $q_1, q_2, \ldots, q_l$, then completely describe the spatial orientation of the entire N-body system. To each of these l coordinates, there corresponds a conjugate momentum p_j, say, defined by Eq. (1–19). The l spatial coordinates $\{q_j\}$ and the l momenta $\{p_j\}$ completely specify the classical mechanical state of the N-body system. These $2l$ coordinates, along with the equations of motion of the system, completely determine the future and past course of the system.

We now construct a conceptual Euclidean space of $2l$ dimensions, with $2l$-rectangular axes, one for each of the spatial coordinates $q_1, \ldots, q_l$ and one for each of the momenta $p_1, \ldots, p_l$. Following Gibbs, we speak of such a conceptual space as a *phase space* for the system under consideration. The state of the classical N-body system at any time t is completely specified by the location of *one* point in phase space. Such a point is called a *phase point*. As the system evolves in time, its dynamics is completely described by the motion or trajectory of the phase point through phase space. The trajectory of the phase point is given by Hamilton's equations of motion:

$$\dot{q}_j = \frac{\partial H}{\partial p_j} \quad \text{and} \quad \dot{p}_j = -\frac{\partial H}{\partial q_j} \qquad j = 1, 2, \ldots, l = sN \tag{7–20}$$

In principle, these $2l$ equations can be integrated to give $\{q_j(t)\}$ and $\{p_j(t)\}$. For notational simplicity, we shall denote the set of l q's by $q(t)$, and the set of l p's by $p(t)$. The $2l$ constants of integration can be fixed by the location of the phase point at some initial time, say t_0. Of course, in practice, such an integration is not feasible.

We now introduce the concept of an ensemble of systems in phase space. For simplicity we shall consider a microcanonical ensemble, that is, an ensemble representative of an isolated system. Consider a large number $\mathscr{A}$ of isolated systems, each of which having the same values of macroscopic variables N, V, and E.

The detailed classical state of each system in the ensemble has a representative phase point in the *same* phase space. The entire ensemble then appears as a cloud of points in phase space. As time evolves, each point will trace out its *independent* trajectory. The trajectories are independent, since each one represents an isolated system and is, therefore, independent of all the others. The postulate of equal a priori probabilities requires that there is a representative phase point in phase space for each and every set of coordinates and momenta consistent with the few fixed macroscopic variables. In particular, the postulate of equal a priori probabilities states that for a microcanonical ensemble, the density is *uniform* over the constant energy "surface" in phase space, where the value of the energy on the surface is that of the isolated system. We consider all parts of phase space equally important, as long as the (p, q)'s are consistent with all that we know macroscopically about the system, that is, consistent with the values of N, V, and E for the system that the ensemble represents. Just as every quantum state was equally likely before, now we consider every classical state to be equally probable.

This cloud of points is very dense then, and we can define a number density $f(p, q, t)$, such that the number of systems in the ensemble that have phase points in $dp\,dq$ about the point p, q at time t is $f(p, q, t)\,dp\,dq$. Clearly we must have

$$\int \cdots \int f(p, q)\,dp\,dq = \mathscr{A} \tag{7-21}$$

The ensemble average of any function, say $\phi(p, q)$, of the momenta and coordinates of the system is defined as

$$\bar{\phi} = \frac{1}{\mathscr{A}} \int \cdots \int \phi(p, q) f(p, q, t)\,dp\,dq \tag{7-22}$$

It is Gibbs' postulate to equate this ensemble average to the corresponding thermodynamic function. Note the similarity between this equation and Eq. (2–5), its quantum mechanical analog.

Since the equations of motion determine the trajectory of each phase point, they must also determine the density $f(p, q, t)$ at any time if the dependence of f on p and q is known at some initial time t_0. The time dependence of f is thus controlled by the laws of mechanics and is not arbitrary. The time dependence of f is given by the Liouville equation, which we now derive.

