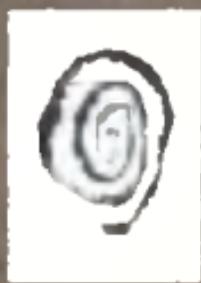


• EIF



FUNDAMENTALS OF STATISTICAL AND THERMAL PHYSICS

***McGraw-Hill series in fundamentals of physics:
an undergraduate textbook program***

E. U. CONDON, EDITOR

MEMBERS OF THE ADVISORY BOARD

*Arthur F. Kip, Department of Physics, University of California, Berkeley
Hugh D. Young, Department of Physics, Carnegie Institute of Technology*

INTRODUCTORY TEXTS

*Young · Fundamentals of Mechanics and Heat
Kip · Fundamentals of Electricity and Magnetism
Young · Fundamentals of Optics and Modern Physics*

UPPER-DIVISION TEXTS

Reif · Fundamentals of Statistical and Thermal Physics

*Fundamentals of statistical
and thermal physics*

F. REIF

Professor of Physics

University of California, Berkeley

McGRAW-HILL BOOK COMPANY

*New York San Francisco St. Louis
Toronto London Sydney*

Fundamentals of statistical and thermal physics

Copyright © 1965 by McGraw-Hill, Inc. All rights reserved.
Printed in the United States of America. This book, or parts
thereof, may not be reproduced in any form without per-
mission of the publishers. Library of Congress Catalog Card
Number 63-22730.

Preface

THIS BOOK is devoted to a discussion of some of the basic physical concepts and methods appropriate for the description of systems involving very many particles. It is intended, in particular, to present the disciplines of thermodynamics, statistical mechanics, and kinetic theory from a unified and modern point of view. Accordingly, the presentation departs from the historical development in which thermodynamics was the first of these disciplines to arise as an independent subject. The history of the ideas concerned with heat, work, and the kinetic theory of matter is interesting and instructive, but it does not represent the clearest or most illuminating way of developing these subjects. I have therefore abandoned the historical approach in favor of one that emphasizes the essential unity of the subject matter and seeks to develop physical insight by stressing the microscopic content of the theory.

Atoms and molecules are constructs so successfully established in modern science that a nineteenth-century distrust of them seems both obsolete and inappropriate. For this reason I have deliberately chosen to base the entire discussion on the premise that all macroscopic systems consist ultimately of atoms obeying the laws of quantum mechanics. A combination of these microscopic concepts with some statistical postulates then leads readily to some very general conclusions on a purely *macroscopic* level of description. These conclusions are valid *irrespective* of any particular models that might be assumed about the nature or interactions of the particles in the systems under consideration; they possess, therefore, the full generality of the classical laws of thermodynamics. Indeed, they are more general, since they make clear that the macroscopic parameters of a system are statistical in nature and exhibit fluctuations which are calculable and observable under appropriate conditions. Despite the microscopic point of departure, the book thus contains much general reasoning on a purely macroscopic level—probably about as much as a text on classical thermodynamics—but the microscopic content of the macroscopic arguments remains clear at all stages. Furthermore, if one is willing to adopt specific microscopic models concerning the particles consti-

tuting a system, then it is also apparent how one can calculate macroscopic quantities on the basis of this microscopic information. Finally, the statistical concepts used to discuss equilibrium situations constitute an appropriate preparation for their extension to the discussion of systems which are not in equilibrium.

This approach has, in my own teaching experience, proved to be no more difficult than the customary one which begins with classical thermodynamics. The latter subject, developed along purely macroscopic lines, is conceptually far from easy. Its reasoning is often delicate and of a type which seems unnatural to many physics students, and the significance of the fundamental concept of entropy is very hard to grasp. I have chosen to forego the subtleties of traditional arguments based on cleverly chosen cycles and to substitute instead the task of assimilating some elementary statistical ideas. The following gains are thereby achieved: (a) Instead of spending much time discussing various arguments based on heat engines, one can introduce the student at an early stage to statistical methods which are of great and recurring importance throughout all of physics. (b) The microscopic approach yields much better physical insight into many phenomena and leads to a ready appreciation of the meaning of entropy. (c) Much of modern physics is concerned with the explanation of macroscopic phenomena in terms of microscopic concepts. It seems useful, therefore, to follow a presentation which stresses at all times the interrelation between microscopic and macroscopic levels of description. The traditional teaching of thermodynamics and statistical mechanics as distinct subjects has often left students with their knowledge compartmentalized and has also left them ill-prepared to accept newer ideas, such as spin temperature or negative temperature, as legitimate and natural. (d) Since a unified presentation is more economical, conceptually as well as in terms of time, it permits one to discuss more material and some more modern topics.

The basic plan of the book is the following: The first chapter is designed to introduce some basic probability concepts. Statistical ideas are then applied to systems of particles in equilibrium so as to develop the basic notions of statistical mechanics and to derive therefrom the purely macroscopic general statements of thermodynamics. The *macroscopic* aspects of the theory are then discussed and illustrated at some length; the same is then done for the *microscopic* aspects of the theory. Some more complicated equilibrium situations, such as phase transformations and quantum gases, are taken up next. At this point the text turns to a discussion of nonequilibrium situations and treats transport theory in dilute gases at varying levels of sophistication. Finally, the last chapter deals with some general questions involving irreversible processes and fluctuations. Several appendices contain mostly various useful mathematical results.

The book is intended chiefly as a text for an introductory course in statistical and thermal physics for college juniors or seniors. The mimeographed notes on which it is based have been used in this way for more than two years by myself and several of my colleagues in teaching such a course at the University of California in Berkeley. No prior knowledge of heat or thermo-

dynamics is presupposed; the necessary prerequisites are only the equivalents of a course in introductory physics and of an elementary course in atomic physics. The latter course is merely supposed to have given the student sufficient background in modern physics (*a*) to know that quantum mechanics describes systems in terms of quantum states and wave functions, (*b*) to have encountered the energy levels of a simple harmonic oscillator and to have seen the quantum description of a free particle in a box, and (*c*) to have heard of the Heisenberg uncertainty and Pauli exclusion principles. These are essentially all the quantum ideas that are needed.

The material included here is more than can be covered in a one-semester undergraduate course. This was done purposely (*a*) to include a discussion of those basic ideas likely to be most helpful in facilitating the student's later access to more advanced works, (*b*) to allow students with some curiosity to read beyond the minimum on a given topic, (*c*) to give the instructor some possibility of selecting between alternate topics, and (*d*) to anticipate current revisions of the introductory physics course curriculum which should make upper-division students in the near future much more sophisticated and better prepared to handle advanced material than they are now. In actual practice I have successfully covered the first 12 chapters (omitting Chapter 10 and most starred sections) in a one-semester course. Chapter 1 contains a discussion of probability concepts more extensive than is needed for the understanding of subsequent chapters. In addition, the chapters are arranged in such a way that it is readily possible, after the first eight chapters, to omit some chapters in favor of others without encountering difficulties with prerequisites.

The book should also be suitable for use in an introductory graduate course if one includes the starred sections and the last three chapters, which contain somewhat more advanced material. Indeed, with students who have studied classical thermodynamics but have had no significant exposure to the ideas of statistical mechanics in their undergraduate career, one cannot hope to cover in a one-semester graduate course appreciably more subject matter than is treated here. One of my colleagues has thus used the material in our Berkeley graduate course on statistical mechanics (a course which is, as yet, mostly populated by students with this kind of preparation).

Throughout the book I have tried to keep the approach well-motivated and to strive for simplicity of presentation. It has not been my aim to pursue rigor in the formal mathematical sense. I have, however, attempted to keep the basic physical ideas in the forefront and to discuss them with care. In the process the book has become longer than it might have otherwise, for I have not hesitated to increase the ratio of words to formulas, to give illustrative examples, or to present several ways of looking at a question whenever I felt that it would enhance understanding. My aim has been to stress physical insight and important methods of reasoning, and I advise most earnestly that the student stress these aspects of the subject instead of trying to memorize various formulas meaningless in themselves. To avoid losing the reader in irrelevant details, I have often refrained from presenting the most general case of a problem and have sought instead to treat relatively simple cases by power-

ful and easily generalizable methods. The book is not meant to be encyclopaedic; it is merely intended to provide a basic skeleton of some fundamental ideas most likely to be useful to the student in his future work. Needless to say, some choices had to be made. For example, I thought it important to introduce the Boltzmann equation, but resisted the temptation of discussing applications of the Onsager relations to various irreversible phenomena such as thermoelectric effects.

It is helpful if a reader can distinguish material of secondary importance from that which is essential to the main thread of the argument. Two devices have been used to indicate subject matter of subsidiary importance: (a) Sections marked by a star (asterisk) contain material which is more advanced or more detailed; they can be omitted (and probably should be omitted in a first reading) without incurring a handicap in proceeding to subsequent sections. (b) Many remarks, examples, and elaborations are interspersed throughout the text and are set off on a gray background. Conversely, black marginal pointers have been used to emphasize important results and to facilitate reference to them.

The book contains about 230 problems, which should be regarded as an essential part of the text. It is indispensable that the student solve an appreciable fraction of these problems if he is to gain a meaningful understanding of the subject matter and not merely a casual hearsay acquaintance with it.

I am indebted to several of my colleagues for many valuable criticisms and suggestions. In particular, I should like to thank Prof. Eyyvind H. Wichmann, who read an older version of the entire manuscript with meticulous care, Prof. Owen Chamberlain, Prof. John J. Hopfield, Dr. Allan N. Kaufman, and Dr. John M. Worlock. Needless to say, none of these people should be blamed for the flaws of the final product.

Acknowledgements are also due to Mr. Roger F. Knacke for providing the answers to the problems. Finally, I am particularly grateful to my secretary, Miss Beverly West, without whose devotion and uncanny ability to transform pages of utterly illegible handwriting into a perfectly typed technical manuscript this book could never have been written.