Consider the small volume element $\delta p_1 \cdots \delta p_l\, \delta q_1 \cdots \delta q_l$ about the point p_1, $p_l, q_1, \cdots q_l$. The number of phase points inside this volume at any instant is

$$\delta N = f(p_1, \ldots, p_l, q_1, \ldots, q_l, t)\, \delta p_1 \cdots \delta p_l\, \delta q_1 \cdots \delta q_l$$

This number will, in general, change with time since the natural trajectories of phase points will take them into and out of this volume element, and the number passing through any one "face" will, in general, be different from the number passing through the opposite "face." Let us calculate the number entering one face and leaving through the opposite. Consider two faces perpendicular to the q_1-axis and located at q_1 and $q_1 + \delta q_1$. The number of phase points entering the first of these faces per unit time is

$$f\dot{q}_1\, \delta q_2 \cdots \delta q_l\, \delta p_1 \cdots \delta p_l \tag{7-23}$$

(See Problem 7–33.) The number passing through the other face per unit time is

$$f(q_1 + \delta q_1, q_2, \ldots, q_l, p_1, \ldots, p_l) \times \dot{q}_1(q_1 + \delta q_1, q_2, \ldots, q_l, p_1, \ldots, p_l)\, \delta q_2 \cdots \delta q_l\, \delta p_1 \cdots \delta p_l$$

which, if we expand f and $\dot{q}_1$ to linear terms in δq_1, gives

$$\left(f + \frac{\partial f}{\partial q_1}\delta q_1\right)\left(\dot{q}_1 + \frac{\partial \dot{q}_1}{\partial q_1}\delta q_1\right)\delta q_2 \cdots \delta q_l\, \delta p_1 \cdots \delta p_l + \cdots \tag{7–24}$$

Subtracting (7–24) from (7–23), we get the *net* flow of phase points in the q_1-direction into the volume element $\delta q_1 \cdots \delta q_l\, \delta p_1 \cdots \delta p_l$:

$$\text{net flow} = -\left(\frac{\partial f}{\partial q_1}\dot{q}_1 + f\frac{\partial \dot{q}_1}{\partial q_1}\right)\delta p_1 \cdots \delta p_l\, \delta q_1 \cdots \delta q_l$$

in the q_1-direction. In a similar manner, the net flow in the p_1-direction is (remember that momenta and spatial coordinates have equal status in phase space):

$$-\left(\frac{\partial f}{\partial p_1}\dot{p}_1 + f\frac{\partial \dot{p}_1}{\partial p_1}\right)\delta p_1 \cdots \delta p_l\, \delta q_1 \cdots \delta q_l$$

Thus the change in the number of phase points through all the faces is

$$\sum_{j=1}^{l}\left(\frac{\partial f}{\partial q_j}\dot{q}_j + f\frac{\partial \dot{q}_j}{\partial q_j} + \frac{\partial f}{\partial p_j}\dot{p}_j + f\frac{\partial \dot{p}_j}{\partial p_j}\right)\delta p_1 \cdots \delta p_l\, \delta q_1 \cdots \delta q_l.$$

This must be equal to the change of δN with time, and so we have

$$\frac{d(\delta N)}{dt} = -\sum_{j=1}^{l}\left[f\left(\frac{\partial \dot{q}_j}{\partial q_j} + \frac{\partial \dot{p}_j}{\partial p_j}\right) + \left(\frac{\partial f}{\partial q_j}\dot{q}_j + \frac{\partial f}{\partial p_j}\dot{p}_j\right)\right]\delta p_1 \cdots \delta p_l\, \delta q_1 \cdots \delta q_l \tag{7–25}$$

This result can be immediately simplified. Since

$$\dot{q}_j = \frac{\partial H}{\partial p_j} \qquad \dot{p}_j = -\frac{\partial H}{\partial q_j} \tag{7–26}$$

the first term in parentheses in Eq. (7–25) is

$$\frac{\partial \dot{q}_j}{\partial q_j} + \frac{\partial \dot{p}_j}{\partial p_j} = 0 \tag{7–27}$$

Furthermore, we divide Eq. (7–25) by the volume element $\delta p_1 \cdots \delta p_l\, \delta q_1 \cdots \delta q_l$. This gives the rate of change in the density itself around the point $p_1, \ldots p_l, q_1, \ldots q_l$, so that we can write

$$\frac{\partial f}{\partial t} = -\sum_{j=1}^{l}\left(\frac{\partial f}{\partial q_j}\dot{q}_j + \frac{\partial f}{\partial p_j}\dot{p}_j\right) \tag{7–28}$$

where we have written $\partial f/\partial t$ to indicate that we have fixed our attention on a given stationary point in the phase space.