It has been said that "an author never finishes a book, he merely abandons it." I have come to appreciate vividly the truth of this statement and dread to see the day when, looking at the manuscript in print, I am sure to realize that so many things could have been done better and explained more clearly. If I abandon the book nevertheless, it is in the modest hope that it may be useful to others despite its shortcomings.

Contents

Preface *vii*

1 *Introduction to statistical methods* *1*

RANDOM WALK AND BINOMIAL DISTRIBUTION

1·1	<i>Elementary statistical concepts and examples</i>	4
1·2	<i>The simple random walk problem in one dimension</i>	7
1·3	<i>General discussion of mean values</i>	11
1·4	<i>Calculation of mean values for the random walk problem</i>	13
1·5	<i>Probability distribution for large N</i>	17
1·6	<i>Gaussian probability distributions</i>	21

GENERAL DISCUSSION OF THE RANDOM WALK

1·7	<i>Probability distributions involving several variables</i>	25
1·8	<i>Comments on continuous probability distributions</i>	27
1·9	<i>General calculation of mean values for the random walk</i>	32
*1·10	<i>Calculation of the probability distribution</i>	35
*1·11	<i>Probability distribution for large N</i>	37

2 *Statistical description of systems of particles* *47*

STATISTICAL FORMULATION OF THE MECHANICAL PROBLEM

2·1	<i>Specification of the state of a system</i>	48
2·2	<i>Statistical ensemble</i>	52
2·3	<i>Basic postulates</i>	53
2·4	<i>Probability calculations</i>	60
2·5	<i>Behavior of the density of states</i>	61

INTERACTION BETWEEN MACROSCOPIC SYSTEMS

- 2·6 *Thermal interaction* 66
- 2·7 *Mechanical interaction* 68
- 2·8 *General interaction* 73
- 2·9 *Quasi-static processes* 74
- 2·10 *Quasi-static work done by pressure* 76
- 2·11 *Exact and “inexact” differentials* 78

3 Statistical thermodynamics 87

IRREVERSIBILITY AND THE ATTAINMENT OF EQUILIBRIUM

- 3·1 *Equilibrium conditions and constraints* 87
- 3·2 *Reversible and irreversible processes* 91

THERMAL INTERACTION BETWEEN MACROSCOPIC SYSTEMS

- 3·3 *Distribution of energy between systems in equilibrium* 94
- 3·4 *The approach to thermal equilibrium* 100
- 3·5 *Temperature* 102
- 3·6 *Heat reservoirs* 106
- 3·7 *Sharpness of the probability distribution* 108

GENERAL INTERACTION BETWEEN MACROSCOPIC SYSTEMS

- 3·8 *Dependence of the density of states on the external parameters* 112
- 3·9 *Equilibrium between interacting systems* 114
- 3·10 *Properties of the entropy* 117

SUMMARY OF FUNDAMENTAL RESULTS

- 3·11 *Thermodynamic laws and basic statistical relations* 122
- 3·12 *Statistical calculation of thermodynamic quantities* 124

4 Macroscopic parameters and their measurement 128

- 4·1 *Work and internal energy* 128
- 4·2 *Heat* 131
- 4·3 *Absolute temperature* 133
- 4·4 *Heat capacity and specific heat* 139
- 4·5 *Entropy* 142
- 4·6 *Consequences of the absolute definition of entropy* 145
- 4·7 *Extensive and intensive parameters* 148

5 Simple applications of macroscopic thermodynamics 152

PROPERTIES OF IDEAL GASES

- 5·1 *Equation of state and internal energy* 153
- 5·2 *Specific heats* 156

5·3	<i>Adiabatic expansion or compression</i>	158,
5·4	<i>Entropy</i>	160

GENERAL RELATIONS FOR A HOMOGENEOUS SUBSTANCE

5·5	<i>Derivation of general relations</i>	161
5·6	<i>Summary of Maxwell relations and thermodynamic functions</i>	164
5·7	<i>Specific heats</i>	166
5·8	<i>Entropy and internal energy</i>	171

FREE EXPANSION AND THROTTLING PROCESSES

5·9	<i>Free expansion of a gas</i>	175
5·10	<i>Throttling (or Joule-Thomson) process</i>	178

HEAT ENGINES AND REFRIGERATORS

5·11	<i>Heat engines</i>	184
5·12	<i>Refrigerators</i>	190

6 Basic methods and results of statistical mechanics 201

ENSEMBLES REPRESENTATIVE OF SITUATIONS OF PHYSICAL INTEREST

6·1	<i>Isolated system</i>	201
6·2	<i>System in contact with a heat reservoir</i>	202
6·3	<i>Simple applications of the canonical distribution</i>	206
6·4	<i>System with specified mean energy</i>	211
6·5	<i>Calculation of mean values in a canonical ensemble</i>	212
6·6	<i>Connection with thermodynamics</i>	214

APPROXIMATION METHODS

6·7	<i>Ensembles used as approximations</i>	219
*6·8	<i>Mathematical approximation methods</i>	221

GENERALIZATIONS AND ALTERNATIVE APPROACHES

*6·9	<i>Grand canonical and other ensembles</i>	225
*6·10	<i>Alternative derivation of the canonical distribution</i>	229

7 Simple applications of statistical mechanics 237

GENERAL METHOD OF APPROACH

7·1	<i>Partition functions and their properties</i>	237
-----	---	-----

IDEAL MONATOMIC GAS

7·2	<i>Calculation of thermodynamic quantities</i>	239
7·3	<i>Gibbs paradox</i>	243
7·4	<i>Validity of the classical approximation</i>	246

THE EQUIPARTITION THEOREM

- 7·5 *Proof of the theorem* 248
7·6 *Simple applications* 250
7·7 *Specific heats of solids* 253

PARAMAGNETISM

- 7·8 *General calculation of magnetization* 257

KINETIC THEORY OF DILUTE GASES IN EQUILIBRIUM

- 7·9 *Maxwell velocity distribution* 262
7·10 *Related velocity distributions and mean values* 265
7·11 *Number of molecules striking a surface* 269
7·12 *Effusion* 273
7·13 *Pressure and momentum transfer* 278

8 Equilibrium between phases or chemical species 288**GENERAL EQUILIBRIUM CONDITIONS**

- 8·1 *Isolated system* 289
8·2 *System in contact with a reservoir at constant temperature* 291
8·3 *System in contact with a reservoir at constant temperature and pressure* 294
8·4 *Stability conditions for a homogeneous substance* 296

EQUILIBRIUM BETWEEN PHASES

- 8·5 *Equilibrium conditions and the Clausius-Clapeyron equation* 301
8·6 *Phase transformations and the equation of state* 306

SYSTEMS WITH SEVERAL COMPONENTS; CHEMICAL EQUILIBRIUM

- 8·7 *General relations for a system with several components* 312
8·8 *Alternative discussion of equilibrium between phases* 315
8·9 *General conditions for chemical equilibrium* 317
8·10 *Chemical equilibrium between ideal gases* 319

9 Quantum statistics of ideal gases 331**MAXWELL-BOLTZMANN, BOSE-EINSTEIN, AND FERMI-DIRAC STATISTICS**

- 9·1 *Identical particles and symmetry requirements* 331
9·2 *Formulation of the statistical problem* 335
9·3 *The quantum distribution functions* 338
9·4 *Maxwell-Boltzmann statistics* 343
9·5 *Photon statistics* 345
9·6 *Bose-Einstein statistics* 346
9·7 *Fermi-Dirac statistics* 350
9·8 *Quantum statistics in the classical limit* 351

IDEAL GAS IN THE CLASSICAL LIMIT

- 9·9 *Quantum states of a single particle* 353
9·10 *Evaluation of the partition function* 360
9·11 *Physical implications of the quantum-mechanical enumeration of states* 363
*9·12 *Partition functions of polyatomic molecules* 367

BLACK-BODY RADIATION

- 9·13 *Electromagnetic radiation in thermal equilibrium inside an enclosure* 373
9·14 *Nature of the radiation inside an arbitrary enclosure* 378
9·15 *Radiation emitted by a body at temperature T* 381

CONDUCTION ELECTRONS IN METALS

- 9·16 *Consequences of the Fermi-Dirac distribution* 388
*9·17 *Quantitative calculation of the electronic specific heat* 393

10 Systems of interacting particles

404

SOLIDS

- 10·1 *Lattice vibrations and normal modes* 407
10·2 *Debye approximation* 411

NONIDEAL CLASSICAL GAS

- 10·3 *Calculation of the partition function for low densities* 418
10·4 *Equation of state and virial coefficients* 422
10·5 *Alternative derivation of the van der Waals equation* 426

FERROMAGNETISM

- 10·6 *Interaction between spins* 428
10·7 *Weiss molecular-field approximation* 430

11 Magnetism and low temperatures

438

- 11·1 *Magnetic work* 439
11·2 *Magnetic cooling* 445
11·3 *Measurement of very low absolute temperatures* 452
11·4 *Superconductivity* 455

12 Elementary kinetic theory of transport processes

461

- 12·1 *Collision time* 463
12·2 *Collision time and scattering cross section* 467

12·3	<i>Viscosity</i>	471
12·4	<i>Thermal conductivity</i>	478
12·5	<i>Self-diffusion</i>	483
12·6	<i>Electrical Conductivity</i>	488

13 *Transport theory using the relaxation time approximation* 494

13·1	<i>Transport processes and distribution functions</i>	494
13·2	<i>Boltzmann equation in the absence of collisions</i>	498
13·3	<i>Path integral formulation</i>	502
13·4	<i>Example: calculation of electrical conductivity</i>	504
13·5	<i>Example: calculation of viscosity</i>	507
13·6	<i>Boltzmann differential equation formulation</i>	508
13·7	<i>Equivalence of the two formulations</i>	510
13·8	<i>Examples of the Boltzmann equation method</i>	511

14 *Near-exact formulation of transport theory* 516

14·1	<i>Description of two-particle collisions</i>	516
14·2	<i>Scattering cross sections and symmetry properties</i>	520
14·3	<i>Derivation of the Boltzmann equation</i>	523
14·4	<i>Equation of change for mean values</i>	525
14·5	<i>Conservation equations and hydrodynamics</i>	529
14·6	<i>Example: simple discussion of electrical conductivity</i>	531
14·7	<i>Approximation methods for solving the Boltzmann equation</i>	534
14·8	<i>Example: calculation of the coefficient of viscosity</i>	539

15 *Irreversible processes and fluctuations* 548

TRANSITION PROBABILITIES AND MASTER EQUATION

15·1	<i>Isolated system</i>	548
15·2	<i>System in contact with a heat reservoir</i>	551
15·3	<i>Magnetic resonance</i>	553
15·4	<i>Dynamic nuclear polarization; Overhauser effect</i>	556

SIMPLE DISCUSSION OF BROWNIAN MOTION

15·5	<i>Langevin equation</i>	560
15·6	<i>Calculation of the mean-square displacement</i>	565

DETAILED ANALYSIS OF BROWNIAN MOTION

15·7	<i>Relation between dissipation and the fluctuating force</i>	567
------	---	-----

15·8	<i>Correlation functions and the friction constant</i>	570
*15·9	<i>Calculation of the mean-square velocity increment</i>	574
*15·10	<i>Velocity correlation function and mean-square displacement</i>	575

CALCULATION OF PROBABILITY DISTRIBUTIONS

*15·11	<i>The Fokker-Planck equation</i>	577
*15·12	<i>Solution of the Fokker-Planck equation</i>	580

FOURIER ANALYSIS OF RANDOM FUNCTIONS

15·13	<i>Fourier analysis</i>	582
15·14	<i>Ensemble and time averages</i>	583
15·15	<i>Wiener-Khintchine relations</i>	585
15·16	<i>Nyquist's theorem</i>	587
15·17	<i>Nyquist's theorem and equilibrium conditions</i>	589

GENERAL DISCUSSION OF IRREVERSIBLE PROCESSES

15·18	<i>Fluctuations and Onsager relations</i>	594
-------	---	-----

Appendices

605

A.1	<i>Review of elementary sums</i>	605
A.2	<i>Evaluation of the integral</i> $\int_{-\infty}^{\infty} e^{-x^2} dx$	606
A.3	<i>Evaluation of the integral</i> $\int_0^{\infty} e^{-x} x^n dx$	607
A.4	<i>Evaluation of integrals of the form</i> $\int_0^{\infty} e^{-ax^2} x^n dx$	608
A.5	<i>The error function</i>	609
A.6	<i>Stirling's formula</i>	610
A.7	<i>The Dirac delta function</i>	614
A.8	<i>The inequality</i> $\ln x \leq x - 1$	618
A.9	<i>Relations between partial derivatives of several variables</i>	619
A.10	<i>The method of Lagrange multipliers</i>	620
A.11	<i>Evaluation of the integral</i> $\int_0^{\infty} (e^x - 1)^{-1} x^3 dx$	622
A.12	<i>The H theorem and the approach to equilibrium</i>	624
A.13	<i>Liouville's theorem in classical mechanics</i>	626

Numerical constants

629

Bibliography

631

Answers to selected problems

637

Index

643

*Fundamentals of statistical
and thermal physics*

Introduction to statistical methods

THIS BOOK will be devoted to a discussion of systems consisting of very many particles. Examples are gases, liquids, solids, electromagnetic radiation (photons), etc. Indeed, most physical, chemical, or biological systems do consist of many molecules; our subject encompasses, therefore, a large part of nature.

The study of systems consisting of many particles is probably the most active area of modern physics research outside the realm of high-energy physics. In the latter domain, the challenge is to understand the fundamental interactions between nucleons, neutrinos, mesons, or other strange particles. But, in trying to discuss solids, liquids, plasmas, chemical or biological systems, and other such systems involving many particles, one faces a rather different task which is no less challenging. Here there are excellent reasons for supposing that the familiar laws of quantum mechanics describe adequately the motions of the atoms and molecules of these systems; furthermore, since the nuclei of atoms are not disrupted in ordinary chemical or biological processes and since gravitational forces between atoms are of negligible magnitude, the forces between the atoms in these systems involve only well-understood electromagnetic interactions. Somebody sufficiently sanguine might therefore be tempted to claim that these systems are "understood in principle." This would, however, be a rather empty and misleading statement. For, although it might be possible to write down the equations of motion for any one of these systems, the complexity of a system containing many particles is so great that it may make the task of deducing any useful consequences or predictions almost hopeless. The difficulties involved are not just questions of quantitative detail which can be solved by the brute force application of bigger and better computers. Indeed, even if the interactions between individual particles are rather simple, the sheer complexity arising from the interaction of a large number of them can often give rise to quite unexpected *qualitative* features in the behavior of a system. It may require very deep analysis to predict the occurrence of these features from a knowledge of the individual particles. For example, it is a striking fact, and one which is difficult to understand in microscopic detail,

that simple atoms forming a gas can condense abruptly to form a liquid with very different properties. It is a fantastically more difficult task to attain an understanding of how an assembly of certain kinds of molecules can lead to a system capable of biological growth and reproduction.

The task of understanding systems consisting of many particles is thus far from trivial, even when the interactions between individual atoms are well known. Nor is the problem just one of carrying out complicated computations. The main aim is, instead, to use one's knowledge of basic physical laws to develop new concepts which can illuminate the essential characteristics of such complex systems and thus provide sufficient insight to facilitate one's thinking, to recognize important relationships, and to make useful predictions. When the systems under consideration are not too complex and when the desired level of description is not too detailed, considerable progress can indeed be achieved by relatively simple methods of analysis.

It is useful to introduce a distinction between the sizes of systems whose description may be of interest. We shall call a system "microscopic" (i.e. "small scale") if it is roughly of atomic dimensions or smaller (say of the order of 10 Å or less). For example, the system might be a molecule. On the other hand, we shall call a system "macroscopic" (i.e., "large scale") when it is large enough to be visible in the ordinary sense (say greater than 1 micron, so that it can at least be observed with a microscope using ordinary light). The system consists then of very many atoms or molecules. For example, it might be a solid or liquid of the type we encounter in our daily experience. When one is dealing with such a macroscopic system, one is, in general, not concerned with the detailed behavior of each of the individual particles constituting the system. Instead, one is usually interested in certain macroscopic parameters which characterize the system as a whole, e.g., quantities like volume, pressure, magnetic moment, thermal conductivity, etc. If the macroscopic parameters of an isolated system do not vary in time, then one says that the system is in equilibrium. If an isolated system is not in equilibrium, the parameters of the system will, in general, change until they attain constant values corresponding to some final equilibrium condition. Equilibrium situations can clearly be expected to be amenable to simpler theoretical discussion than more general time-dependent nonequilibrium situations.

Macroscopic systems (like gases, liquids, or solids) began first to be systematically investigated from a macroscopic phenomenological point of view in the last century. The laws thus discovered formed the subject of "thermodynamics." In the second half of the last century the theory of the atomic constitution of all matter gained general acceptance, and macroscopic systems began to be analyzed from a fundamental microscopic point of view as systems consisting of very many atoms or molecules. The development of quantum mechanics after 1926 provided an adequate theory for the description of atoms and thus opened the way for an analysis of such systems on the basis of realistic microscopic concepts. In addition to the most modern methods of the "many-body problem," there have thus developed several disciplines of physics which deal with systems consisting of very many particles. Although

the boundaries between these disciplines are not very sharp, it may be useful to point out briefly the similarities and differences between their respective methods of approach.

a. For a system in *equilibrium*, one can try to make some very general statements concerning relationships existing between the *macroscopic* parameters of the system. This is the approach of classical "thermodynamics," historically the oldest discipline. The strength of this method is its great generality, which allows it to make valid statements based on a minimum number of postulates *without* requiring any detailed assumptions about the *microscopic* (i.e., molecular) properties of the system. The strength of the method also implies its weakness: only relatively few statements can be made on such general grounds, and many interesting properties of the system remain outside the scope of the method.

b. For a system in equilibrium, one can again try to make very general statements consistently based, however, on the *microscopic* properties of the particles in the system and on the laws of mechanics governing their behavior. This is the approach of "statistical mechanics." It yields *all* the results of thermodynamics *plus* a large number of general relations for calculating the macroscopic parameters of the system from a knowledge of its microscopic constituents. This method is one of great beauty and power.

c. If the system is *not* in equilibrium, one faces a much more difficult task. One can still attempt to make very general statements about such systems, and this leads to the methods of "irreversible thermodynamics," or, more generally, to the study of "statistical mechanics of irreversible processes." But the generality and power of these methods is much more limited than in the case of systems in equilibrium.

d. One can attempt to study *in detail* the interactions of all the particles in the system and thus to calculate parameters of macroscopic significance. This is the method of "kinetic theory." It is in principle always applicable, even when the system is *not* in equilibrium so that the powerful methods of equilibrium statistical mechanics are not available. Although kinetic theory yields the most detailed description, it is by the same token also the most difficult method to apply. Furthermore, its detailed point of view may tend to obscure general relationships which are of wider applicability.

Historically, the subject of thermodynamics arose first before the atomic nature of matter was understood. The idea that heat is a form of energy was first suggested by the work of Count Rumford (1798) and Davy (1799). It was stated explicitly by the German physician R. J. Mayer (1842) but gained acceptance only after the careful experimental work of Joule (1843–1849). The first analysis of heat engines was given by the French engineer S. Carnot in 1824. Thermodynamic theory was formulated in consistent form by Clausius and Lord Kelvin around 1850, and was greatly developed by J. W. Gibbs in some fundamental papers (1876–1878).