Equation (7–28) can be written in a more conventional form by using Eqs. (7–26) for $\dot{q}_j$ and $\dot{p}_j$. The result is

$$\frac{\partial f}{\partial t} = -\sum_{j=1}^{l}\left(\frac{\partial H}{\partial p_j}\frac{\partial f}{\partial q_j} - \frac{\partial H}{\partial q_j}\frac{\partial f}{\partial p_j}\right) \tag{7–29}$$

This is the Liouville equation, the most fundamental equation of classical statistical mechanics. In fact, it can be shown that the Liouville equation is equivalent to the $6N$ Hamiltonian equations of motion of the N-body system. See Mazo under "Additional Reading" for a proof of this. In Cartesian coordinates, the Liouville equation for N point masses is (see Problem 7–11):

$$\frac{\partial f}{\partial t} + \sum_{j=1}^{N} \frac{\mathbf{p}_j}{m_j} \cdot \nabla_{\mathbf{r}_j} f + \sum_{j=1}^{N} \mathbf{F}_j \cdot \nabla_{\mathbf{p}_j} f = 0 \tag{7–30}$$

In this equation $\nabla_{\mathbf{r}_j}$ denotes the gradient with respect to the spatial variables in f; $\nabla_{\mathbf{p}_j}$ denotes the gradient with respect to the momentum variables in f; and $\mathbf{F}_j$ is the total force on the jth particle. The Liouville equation forms the starting point of most theories of nonequilibrium statistical mechanics.

There are several interesting deductions from the Liouville equation which we now discuss. Consider Eq. (7–28)

$$\frac{\partial f}{\partial t} + \sum_{j=1}^{l} \left(\frac{\partial f}{\partial p_j}\right) \dot{p}_j + \sum_{j=1}^{l} \left(\frac{\partial f}{\partial q_j}\right) \dot{q}_j = 0 \tag{7–31}$$

Since $f = f(p, q, t)$, this equation is equivalent to

$$\frac{df}{dt} = 0 \tag{7–32}$$

Physically, this equation says that the density in the neighborhood of any selected moving phase point is a constant along the trajectory of that phase point. Thus the cloud of phase points behaves as an incompressable fluid. Gibbs called this the principle of the conservation of density in phase. An equivalent statement of this is that if p, q are the coordinates of a phase point at time t, which at time t_0 were (p_0, q_0), then Liouville's equation implies that (see Problem 7–12)

$$f(p, q; t) = f(p_0, q_0; t_0) \tag{7–33}$$

Because of the equations of motions, the point (p, q) should be considered a function of the initial point (p_0, q_0) and the elapsed time t. That is

$$p = p(p_0, q_0; t)$$
$$q = q(p_0, q_0; t)$$

Now let us select a small element of volume at (p_0, q_0) at time t_0. At a later time, $t_0 + t$, the phase points originally on the surface of this volume element will have formed a new surface enclosing a volume element of different shape at the phase point (p, q). The volume element at (p, q) must contain the same number of phase points as the original volume element at (p_0, q_0). This follows because a phase point outside or inside the volume element can never cross the surface as the element moves through phase space, for otherwise there would be two different trajectories through the same point in phase space. This is impossible, however, because of the uniqueness of the equations of motion of a phase point. Trajectories of phase points can never cross. Now since the density and number of phase points in the volume element are the same at p_0, q_0 and p, q, it follows that although the shape of this volume may change and contort itself as it moves through phase space, its volume remains constant. Gibbs called this result *conservation of extension in phase space*. This fact is expressed mathematically by writing

$$\delta p \, \delta q = \delta p_0 \, \delta q_0 \qquad \text{for all } t \tag{7–34}$$

Another way of expressing this is to say that the Jacobian of the set (p, q) to (p_0, q_0) is unity. This can be proved directly from the equations of motion of the system. See Mazo under "Additional Reading."

A corollary of this theorem, whose proof demands a more extensive knowledge of classical mechanics, is that if we are given two sets of coordinates and their conjugate momenta, say,

$$q_1, q_2, \ldots, q_{3n}, p_1, p_2, \ldots, p_{3n}$$
$$Q_1, Q_2, \ldots, Q_{3n}, P_1, P_2, \ldots, P_{3n}$$

which can describe a system in phase space equally well, then

$$dq_1\, dq_2 \cdots dq_{3n}\, dp_1 \cdots dp_{3n} = dQ_1 \cdots dQ_{3n}\, dP_1 \cdots dP_{3n}$$

For example, a single particle in three dimensions may be described by the coordinates (x, y, z) or the spherical coordinates (r, θ, ϕ). It is straightforward, albeit lengthy to show that

$$dp_x\, dp_y\, dp_z\, dx\, dy\, dz = dp_r\, dp_\theta\, dp_\phi\, dr\, d\theta\, d\phi \tag{7-35}$$

Notice that although the volume elements in ordinary coordinate space are $dx\, dy\, dz$ and $r^2 \sin\theta\, dr\, d\theta\, d\phi$, the $r^2 \sin\theta$ factor does not occur in the phase space transformation. These simple volume element transformations would not generally be true if we have chosen the generalized coordinates and velocities instead of momenta. This is one reason why momenta and not velocities are used to describe classical systems.

7-3 EQUIPARTITION OF ENERGY

We have seen that classical statistical mechanics is applicable when the temperature is high enough to replace the quantum statistical summation by an integral. Under these conditions, it is not necessary to know the eigenvalues of the quantum mechanical problem, only the classical Hamiltonian is required. There is an interesting theorem of classical statistical mechanics which can be used to understand more fully some of the results of the last two chapters.

Consider the expression for the average energy of a molecule in a system of independent molecules,

$$\bar{\varepsilon} = \frac{\iint He^{-\beta H}\, dp_1 \cdots dq_s}{\iint e^{-\beta H}\, dp_1 \cdots dq_s} \tag{7-36}$$

which can be evaluated in principle for any known dependence of H on the p's and the q's. Multiplying by the total number of molecules gives an expression for the total energy of the system, and by differentiating with respect to T, we obtain an expression for its heat capacity at constant volume.

If it so happens that the Hamiltonian is of the form

$$H(p_1, p_2, \ldots, q_s) = \sum_{j=1}^{m} a_j p_j^2 + \sum_{j=1}^{n} b_j q_j^2 + H(p_{m+1}, \ldots, p_s, q_{n+1}, \ldots, q_s) \tag{7-37}$$

where the a_j and b_j are constants, then it is easy to show that each of these quadratic terms will contribute $kT/2$ to the energy and $k/2$ to the heat capacity. (See Problem 7-29.) This result is called the principle of equipartition of energy. It should be emphasized that the principle is a consequence of the quadratic form of terms in the Hamiltonian, rather than a general consequence of classical statistical mechanics.

Let us apply this general theorem to some of the cases we have treated in Chapters 5 and 6. For instance, for a monatomic ideal gas, the Hamiltonian is

$$H = \frac{p_x^2 + p_y^2 + p_z^2}{2m} \tag{7–38}$$

Since there are three quadratic terms, each atom contributes $3kT/2$ to the total energy and so $3k/2$ to the constant volume heat capacity. This is exactly our result in Chapter 5. For the case of a rigid rotor, the Hamiltonian is

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

The $\sin^2\theta$ in the p_ϕ^2 term would seem to exclude the p_ϕ^2 term from the principle of equipartition, since Eq. (7–37) requires that the coefficients a_j be constants. There is a more general version, however, that allows the a_j and b_j to be functions of the momenta and coordinates not involved in the quadratic terms, that is, to be functions of $p_{m+1}, \ldots, p_s$ and $q_{n+1}, \ldots, q_s$ in Eq. (7–37). The proof of this is more difficult than the proof of the simpler version. (See either Problem 7–30 or Tolman under "Additional Reading.") Because of this, each quadratic term above still contributes its equipartition value, and so the rotational contribution of a rigid rotor to the energy is kT per molecule, just as we obtained in Chapter 6 [*cf.* Eq. (6–36)].

Note that equipartition is a classical concept, that is, the degree of freedom contributing must be such that $\Delta\varepsilon/kT$ is small in passing from one level to another. We have seen that this is true for translational and rotational degrees of freedom at ordinary temperatures, but not vibrational degrees of freedom. The heat capacity for an ideal diatomic gas in the rigid rotor–harmonic oscillator approximation is [*cf.* Eq. (6–53)]

$$C_V = \tfrac{5}{2}Nk + \frac{Nk(\Theta_v/T)^2 e^{\Theta_v/T}}{(e^{\Theta_v/T} - 1)^2} \tag{7–39}$$

where the $\frac{5}{2}Nk$ comes from the translational plus rotational degrees of freedom which, we have seen, are excited enough to be treated classically. The second term is the vibrational contribution, which reaches its expected classical limit of Nk, since the classical Hamiltonian for a harmonic oscillator is $(p^2/2m) + (k/2)x^2)$, when Θ_v/T becomes small, which is far above room temperature for most molecules. A value of the vibrational contribution to C_V differing from Nk is thus a quantum mechanical result.