The atomic approach to macroscopic problems began with the study of the kinetic theory of dilute gases. This subject was developed through the

pioneering work of Clausius, Maxwell, and Boltzmann. Maxwell discovered the distribution law of molecular velocities in 1859, while Boltzmann formulated his fundamental integrodifferential equation (the Boltzmann equation) in 1872. The kinetic theory of gases achieved its modern form when Chapman and Enskog (1916–1917) succeeded in approaching the subject by developing systematic methods for solving this equation.

The more general discipline of statistical mechanics also grew out of the work of Boltzmann who, in 1872, succeeded further in giving a fundamental microscopic analysis of irreversibility and of the approach to equilibrium. The theory of statistical mechanics was then developed greatly in generality and power by the basic contributions of J. W. Gibbs (1902). Although the advent of quantum mechanics has brought many changes, the basic framework of the modern theory is still the one which he formulated.

In discussing systems consisting of very many particles, we shall not aim to recapitulate the historical development of the various disciplines dealing with the physical description of such systems. Instead we shall, from the outset, adopt a modern point of view based on our present-day knowledge of atomic physics and quantum mechanics. We already mentioned the fact that one can make very considerable progress in understanding systems of many particles by rather simple methods of analysis. This may seem rather surprising at first sight; for are not systems such as gases or liquids, which consist of a number of particles of the order of Avogadro's number (10^{23}), hopelessly complicated? The answer is that the very complexity of these systems contains within it the key to a successful method of attack. Since one is not concerned with the detailed behavior of each and every particle in such systems, it becomes possible to apply statistical arguments to them. But as every gambler, insurance agent, or other person concerned with the calculation of probabilities knows, statistical arguments become most satisfactory when they can be applied to large numbers. What a pleasure, then, to be able to apply them to cases where the numbers are as large as 10^{23} , i.e., Avogadro's number! In systems such as gases, liquids, or solids where one deals with very many identical particles, statistical arguments thus become particularly effective. This does not mean that all problems disappear; the physics of many-body problems does give rise to some difficult and fascinating questions. But many important problems do indeed become quite simple when approached by statistical means.

RANDOM WALK AND BINOMIAL DISTRIBUTION

1 · 1 *Elementary statistical concepts and examples*

The preceding comments make it apparent that statistical ideas will play a central role throughout this book. We shall therefore devote this first chapter to a discussion of some elementary aspects of probability theory which are of great and recurring usefulness.

The reader will be assumed to be familiar with the most rudimentary probability concepts. It is important to keep in mind that whenever it is desired to describe a situation from a statistical point of view (i.e., in terms of probabilities), it is always necessary to consider an assembly (or "ensemble") consisting of a very large number π of similarly prepared systems. The probability of occurrence of a particular event is then defined with respect to this particular ensemble and is given by the fraction of systems in the ensemble which are characterized by the occurrence of this specified event. For example, in throwing a pair of dice, one can give a statistical description by considering that a very large number π (in principle, $\pi \rightarrow \infty$) of similar pairs of dice are thrown under similar circumstances. (Alternatively one could imagine the same pair of dice thrown π times in succession under similar circumstances.) The probability of obtaining a double ace is then given by the fraction of these experiments in which a double ace is the outcome of a throw.

Note also that the probability depends very much on the nature of the ensemble which is contemplated in defining this probability. For example, it makes no sense to speak simply of the probability that an individual seed will yield red flowers. But one can meaningfully ask for the probability that such a seed, regarded as a member of an ensemble of similar seeds derived from a specified set of plants, will yield red flowers. The probability depends crucially on the ensemble of which the seed is regarded as a member. Thus the probability that a given seed will yield red flowers is in general different if this seed is regarded as (a) a member of a collection of similar seeds which are known to be derived from plants that had produced red flowers or as (b) a member of a collection of seeds which are known to be derived from plants that had produced pink flowers.

In the following discussion of basic probability concepts it will be useful to keep in mind a specific simple, but important, illustrative example—the so-called "random-walk problem." In its simplest idealized form, the problem can be formulated in the following traditional way: A drunk starts out from a lamppost located on a street. Each step he takes is of equal length l . The man is, however, so drunk that the direction of each step—whether it is to the right or to the left—is completely independent of the preceding step. All one can say is that each time the man takes a step, the probability of its being to the right is p , while the probability of its being to the left is $q = 1 - p$. (In the simplest case $p = q$, but in general $p \neq q$. For example, the street might be inclined with respect to the horizontal, so that a step downhill to the right is more likely than one uphill to the left.)

Choose the x axis to lie along the street so that $x = 0$ is the position of the

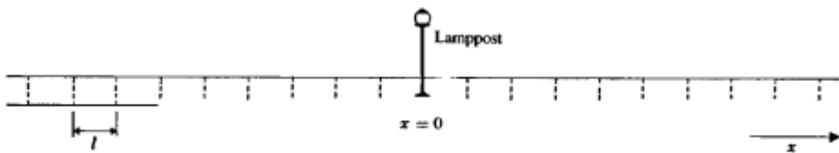


Fig. 1·1·1 *The drunkard's random walk in one dimension.*



Fig. 1·1·2 Example of a random walk in two dimensions.

point of origin, the lamppost. Clearly, since each step is of length l , the location of the man along the x axis must be of the form $x = ml$, where m is an integer (positive, negative, or zero). The question of interest is then the following: After the man has taken N steps, what is the probability of his being located at the position $x = ml$?

This statistical formulation of the problem again implies that one considers a very large number π of similar men starting from similar lampposts. (Alternatively, if the situation does not change in time, e.g., if the man does not gradually sober up, one could also repeat the same experiment π times with the same man.) At each step, one finds that a fraction p of the men moves to the right. One then asks what fraction of the men will be located at the position $x = ml$ after N steps.

One can readily generalize this one-dimensional problem to more dimensions whether it be two (a drunk starting out from a lamppost in the middle of a parking lot), three, or more. One again asks for the probability that after N steps the man is located at a given distance from the origin (although the distance is no longer of the form ml , with m integral).

Now the main concern of physics is not with drunks who stagger home from lampposts. But the problem which this illustrates is the one of adding N vectors of equal length but of random directions (or directions specified by some probability distribution) and then asking for the probability that their resultant vector sum has a certain magnitude and direction (see Fig. 1·1·2). We mention a few physical examples where this question is relevant.

a. Magnetism: An atom has a spin $\frac{1}{2}$ and a magnetic moment μ ; in accordance with quantum mechanics, its spin can therefore point either "up" or "down" with respect to a given direction. If both these possibilities are equally likely, what is the net total magnetic moment of N such atoms?

b. Diffusion of a molecule in a gas: A given molecule travels in three dimensions a mean distance l between collisions with other molecules. How far is it likely to have gone after N collisions?

c. Light intensity due to N incoherent light sources: The light amplitude due to each source can be represented by a two-dimensional vector whose direction specifies the phase of the disturbance. Here the phases are random, and the resultant amplitude, which determines the total intensity of the light from all the sources, must be computed by statistical means.

The random walk problem illustrates some very fundamental results of probability theory. The techniques used in the study of this problem are powerful and basic, and recur again and again throughout statistical physics. It is therefore very instructive to gain a good understanding of this problem.

1 · 2 *The simple random walk problem in one dimension*

For the sake of simplicity we shall discuss the random walk problem in one dimension. Instead of talking in terms of a drunk taking steps, let us revert to the less alcoholic vocabulary of physics and think of a particle performing successive steps, or displacements, in one dimension. After a total of N such steps, each of length l , the particle is located at

$$x = ml$$

where m is an integer lying between

$$-N \leq m \leq N$$

We want to calculate the probability $P_N(m)$ of finding the particle at the position $x = ml$ after N such steps.

Let n_1 denote the number of steps to the right and n_2 the corresponding number of steps to the left. Of course, the total number of steps N is simply

$$N = n_1 + n_2 \quad (1 \cdot 2 \cdot 1)$$

The net displacement (measured to the right in units of a step length) is given by

$$m = n_1 - n_2 \quad (1 \cdot 2 \cdot 2)$$

If it is known that in some sequence of N steps the particle has taken n_1 steps to the right, then its net displacement from the origin is determined. Indeed, the preceding relations immediately yield

$$m = n_1 - n_2 = n_1 - (N - n_1) = 2n_1 - N \quad (1 \cdot 2 \cdot 3)$$

This shows that if N is odd, the possible values of m must also be odd. Conversely, if N is even, m must also be even.

Our fundamental assumption was that successive steps are statistically

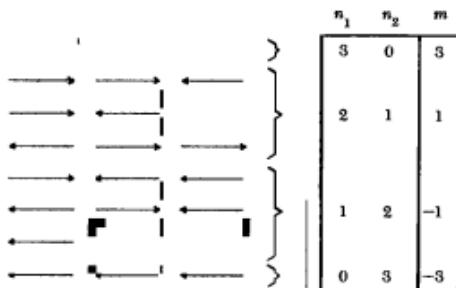


Fig. 1·2·1 Illustration showing the eight sequences of steps which are possible if the total number of steps is $N = 3$.

independent of each other. Thus one can assert simply that, irrespective of past history, each step is characterized by the respective probabilities

$$\begin{aligned} p &= \text{probability that the step is to the right} \\ \text{and} \quad q &= 1 - p = \text{probability that the step is to the left} \end{aligned}$$

Now, the probability of any *one* given sequence of n_1 steps to the right and n_2 steps to the left is given simply by multiplying the respective probabilities, i.e., by

$$\underbrace{p \cdot p \cdot \dots \cdot p}_{n_1 \text{ factors}} \cdot \underbrace{q \cdot q \cdot \dots \cdot q}_{n_2 \text{ factors}} = p^{n_1} q^{n_2} \quad (1 \cdot 2 \cdot 4)$$

But there are many different possible ways of taking N steps so that n_1 of them are to the right and n_2 are to the left (see illustration in Fig. 1·2·1). Indeed, the number of distinct possibilities (as shown below) is given by

$$\frac{N!}{n_1! n_2!} \quad (1 \cdot 2 \cdot 5)$$

Hence the probability $W_N(n_1)$ of taking (in a total of N steps) n_1 steps to the right and $n_2 = N - n_1$ steps to the left, in any order, is obtained by multiplying the probability (1·2·4) of this sequence by the number (1·2·5) of possible sequences of such steps. This gives

$$\blacktriangleright \quad W_N(n_1) = \frac{N!}{n_1! n_2!} p^{n_1} q^{n_2} \quad (1 \cdot 2 \cdot 6)$$

Simple example Consider the simple illustration of Fig. 1·2·1, which shows the case of a total of $N = 3$ steps. There is only one way in which all three successive steps can be to the right; the corresponding probability $W(3)$ that all three steps are to the right is then simply $p \cdot p \cdot p = p^3$. On the other hand, the probability of a sequence of steps where two steps are to the right

while the third step is to the left is p^2q . But there are three such possible sequences. Thus the total probability of occurrence of a situation where two steps are to the right and one is to the left is given by $3p^2q$.