There are more general formulations of the principle of equipartition of energy than we have given here, but they are not necessary for most purposes. In fact, the principle itself is perhaps more of historical interest today than actual practical interest. It is interesting to note in this regard that when the electronic structure of atoms and metals evolved toward the end of the nineteenth century, it was of great concern to Gibbs that the electrons contributed only a very small fraction of their equipartition value to the heat capacities of metals. He did not live to see this anomalous result completely explained by quantum statistics. Since electrons have such a small mass, they behave not at all classically and should, therefore, not be governed by the equipartition of energy (*cf.* Section 10–2).

We have made this long detour through phase space for more than just historical reasons. As we said earlier, most of the systems of interest to chemists can be treated very satisfactorily by classical methods. In fact, the quantum statistical theories of systems of interacting particles are quite a demanding and specialized subject whose techniques are still being developed. Fortunately, being chemists, we are spared from having to master these techniques. Even today the classical Liouville equation forms the starting point for most of the rigorous approaches to nonequilibrium statistical mechanics. We shall now discuss the problem that sent us here in the first place, namely, the study of ideal polyatomic gases.

ADDITIONAL READING

General

EYRING, H., HENDERSON, D., STOVER, B. J., and EYRING, E. M. 1964. *Statistical mechanics and dynamics*. New York: Wiley. Chapter 7.

GIBBS, J. W. 1960. *Elementary principles in statistical mechanics*. New York: Dover.

HILL, T. L. 1956. *Statistical mechanics*. New York: McGraw-Hill. Chapter 1.

HUANG, K. 1963. *Statistical mechanics*. New York: Wiley. Chapter 7.

KHINCHIN, A. I. 1949. *Mathematical foundations of statistical mechanics*. New York: Dover.

KILPATRICK, J. E. 1967. In *Physical chemistry, an advanced treatise*, Vol. II, ed. by H. Eyring, D. Henderson, and W. Jost. New York: Academic.

KUBO, R. 1965. *Statistical mechanics*. Amsterdam: North-Holland Publishing Co. Sections 1–1 to 1–14.

MAZO, R. M. 1967. *Statistical mechanical theories of transport processes*. New York: Pergamon. Chapter 2.

MÜNSTER, A. 1969. *Statistical thermodynamics*, Vol. I. Berlin: Springer-Verlag. Chapter 1.

RUSHBROOKE, G. S. 1949. *Statistical mechanics*. London: Oxford University Press. Chapter 4.

TER HAAR, D. 1966. *Elements of thermostatistics*. London: Oxford University Press. Chapter 5.

———. 1955. *Rev. Mod. Phys.* **27**, p. 289, 1955.

TOLMAN, R. C. 1938. *Statistical mechanics*. London: Oxford University Press. Chapter 3.

PROBLEMS

7–1. Show that at room temperature the translational quantum numbers are typically around 10^8 or so.

7–2. What is the constant energy surface in phase space for a simple linear harmonic oscillator? What is it for a single-point mass? What is it for an ideal gas of N-point masses?

7–3. Convince yourself that trajectories in phase space can never cross, also that surfaces (really hypersurfaces) of constant energy can never intersect if the energies are different.

7–4. Consider a classical ideal gas enclosed in an infinitely tall cylinder in a gravitational field. Assuming that the temperature is uniform up the cylinder, derive the barometric formula

$$p(z) - p(0)\exp\left(\frac{-mgz}{kT}\right)$$

From this calculate the atmospheric pressure at the top of Mt. Everest.

7–5. An ideal gas consisting of N particles of mass m is enclosed in an infinitely tall cylindrical container placed in a uniform gravitational field, and is in thermal equilibrium. Calculate the classical partition function, Helmholtz free energy, mean energy, and heat capacity of this system.

7–6. Consider a perfect gas of molecules with permanent electric dipole moments $\boldsymbol{\mu}$ in an electric field $\mathscr{E}$. Neglecting the polarizability of the molecules, the potential energy is

$$U = -\mu\mathscr{E}\cos\theta$$

where θ is the angle between $\boldsymbol{\mu}$ and $\mathscr{E}$. Using classical mechanics, derive an expression for the additional effect of $\mathscr{E}$ on the energy E and heat capacity of the gas.