Reasoning leading to Eq. (1·2·5) The problem is to count the number of distinct ways in which N objects (or steps), of which n_1 are indistinguishably of one type and n_2 of a second type, can be accommodated in a total of $N = n_1 + n_2$ possible places. In this case

the 1st place can be occupied by any one of the N objects
the 2nd place can be occupied by any one of the remaining $(N - 1)$ objects

the N th place can be occupied only by the last 1 object

Hence all the available places can be occupied in

$$N(N - 1)(N - 2) \cdots 1 = N!$$

possible ways. The above enumeration considers each object as distinguishable. But since the n_1 objects of the first type are indistinguishable (e.g., all are right steps), all the $n_1!$ permutations of these objects among themselves lead to the same situation. Similarly, all the $n_2!$ permutations of objects of the second type among themselves lead to the same situation. Hence, by dividing the total number $N!$ of arrangements of the objects by the number $n_1!n_2!$ of irrelevant permutations of objects of each type, one obtains the total number $N!/n_1!n_2!$ of distinct ways in which N objects can be arranged if n_1 are of one type and n_2 of another type.

Example For instance, in the previous example of three steps, there are $N = 3$ possible events (or places) designated in Fig. 1·2·2 by B_1, B_2, B_3 and capable of being filled by the three particular steps labeled A_1, A_2, A_3 . The event B_1 can then occur in any of three ways, B_2 in any of two ways, and B_3 in only one way. There are thus $3 \times 2 \times 1 = 3! = 6$ possible sequences

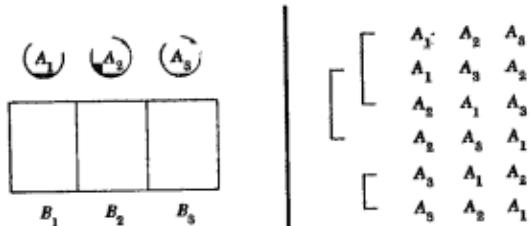


Fig. 1·2·2 Diagram illustrating the problem of distributing three objects A_1, A_2, A_3 among three places B_1, B_2, B_3 . The right part of the diagram lists the possible arrangements and indicates by brackets those arrangements which are identical when A_1 and A_2 are considered indistinguishable.

of these three steps. But suppose that A_1 and A_2 denote both right steps ($n_1 = 2$), while A_3 denotes a left step ($n_2 = 1$). Then sequences of steps differing only by the two permutations of A_1 and A_2 are really identical. Thus one is left with only $\frac{N}{2} = 3$ distinct sequences for which two steps are to the right and one step is to the left.

The probability function (1·2·6) is called the binomial distribution. The reason is that (1·2·5) represents a typical term encountered in expanding $(p + q)^N$ by the binomial theorem. Indeed, we recall that the binomial expansion is given by the formula

$$(p + q)^N = \sum_{n=0}^N \frac{N!}{n!(N-n)!} p^n q^{N-n} \quad (1\cdot2\cdot7)$$

We already pointed out in (1·2·3) that if it is known that the particle has performed n_1 steps to the right in a total of N steps, then its net displacement m from the origin is determined. Thus the probability $P_N(m)$ that the particle is found at position m after N steps is the same as $W_N(n_1)$ given by (1·2·6), i.e.,

$$P_N(m) = W_N(n_1) \quad (1\cdot2\cdot8)$$

By (1·2·1) and (1·2·2) one finds explicitly*

$$n_1 = \frac{1}{2}(N + m), \quad n_2 = \frac{1}{2}(N - m) \quad (1\cdot2\cdot9)$$

Substitution of these relations in (1·2·6) thus yields

$$P_N(m) = \frac{N!}{[(N+m)/2]![(N-m)/2]!} p^{(N+m)/2} (1-p)^{(N-m)/2} \quad (1\cdot2\cdot10)$$

In the special case where $p = q = \frac{1}{2}$ this assumes the symmetrical form

$$P_N(m) = \frac{N!}{[(N+m)/2]![(N-m)/2]!} \left(\frac{1}{2}\right)^N$$

Examples Suppose that $p = q = \frac{1}{2}$ and that $N = 3$ as illustrated in Fig. 1·2·1. Then the possible numbers of steps to the right are $n_1 = 0, 1, 2$, or 3 ; the corresponding displacements are $m = -3, -1, 1$, or 3 ; the corresponding probabilities are (as is clear from Fig. 1·2·1)

$$W_3(n_1) = P_3(m) = \frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8} \quad (1\cdot2\cdot11)$$

Figure (1·2·3) illustrates the binomial distribution for the same case where $p = q = \frac{1}{2}$, but with the total number of steps $N = 20$. The envelope of these discrete values of $P_N(m)$ is a bell-shaped curve. The physical sig-

* Note that (1·2·3) shows that $(N+m)$ and $(N-m)$ are even integers, since they equal $2n_1$ and $2n_2$, respectively.

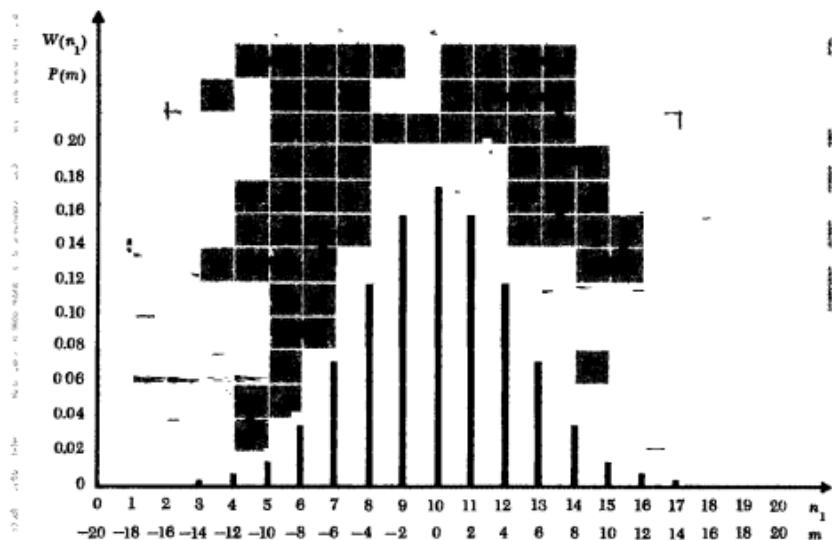


Fig. 1·2·3 Binomial probability distribution for $p = q = \frac{1}{2}$ when $N = 20$ steps. The graph shows the probability $W_N(n_1)$ of n_1 right steps, or equivalently the probability $P_N(m)$ of a net displacement of m units to the right.

nificance of this is obvious. After N random steps, the probability of the particle being a distance of N steps away from the origin is very small, while the probability of its being located in the vicinity of the origin is largest.

1·3 General discussion of mean values

Let u be a variable which can assume any of the M discrete values

$$u_1, u_2, \dots, u_M$$

with respective probabilities

$$P(u_1), P(u_2), \dots, P(u_M)$$

The mean (or average) value of u is denoted by \bar{u} and is defined by

$$\bar{u} = \frac{P(u_1)u_1 + P(u_2)u_2 + \dots + P(u_M)u_M}{P(u_1) + P(u_2) + \dots + P(u_M)}$$

or, in shorter notation, by

$$\bar{u} = \frac{\sum_{i=1}^M P(u_i)u_i}{\sum_{i=1}^M P(u_i)} \quad (1·3·1)$$

This is, of course, the familiar way of computing averages. For example, if u represents the grade of a student on an examination and $P(u)$ the number of students obtaining this grade, then Eq. (1·3·1) asserts that the mean grade is computed by multiplying each grade by the number of students having this grade, adding this up, and then dividing by the total number of students.