7–7. The potential energy of N molecules in a container V can often be fairly well approximated by a sum of pair-wise potentials:

$$U(\mathbf{r}_1, \ldots, \mathbf{r}_N) \quad \sum_{i<j} u(\mathbf{r}_i, \mathbf{r}_j)$$

In addition, the pair-wise potentials $u(\mathbf{r}_i, \mathbf{r}_j)$ are often assumed to depend only upon the distance $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ between the two molecules. Thus one often writes

$$U(\mathbf{r}_1, \ldots, \mathbf{r}_N) \quad \sum_{i<j} u(r_{ij})$$

Convince yourself that even these two simplifications of U do not help in trying to evaluate the configuration integral Z.

7–8. It is possible to determine the value of Boltzmann's constant by observing the distribution of suspended Brownian particles in a gravitational field as a function of their height z. Given that the particles have a mass of 1.0×10^{-14} g, that the temperature is 300°K, and the following data:

z(cm)	Number of particles
0.0000	100
0.0025	55
0.0050	31
0.0075	17
0.0100	9

calculate the value of the Boltzmann constant.

7–9. We can calculate the microcanonical ensemble partition function for a classical monatomic ideal gas in the following way. This partition function is given by

$$\Omega(E, \Delta E) = \frac{1}{N!h^{3N}} \int \cdots \int^{*} dp_1\, dp_2 \cdots dq_{3N}$$

where the asterisk indicates that one integrates over the region of phase space such that

$$E - \Delta E \leq \frac{1}{2m} \sum_{j=1}^{3N} p_j^2 \leq E$$

We have seen in the quantum mechanical case that the thermodynamic consequences of this equation are remarkably insensitive to the value of ΔE. (See Problem 3–14.) We can find Ω most readily by first evaluating

$$I(E) = \frac{1}{N!h^{3N}} \int \cdots \int^{*} dp_1\, dp_2 \cdots dq_{3N}$$

where now the asterisk signifies the constraint

$$0 \leq \frac{1}{2m} \sum_{j=1}^{3N} p_j^2 \leq E$$

Note that $\Omega(E, \Delta E)$ is given by $I(E) - I(E - \Delta E)$. The integration of $dq_1 \cdots dq_{3N}$ in $I(E)$ immediately gives V^N, and the remaining integration over the momenta is just the volume of a $3N$-dimensional sphere of radius $(2mE)^{1/2}$. The volume of a $3N$-dimensional sphere of radius R is (see Problem 1–24)

$$\frac{\pi^{3N/2}}{(3N/2)!} R^{3N}$$

(Note that this reduces correctly when $3N = 2$ and 3.) Using this formula then, show that

$$I(E) = \frac{\pi^{3N/2} V^N (2mE)^{3N/2}}{N!\, h^{3N} (3N/2)!}$$

is in agreement with Eq. (1–36).

7–10. In Problem 3–14 we showed that the entropy could be calculated from $k \ln \Omega(E)\Delta E$ or $k \ln \Phi(E)$, where $\Omega(E)\Delta E$ is the number of states with energies between E and $E + \Delta E$, and $\Phi(E)$ is the total number of states with energies less than E. In addition to this, we showed that the result is remarkably insensitive to the choice of ΔE. We shall now discuss the classical analog of this. In particular, this problem involves showing that the volume of an N-dimensional sphere is essentially the same as the volume of the hypershell of thickness s. First write the volume of the hypersphere as

$$V_{\text{sphere}}(R) = \text{const} \times R^N$$

Now show that if N is large enough such that $sN \gg R$, then

$$\begin{aligned} V_{\text{shell}} &= V(R) - V(R - s) \\ &= \text{const} \times R^N(1 - e^{-sN/R}) \\ &\approx V_{\text{sphere}} \end{aligned}$$

7–11. Show that in Cartesian coordinates, the Liouville equation takes the form of Eq. (7–30).

7–12. Convince yourself that a corollary of Liouville's equation is

$$f(p, q; t) = f(p_0, q_0; t_0)$$

Although we did not discuss it explicitly, much of the kinetic theory of gases is contained in this chapter. Problems 7–13 through 7–25 develop some of the kinetic theory of gases.