More generally, if $f(u)$ is any function of u , then the mean value of $f(u)$ is defined by

$$\overline{f(u)} \equiv \frac{\sum_{i=1}^M P(u_i)f(u_i)}{\sum_{i=1}^M P(u_i)} \quad (1\cdot3\cdot2)$$

This expression can be simplified. Since $P(u_i)$ is defined as a probability, the quantity

$$P(u_1) + P(u_2) + \dots + P(u_M) \equiv \sum_{i=1}^M P(u_i)$$

represents the probability that u assumes any one of its possible values and this must be unity. Hence one has quite generally

► $\sum_{i=1}^M P(u_i) = 1 \quad (1\cdot3\cdot3)$

This is the so-called "normalization condition" satisfied by every probability. As a result, the general definition (1·3·2) becomes

► $\overline{f(u)} \equiv \sum_{i=1}^M P(u_i)f(u_i) \quad (1\cdot3\cdot4)$

Note the following simple results. If $f(u)$ and $g(u)$ are any two functions of u , then

$$\overline{f(u) + g(u)} = \sum_{i=1}^M P(u_i)[f(u_i) + g(u_i)] = \sum_{i=1}^M P(u_i)f(u_i) + \sum_{i=1}^M P(u_i)g(u_i)$$

or

► $\overline{f(u) + g(u)} = \overline{f(u)} + \overline{g(u)} \quad (1\cdot3\cdot5)$

Furthermore, if c is any constant, it is clear that

► $\overline{cf(u)} = c\overline{f(u)} \quad (1\cdot3\cdot6)$

Some simple mean values are particularly useful for describing characteristic features of the probability distribution P . One of these is the mean value \bar{u} (e.g., the mean grade of a class of students). This is a measure of the central value of u about which the various values u_i are distributed. If one measures the values of u from their mean value \bar{u} , i.e., if one puts

$$\Delta u \equiv u - \bar{u} \quad (1\cdot3\cdot7)$$

then $\overline{\Delta u} = \overline{(u - \bar{u})} = \bar{u} - \bar{u} = 0 \quad (1\cdot3\cdot8)$

This says merely that the mean value of the deviation from the mean vanishes.

Another useful mean value is

$$\overline{(\Delta u)^2} \equiv \sum_{i=1}^M P(u_i)(u_i - \bar{u})^2 \geq 0 \quad (1 \cdot 3 \cdot 9)$$

which is called the “second moment of u about its mean,” or more simply the “dispersion of u .” This can never be negative, since $(\Delta u)^2 \geq 0$ so that each term in the sum contributes a nonnegative number. Only if $u_i = \bar{u}$ for all values u_i will the dispersion vanish. The larger the spread of values of u_i about \bar{u} , the larger the dispersion. The dispersion thus measures the amount of scatter of values of the variable about its mean value (e.g., scatter in grades about the mean grade of the students). Note the following general relation, which is often useful in computing the dispersion:

$$\overline{(u - \bar{u})^2} = \overline{(u^2 - 2u\bar{u} + \bar{u}^2)} = \bar{u}^2 - 2\bar{u}\bar{u} + \bar{u}^2$$

or

►
$$\overline{(u - \bar{u})^2} = \bar{u}^2 - \bar{u}^2 \quad (1 \cdot 3 \cdot 10)$$

Since the left side must be positive it also follows that

$$\bar{u}^2 \geq \bar{u}^2 \quad (1 \cdot 3 \cdot 11)$$

One can define further mean values such as $\overline{(\Delta u)^n}$, the “ n th moment of u about its mean,” for integers $n > 2$. These are, however, less commonly useful.

Note that a knowledge of $P(u)$ implies complete information about the actual distribution of values of the variable u . A knowledge of a few moments, like \bar{u} and $\overline{(\Delta u)^2}$, implies only partial, though useful, knowledge of the characteristics of this distribution. A knowledge of *some* mean values is *not* sufficient to determine $P(u)$ completely (unless one knows the moments $\overline{(\Delta u)^n}$ for *all* values of n). But by the same token it is often true that a calculation of the probability distribution function $P(u)$ may be quite difficult, whereas some simple mean values can be readily calculated directly without an explicit knowledge of $P(u)$. We shall illustrate some of these comments in the following pages.

1 · 4 Calculation of mean values for the random walk problem

In (1 · 2 · 6) we found that the probability, in a total of N steps, of making n_1 steps to the right (and $N - n_1 \equiv n_2$ steps to the left) is

$$W(n_1) = \frac{N!}{n_1!(N - n_1)!} p^{n_1} q^{N - n_1} \quad (1 \cdot 4 \cdot 1)$$

(For the sake of simplicity, we omit attaching the subscript N to W when no confusion is likely to arise.)

Let us first verify the normalization, i.e., the condition

$$\sum_{n_1=0}^N W(n_1) = 1 \quad (1 \cdot 4 \cdot 2)$$

which says that the probability of making any number of right steps between 0 and N must be unity. Substituting (1·4·1) into (1·4·2), we obtain

$$\begin{aligned} \sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} &= (p+q)^N \quad \text{by the binomial theorem} \\ &= 1^N = 1 \quad \text{since } q \equiv 1-p \end{aligned}$$

which verifies the result.

What is the mean number \bar{n}_1 of steps to the right? By definition

$$\bar{n}_1 \equiv \sum_{n_1=0}^N W(n_1) n_1 = \sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} n_1 \quad (1 \cdot 4 \cdot 3)$$

If it were not for that extra factor of n_1 in each term of the last sum, this would again be the binomial expansion and hence trivial to sum. The factor n_1 spoils this lovely situation. But there is a very useful general procedure for handling such an extra factor so as to reduce the sum to simpler form. Let us consider the purely mathematical problem of evaluating the sum occurring in (1·4·3), where p and q are considered to be any two *arbitrary* parameters. Then one observes that the extra factor n_1 can be produced by differentiation so that

$$n_1 p^{n_1} = p \frac{\partial}{\partial p} (p^{n_1})$$

Hence the sum of interest can be written in the form

$$\begin{aligned} \sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} n_1 &= \sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} \left[p \frac{\partial}{\partial p} (p^{n_1}) \right] q^{N-n_1} \\ &= p \frac{\partial}{\partial p} \left[\sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} \right] \quad \text{by interchanging order} \\ &\quad \text{of summation and differentiation} \\ &= p \frac{\partial}{\partial p} (p+q)^N \quad \text{by the binomial theorem} \\ &= pN(p+q)^{N-1} \end{aligned}$$

Since this result is true for arbitrary values of p and q , it must also be valid in our particular case of interest where p is some specified constant and $q \equiv 1 - p$. Then $p + q = 1$ so that (1·4·3) becomes simply

► $\bar{n}_1 = Np \quad (1 \cdot 4 \cdot 4)$

We could have guessed this result. Since p is the probability of making a right step, the mean number of right steps in a *total* of N steps is simply given

by $N \cdot p$. Clearly, the mean number of left steps is similarly equal to

$$\bar{n}_2 = Nq \quad (1 \cdot 4 \cdot 5)$$

Of course

$$\bar{n}_1 + \bar{n}_2 = N(p + q) = N$$

add up properly to the total number of steps.

The displacement (measured to the right in units of the step length l) is $m = n_1 - n_2$. Hence we get for the mean displacement

$$\blacktriangleright \quad \bar{m} = \overline{n_1 - n_2} = \bar{n}_1 - \bar{n}_2 = N(p - q) \quad (1 \cdot 4 \cdot 6)$$

If $p = q$, then $\bar{m} = 0$. This must be so since there is then complete symmetry between right and left directions.

Calculation of the dispersion Let us now calculate $(\overline{\Delta n_1})^2$. By (1 · 3 · 10) one has

$$\overline{(\Delta n_1)^2} \equiv \overline{(n_1 - \bar{n}_1)^2} = \overline{n_1^2} - \bar{n}_1^2 \quad (1 \cdot 4 \cdot 7)$$

We already know \bar{n}_1 . Thus we need to compute $\overline{n_1^2}$.

$$\begin{aligned} \overline{n_1^2} &\equiv \sum_{n_1=0}^N W(n_1) n_1^2 \\ &= \sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} n_1^2 \end{aligned} \quad (1 \cdot 4 \cdot 8)$$

Considering p and q as arbitrary parameters and using the same trick of differentiation as before, one can write

$$n_1^2 p^{n_1} = n_1 \left(p \frac{\partial}{\partial p} \right) (p^{n_1}) = \left(p \frac{\partial}{\partial p} \right)^2 (p^{n_1})$$

Hence the sum in (1 · 4 · 8) can be written in the form

$$\begin{aligned} \sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} \left(p \frac{\partial}{\partial p} \right)^2 p^{n_1} q^{N-n_1} &= \left(p \frac{\partial}{\partial p} \right)^2 \sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} && \text{by interchanging order of summation and differentiation} \\ &= \left(p \frac{\partial}{\partial p} \right)^2 (p+q)^N && \text{by the binomial theorem} \\ &= \left(p \frac{\partial}{\partial p} \right) [pN(p+q)^{N-1}] \\ &= p[N(p+q)^{N-1} + pN(N-1)(p+q)^{N-2}] \end{aligned}$$

The case of interest in (1·4·8) is that where $p + q = 1$. Thus (1·4·8) becomes simply

$$\begin{aligned}\overline{n_1^2} &= p[N + pN(N - 1)] \\ &= Np[1 + pN - p] \\ &= (Np)^2 + Npq \quad \text{since } 1 - p = q \\ &= \bar{n}_1^2 + Npq \quad \text{by (1·4·4)}\end{aligned}$$

Hence (1·4·7) gives for the dispersion of n_1 the result

► $\overline{(\Delta n_1)^2} = Npq \quad (1·4·9)$

The quantity $\overline{(\Delta n_1)^2}$ is quadratic in the displacement. Its square root, i.e., the rms (root-mean-square) deviation $\Delta^* n_1 \equiv [(\overline{(\Delta n_1)^2})^{1/2}]$, is a linear measure of the width of the range over which n_1 is distributed. A good measure of the *relative* width of this distribution is then

$$\frac{\Delta^* n_1}{\bar{n}_1} = \frac{\sqrt{Npq}}{Np} = \sqrt{\frac{p}{q}} \frac{1}{\sqrt{N}}$$

In particular,

$$\text{for } p = q = \frac{1}{2}, \quad \frac{\Delta^* n_1}{\bar{n}_1} = \frac{1}{\sqrt{N}}$$

Note that as N increases, the mean value \bar{n}_1 increases like N , but the width $\Delta^* n_1$ increases only like $N^{1/2}$. Hence the *relative* width $\Delta^* n_1/\bar{n}_1$ decreases with increasing N like $N^{-1/2}$.

One can also compute the dispersion of m , i.e., the dispersion of the net displacement to the right. By (1·2·3)

$$m = n_1 - n_2 = 2n_1 - N \quad (1·4·10)$$

Hence one obtains

$$\begin{aligned}\Delta m \equiv m - \bar{m} &= (2n_1 - N) - (2\bar{n}_1 - N) = 2(n_1 - \bar{n}_1) = 2\Delta n_1 \quad (1·4·11) \\ \text{and} \quad (\Delta m)^2 &= 4(\Delta n_1)^2\end{aligned}$$

Taking averages, one gets by (1·4·9)

► $\overline{(\Delta m)^2} = 4\overline{(\Delta n_1)^2} = 4Npq \quad (1·4·12)$

In particular,

$$\text{for } p = q = \frac{1}{2}, \quad \overline{(\Delta m)^2} = N$$

Example Consider the case of $N = 100$ steps, where $p = q = \frac{1}{2}$. Then the mean number \bar{n}_1 of steps to the right (or to the left) is 50; the mean displacement $\bar{m} = 0$. The root-mean-square displacement $[(\Delta m)^2]^{1/2} = 10$ steps.

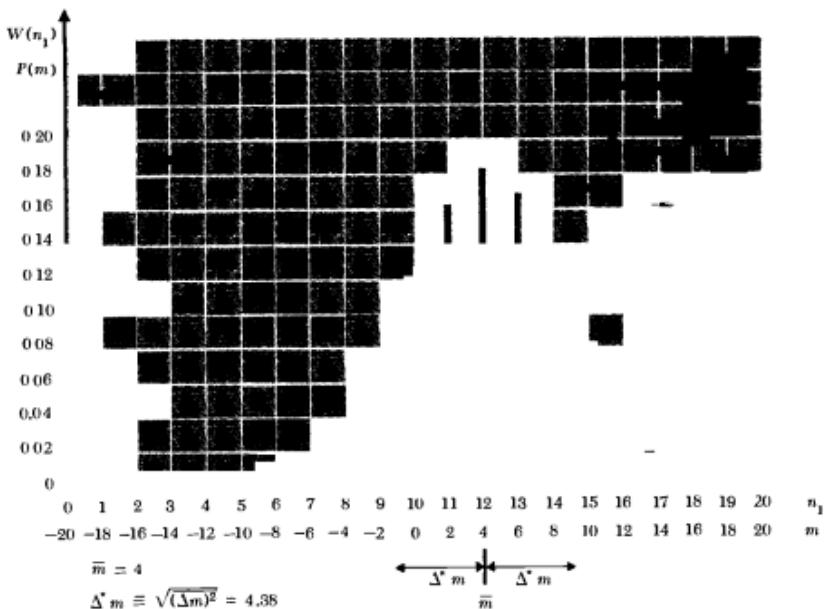


Fig. 1·4·1 Binomial probability distribution for $p = 0.6$ and $q = 0.4$, when $N = 20$ steps. The graph shows again the probability $W(n_1)$ of n_1 right steps, or equivalently, the probability $P(m)$ of a net displacement of m units to the right. The mean values \bar{m} and $(\Delta m)^2$ are also indicated.

1·5 Probability distribution for large N

When N is large, the binomial probability distribution $W(n_1)$ of (1·4·1) tends to exhibit a pronounced maximum at some value $n_1 = \bar{n}_1$, and to decrease rapidly as one goes away from \bar{n}_1 (see, for example, Fig. 1·4·1). Let us exploit this fact to find an approximate expression for $W(n_1)$ valid when N is sufficiently large.

If N is large and we consider regions near the maximum of W where n_1 is also large, the fractional change in W when n_1 changes by unity is relatively quite small, i.e.,

$$|W(n_1 + 1) - W(n_1)| \ll W(n_1) \quad (1·5·1)$$

Thus W can, to good approximation, be considered as a continuous function of the continuous variable n_1 , although only integral values of n_1 are of physical relevance. The location $n_1 = \bar{n}$ of the maximum of W is then approximately determined by the condition

$$\frac{dW}{dn_1} = 0 \quad \text{or equivalently} \quad \frac{d \ln W}{dn_1} = 0 \quad (1 \cdot 5 \cdot 2)$$

where the derivatives are evaluated for $n_1 = \bar{n}_1$. To investigate the behavior of $W(n_1)$ near its maximum, we shall put

$$n_1 = \bar{n}_1 + \eta \quad (1 \cdot 5 \cdot 3)$$

and expand $\ln W(n_1)$ in a Taylor's series about \bar{n}_1 . The reason for expanding $\ln W$, rather than W itself, is that $\ln W$ is a much more slowly varying function of n_1 than W . Thus the power series expansion for $\ln W$ should converge much more rapidly than the one for W .

An example may make this clearer. Suppose one wishes to find an approximate expression, valid for $y \ll 1$, of the function

$$f = (1 + y)^{-N}$$

where N is large. Direct expansion in Taylor's series (or by the binomial theorem) would give

$$f = 1 - Ny + \frac{1}{2}N(N + 1)y^2 \dots$$

Since N is large, $Ny \gtrsim 1$ even for very small values of y and then the above expansion no longer converges. One can get around this difficulty by taking the logarithm first:

$$\ln f = -N \ln(1 + y)$$

Expanding this in Taylor's series, one obtains

$$\ln f = -N(y - \frac{1}{2}y^2 \dots)$$

or

$$f = e^{-N(y - \frac{1}{2}y^2 \dots)}$$

which is valid as long as $|y| \leq 1$.

Expanding $\ln W$ in Taylor's series, one obtains

$$\ln W(n_1) = \ln W(\bar{n}_1) + B_1\eta + \frac{1}{2}B_2\eta^2 + \frac{1}{6}B_3\eta^3 + \dots \quad (1 \cdot 5 \cdot 4)$$

where

$$B_k = \frac{d^k \ln W}{dn_1^k} \quad (1 \cdot 5 \cdot 5)$$

is the k th derivative of $\ln W$ evaluated at $n_1 = \bar{n}_1$. Since one is expanding about a maximum, $B_1 = 0$ by (1 · 5 · 2). Also, since W is a maximum, it follows that the term $\frac{1}{2}B_2\eta^2$ must be negative, i.e., B_2 must be negative. To make this explicit, let us write $B_2 = -|B_2|$. Hence (1 · 5 · 4) yields, putting $\tilde{W} = \tilde{W}(\bar{n}_1)$,

$$W(n_1) = \tilde{W} e^{iB_3\eta^3 + iB_4\eta^4 \dots} = \tilde{W} e^{-\frac{1}{2}|B_2|\eta^2} e^{iB_3\eta^3 \dots} \quad (1 \cdot 5 \cdot 6)$$

In the region where η is sufficiently small, higher-order terms in the expansion can be neglected so that one obtains in first approximation an expression of the simple form

$$W(n_1) = \tilde{W} e^{-\frac{1}{2}|B_2|\eta^2} \quad (1 \cdot 5 \cdot 7)$$

Let us now investigate the expansion (1·5·4) in greater detail. By (1·4·1) one has

$$\ln W(n_1) = \ln N! - \ln n_1! - \ln(N - n_1)! + n_1 \ln p + (N - n_1) \ln q \quad (1\cdot 5\cdot 8)$$

But, if n is any large integer so that $n \gg 1$, then $\ln n!$ can be considered an almost continuous function of n , since $\ln n!$ changes only by a small fraction of itself if n is changed by a small integer. Hence

$$\frac{d \ln n!}{dn} \approx \frac{\ln(n+1)! - \ln n!}{1} = \ln \frac{(n+1)!}{n!} = \ln(n+1)$$

Thus

$$\text{for } n \gg 1, \quad \frac{d \ln n!}{dn} \approx \ln n \quad (1\cdot 5\cdot 9)$$

Hence (1·5·8) yields

$$\frac{d \ln W}{dn_1} = -\ln n_1 + \ln(N - n_1) + \ln p - \ln q \quad (1\cdot 5\cdot 10)$$

By equating this first derivative to zero one finds the value $n_1 = \bar{n}_1$, where W is maximum. Thus one obtains the condition

$$\ln \left[\frac{(N - \bar{n}_1)}{\bar{n}_1} \frac{p}{q} \right] = 0$$

or

$$(N - \bar{n}_1)p = \bar{n}_1q$$

so that

$$\blacktriangleright \quad \bar{n}_1 = Np \quad (1\cdot 5\cdot 11)$$

since $p + q = 1$.

Further differentiation of (1·5·10) yields

$$\frac{d^2 \ln W}{dn_1^2} = -\frac{1}{n_1} - \frac{1}{N - n_1} \quad (1\cdot 5\cdot 12)$$

Evaluating this for the value $n_1 = \bar{n}_1$ given in (1·5·11), one gets

$$B_2 = -\frac{1}{Np} - \frac{1}{N - Np} = -\frac{1}{N} \left(\frac{1}{p} + \frac{1}{q} \right)$$

or

$$\blacktriangleright \quad B_2 = -\frac{1}{Npq} \quad (1\cdot 5\cdot 13)$$

since $p + q = 1$. Thus B_2 is indeed negative, as required for W to exhibit a maximum.