7–13. Consider a system of N interacting molecules, whose vibrational degrees of freedom are treated quantum mechanically and whose translational and rotational degrees of freedom are treated classically with Hamiltonian

$$H_{\text{class}} = K_{\text{trans}} + K_{\text{rot}} + U$$

where K represents kinetic energy, and U represents potential energy. Substitute this into Eq. (7–19); integrate over all the coordinates except the $3N$ translational momentum coordinates; and derive

$$\text{prob}\{K_{\text{trans}}\} = \frac{e^{-K_{\text{trans}}/kT}\, dp_{\text{trans}}}{\int e^{-K_{\text{trans}}/kT}\, dp_{\text{trans}}}$$

Now realize that

$$K_{\text{trans}} = \sum_{j=1}^{N} \frac{1}{2m} (p_{xj}^2 + p_{yj}^2 + p_{zj}^2)$$

and derive the normalized Maxwell-Boltzmann distribution, namely,

$$f(p_x, p_y, p_z)\, dp_x\, dp_y\, dp_z = (2\pi mkT)^{-3/2}\, e^{-(p_x^2 + p_y^2 + p_z^2)/2mkT}\, dp_x\, dp_y\, dp_z \qquad (7\text{–}40)$$

One can derive all of the usual expressions of the kinetic theory of gases from this.

7–14. An integral that appears often in statistical mechanics and particularly in the kinetic theory of gases is

$$I_n = \int_0^\infty x^n e^{-ax^2}\, dx$$

This integral can be readily generated from two basic integrals. For even values of n, we first consider

$$I_0 = \int_0^\infty e^{-ax^2}\,dx$$

The standard trick to evaluate this integral is to square it, and then transform the variables into polar coordinates:

$$\begin{aligned} I_0^2 &= \int_0^\infty \int_0^\infty e^{-ax^2} e^{-ay^2}\,dx\,dy \\ &= \int_0^\infty \int_0^{\pi/2} e^{-ar^2} r\,dr\,d\theta \\ &= \frac{\pi}{4a} \end{aligned}$$

$$I_0 = \frac{1}{2}\left(\frac{\pi}{a}\right)^{1/2}$$

Using this result, show that for even n

$$I_n = \frac{1 \cdot 3 \cdot 5 \cdots (n-1)}{2(2a)^{n/2}}\left(\frac{\pi}{a}\right)^{1/2} \qquad n \text{ even}$$

For odd values of n, the basic integral I_1 is easy. Using I_1, show that

$$I_n = \frac{\Gamma\left(\frac{n+1}{2}\right)}{2a^{(n+1)/2}} \qquad n \text{ odd}$$

7–15. Convert Eq. (7–40) (see Problem 7–13) from a Cartesian coordinate to a spherical coordinate representation by writing

$$\begin{aligned} &p^2 = p_x^2 + p_y^2 + p_z^2 \\ &p_z = p \cos\theta \\ &p_x = p \sin\theta \cos\phi \\ &p_y = p \sin\theta \sin\phi \\ &dp_x\,dp_y\,dp_z \rightarrow p^2 \sin\theta\,dp\,d\theta\,d\phi \end{aligned}$$

and integrating over θ and ϕ to get

$$f(p)\,dp = 4\pi(2\pi mkT)^{-3/2} p^2 e^{-p^2/2mkT}\,dp$$

for the fraction of molecules with momentum between p and $p + dp$. By substituting $p = mv$, we get the fraction of molecules with *speeds* between v and $v + dv$:

$$f(v)\,dv = 4\pi\left(\frac{m}{2\pi kT}\right)^{3/2} v^2 e^{-mv^2/2kT}\,dv$$

7–16. Prove that the most probable molecular speed is $v_{mp} = (2kT/m)^{1/2}$, that the mean speed is $\langle v\rangle = (8kT/\pi m)^{1/2}$, and that the root-mean-square speed is $\langle v^2\rangle^{1/2} = (3kT/m)^{1/2}$. Evaluate these for H_2 and N_2 at 25°C.

7–17. Show that the mean-square fluctuation of the velocity of the Maxwell-Boltzmann distribution is

$$\overline{v^2} - \bar{v}^2 = \frac{kT}{m}\left(3 - \frac{8}{\pi}\right)$$

7–18. Show that the average velocity in any direction (say x, y, or z) vanishes. What does this mean?

7–19. Derive an expression for the fraction of molecules with translational energy between ε and $\varepsilon + d\varepsilon$ from both Eqs. (7–40) and (5–7).

7–20. According to Problem 1–35, the speed of sound in an ideal gas is given by

$$c_0 = \left(\gamma \frac{RT}{M}\right)^{1/2}$$

where M is the molecular weight of the gas, and $\gamma = C_p/C_V$. Show that $c_0 = 0.81\bar{v}$ for an ideal monatomic gas.

7–21. Calculate the probability that two molecules will have a total kinetic energy between ε and $\varepsilon + d\varepsilon$.

7–22. Calculate the fraction of molecules with x-component of velocity between $\pm n(2kT/m)^{1/2}$, where $n = 1$, 2, and 3. Remember that the integral of $\exp(-x^2)$ with finite limits cannot be evaluated in closed form and is expressed in terms of the error function erf (x) by

$$\text{erf}\,(x) = \frac{2}{\pi^{1/2}} \int_0^x e^{-t^2}\, dt$$

7–23. What is the average kinetic energy $\bar{\varepsilon}$ and the most probable kinetic energy ε_{mp} of a gas molecule?

7–24. Show that the number of molecules striking a unit area per unit time is $\rho\bar{v}/4$, where $\rho = N/V$.

7–25. How would you interpret the velocity distribution

$$\phi(\mathbf{v}) = \left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left\{-\frac{m}{2kT}\,[(v_x - a)^2 + (v_y - b)^2 + (v_z - c)^2]\right\}$$

in which a, b, and c are constants?

7–26. The relativistic dependence of the kinetic energy on momentum is

$$\varepsilon = c(p_x^2 + p_y^2 + p_z^2 + m_0^2c^2)^{1/2}$$

where m_0 is the rest mass of the particle, and c is the speed of light. Determine the thermodynamic properties of an ideal gas in the extreme relativistic limit, where $p \gg m_0 c$.

7–27. If an atom is radiating light of wavelength λ_0, the wavelength measured by an observer will be

$$\lambda = \lambda_0\left(1 + \frac{v_z}{c}\right)$$

if moving away from or toward the observer with velocity v_z. In this equation c is the speed of light. This is known as the Doppler effect. If one observes the radiation emitted from a gas at temperature T, it is found that the line at λ_0 will be spread out by the Maxwellian distribution of velocities v_z of the molecules emitting the radiation. Show that $I(\lambda)\, d\lambda$, the intensity of radiation observed between wavelengths λ and $\lambda + d\lambda$, is

$$I(\lambda) \propto \exp\left\{-\frac{mc^2(\lambda - \lambda_0)^2}{2\lambda_0^2 kT}\right\}$$

This spreading about the line at λ_0 is known as Doppler broadening. Estimate the Doppler line width for HCl radiating microwave radiation at room temperature.

7–28. Plot C_V in Eq. (7–39) versus temperature and see that the vibrational contribution does not contribute until the temperature approaches Θ_V.

7–29. Prove that if the Hamiltonian is given by Eq. (7–37), then each of the quadratic terms will contribute $kT/2$ to the average molecular energy and $k/2$ to the molecular heat capacity.

7–30. Prove that even if the a_j and b_j in the Hamiltonian of Eq. (7–37) are functions of the momenta and coordinates not involved in the quadratic terms, the law of equipartition still applies. In particular, show how this more general version of the law of equipartition applies to the rigid rotor Hamiltonian.

7–31. Let $H(p, q)$ be the classical Hamiltonian for a classical system of N interacting particles. Let x_j be one of the $3N$ momentum components or one of the $3N$ spatial coordinates. Prove the generalized equipartition theorem, namely, that

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

and from this derive the principle of equipartition of energy that we discussed earlier. Hint: Realize that the potential $U \to \infty$ at the walls of the container.

7–32. Consider a two-dimensional harmonic oscillator with Hamiltonian

$$H = \frac{1}{2m}(p_x^2 + p_y^2) + \frac{k}{2}(x^2 + y^2)$$

According to the principle of equipartition of energy, the average energy will be $2kT$. Now transform this Hamiltonian to plane polar coordinates to get

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2}\right) + \frac{k}{2}r^2$$

What would you predict for the average energy now? Show by direct integration in plane polar coordinates that $\bar{\varepsilon} = 2kT$. Is anything wrong here? Why not?

7–33. Convince yourself that the number of phase points passing through a face perpendicular to q_1 per unit time is

$$f\dot{q}_1\, \delta q_2\, \delta q_3 \cdots \delta p_l$$