By taking further derivatives, one can examine the higher-order terms in the expansion (1·5·4). Thus, differentiating (1·5·12), one obtains

$$B_3 = \frac{1}{\bar{n}_1^2} - \frac{1}{(N - \bar{n}_1)^2} = \frac{1}{N^2 p^2} - \frac{1}{N^2 q^2}$$

$$\text{or} \quad |B_3| = \frac{|q^2 - p^2|}{N^2 p^2 q^2} < \frac{1}{N^2 p^2 q^2}$$

$$\text{Similarly, } B_4 = -\frac{2}{\tilde{n}_1^3} - \frac{2}{(N - \tilde{n}_1)^3} = -2 \left(\frac{1}{N^2 p^3} + \frac{1}{N^2 q^3} \right)$$

or $|B_4| = \frac{2(p^3 + q^3)}{N^2 p^2 q^2} < \frac{4}{N^3 q^2 p^3}$

Thus it is seen that the k th term in (1·5·4) is smaller in magnitude than $\eta^k / (Npq)^{k-1}$. The neglect of terms beyond $B_2 n^2$, which leads to (1·5·7), is therefore justified if η is sufficiently small so that*

$$\eta \ll Npq \quad (1\cdot5\cdot14)$$

On the other hand, the factor $\exp(-\frac{1}{2}|B_2|\eta^2)$ in (1·5·6) causes W to fall off very rapidly with increasing values of $|\eta|$ if $|B_2|$ is large. Indeed, if

$$|B_2|\eta^2 = \frac{\eta^2}{Npq} \gg 1 \quad (1\cdot5\cdot15)$$

the probability $W(n_1)$ becomes negligibly small compared to $W(\tilde{n}_1)$. Hence it follows that if η is still small enough to satisfy (1·5·14) up to values of η so large that (1·5·15) is also satisfied, then (1·5·7) provides an excellent approximation for W throughout the *entire* region where W has an appreciable magnitude. This condition of simultaneous validity of (1·5·14) and (1·5·15) requires that

$$\begin{aligned} \sqrt{Npq} &\ll \eta \ll Npq \\ Npq &\gg 1 \end{aligned} \quad (1\cdot5\cdot16)$$

i.e., that

This shows that, throughout the entire domain where the probability W is not negligibly small, the expression (1·5·7) provides a very good approximation to the extent that N is large and that neither p nor q is too small.†

The value of the constant \bar{W} in (1·5·7) can be determined from the normalization condition (1·4·2). Since W and n_1 can be treated as quasicontinuous variables, the sum over all integral values of n_1 can be approximately replaced by an integral. Thus the normalization condition can be written

$$\sum_{n_1=0}^N W(n_1) \approx \int W(n_1) dn_1 = \int_{-\infty}^{\infty} W(\tilde{n}_1 + \eta) d\eta = 1 \quad (1\cdot5\cdot17)$$

Here the integral over η can to excellent approximation be extended from $-\infty$ to $+\infty$, since the integrand makes a negligible contribution to the integral wherever $|\eta|$ is large enough so that W is far from its pronounced maximum value. Substituting (1·5·7) into (1·5·17) and using (A·4·2), one obtains

$$\bar{W} \int_{-\infty}^{\infty} e^{-\frac{1}{2}|B_2|\eta^2} = \bar{W} \sqrt{\frac{2\pi}{|B_2|}} = 1$$

* Note that the condition (1·5·1) is equivalent to $|\partial W/\partial n_1| \ll W$, i.e., to $|B_2| = (Npq)^{-1}|\eta| \ll 1$ by virtue of (1·5·7) and (1·5·13). Thus it is also satisfied in the domain (1·5·14) where W is not too small.

† When $p \ll 1$ or $q \ll 1$, it is possible to obtain a different approximation for the binomial distribution, namely the so-called "Poisson distribution" (see Problem 1.9).

Thus (1·5·7) becomes

$$\blacktriangleright \quad W(n_1) = \sqrt{\frac{|B_2|}{2\pi}} e^{-\frac{1}{2}|B_2|(n_1 - \bar{n}_1)^2} \quad (1\cdot 5\cdot 18)$$

The reasoning leading to the functional form (1·5·7) or (1·5·18), the so-called "Gaussian distribution," has been very general in nature. Hence it is not surprising that Gaussian distributions occur very frequently in statistics whenever one is dealing with large numbers. In our case of the binomial distribution, the expression (1·5·18) becomes, by virtue of (1·5·11) and (1·5·13),

$$\blacktriangleright \quad W(n_1) = (2\pi Npq)^{-\frac{1}{2}} \exp \left[-\frac{(n_1 - Np)^2}{2Npq} \right] \quad (1\cdot 5\cdot 19)$$

Note that (1·5·19) is, for large values of N and n_1 , much simpler than (1·4·1), since it does not require the evaluation of large factorials. Note also that, by (1·4·4) and (1·4·9), the expression (1·5·19) can be written in terms of the mean values \bar{n}_1 and $(\Delta n_1)^2$ as

$$W(n_1) = [2\pi(\Delta n_1)^2]^{-\frac{1}{2}} \exp \left[-\frac{(n_1 - \bar{n}_1)^2}{2(\Delta n_1)^2} \right]$$

1 · 6 Gaussian probability distributions

The Gaussian approximation (1·5·19) also yields immediately the probability $P(m)$ that in a large number of N steps the net displacement is m . The corresponding number of right steps is, by (1·2·9), $n_1 = \frac{1}{2}(N + m)$. Hence (1·5·19) gives

$$P(m) = W\left(\frac{N+m}{2}\right) = [2\pi Npq]^{-\frac{1}{2}} \exp \left\{ -\frac{[m - N(p-q)]^2}{8Npq} \right\} \quad (1\cdot 6\cdot 1)$$

since $n_1 - Np = \frac{1}{2}[N + m - 2Np] = \frac{1}{2}[m - N(p-q)]$. By (1·2·3) one has $m = 2n_1 - N$, so that m assumes here integral values separated by an amount $\Delta m = 2$.

We can also express this result in terms of the actual displacement variable x ,

$$x = ml \quad (1\cdot 6\cdot 2)$$

where l is the length of each step. If l is small compared to the smallest length of interest in the physical problem under consideration,* the fact that x can only assume values in discrete increments of $2l$, rather than all values continuously, becomes unimportant. Furthermore, when N is large, the probability $P(m)$ of occurrence of a displacement m does not change significantly from

* For example, if one considers the random motion (diffusion) of an atom in a solid, the step length l is of the order of one lattice spacing, i.e., about 10^{-8} cm. But on the macroscopic scale of experimental measurement the smallest length L of relevance might be 1 micron = 10^{-4} cm.

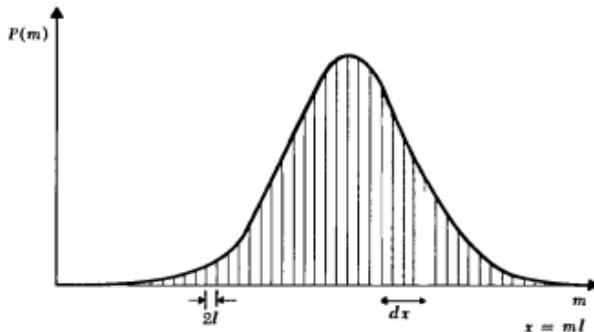


Fig. 1·6·1 The probability $P(m)$ of a net displacement of m units when the total number N of steps is very large and the step length l is very small.

one possible value of m to an adjacent one; i.e., $|P(m+2) - P(m)| \ll P(m)$. Therefore $P(m)$ can be regarded as a smooth function of x . A bar graph of the type shown in Fig. 1·4·1 then assumes the character illustrated in Fig. 1·6·1, where the bars are very densely spaced and their envelope forms a smooth curve.

Under these circumstances it is possible to regard x as a continuous variable on a macroscopic scale and to ask for the probability that the particle is found after N steps in the range between x and $x + dx$.* Since m assumes only integral values separated by $\Delta m = 2$, the range dx contains $dx/2l$ possible values of m , all of which occur with nearly the same probability $P(m)$. Hence the probability of finding the particle anywhere in the range between x and $x + dx$ is simply obtained by summing $P(m)$ over all values of m lying in dx , i.e., by multiplying $P(m)$ by $dx/2l$. This probability is thus proportional to dx (as one would expect) and can be written as

$$\Phi(x) dx = P(m) \frac{dx}{2l} \quad (1·6·3)$$

where the quantity $\Phi(x)$, which is independent of the magnitude of dx , is called a “probability density.” Note that it must be multiplied by a differential element of length dx to yield a probability.

By using (1·6·1) one then obtains

$$\Phi(x) dx = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} dx \quad (1·6·4)$$

where we have used the abbreviations

$$\mu = (p - q)Nl \quad (1·6·5)$$

$$\text{and } \sigma = 2 \sqrt{Npq} l \quad (1·6·6)$$

* Here dx is understood to be a differential in the *macroscopic* sense, i.e., $dx \ll L$, where L is the smallest dimension of relevance in the *macroscopic* discussion, but $dx \gg l$. (In other words, dx is *macroscopically* small, but *microscopically* large.)

The expression (1·6·4) is the standard form of the Gaussian probability distribution. The great generality of the argument leading to (1·5·19) suggests that such Gaussian distributions occur very frequently in probability theory whenever one deals with large numbers.

Using (1·6·4), one can quite generally compute the mean values \bar{x} and $(x - \bar{x})^2$. In calculating these mean values, sums over all possible intervals dx become, of course, integrations. (The limits of x can be taken as $-\infty < x < \infty$, since $\varphi(x)$ itself becomes negligibly small whenever $|x|$ is so large as to lead to a displacement inaccessible in N steps.)

First we verify that $\varphi(x)$ is properly normalized, i.e., that the probability of the particle being somewhere is unity. Thus

$$\begin{aligned}\int_{-\infty}^{\infty} \varphi(x) dx &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2} dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-y^2/2\sigma^2} dy \\ &= \frac{1}{\sqrt{2\pi}\sigma} \sqrt{\pi 2\sigma^2} \\ &= 1\end{aligned}\quad (1·6·7)$$

Here we have put $y = x - \mu$ and evaluated the integral by (A·4·2).

Next we calculate the mean value

$$\begin{aligned}\bar{x} &\equiv \int_{-\infty}^{\infty} x\varphi(x) dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x e^{-(x-\mu)^2/2\sigma^2} dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \left[\int_{-\infty}^{\infty} y e^{-y^2/2\sigma^2} dy + \mu \int_{-\infty}^{\infty} e^{-y^2/2\sigma^2} dy \right]\end{aligned}$$

Since the integrand in the first integral is an odd function of y , the first integral

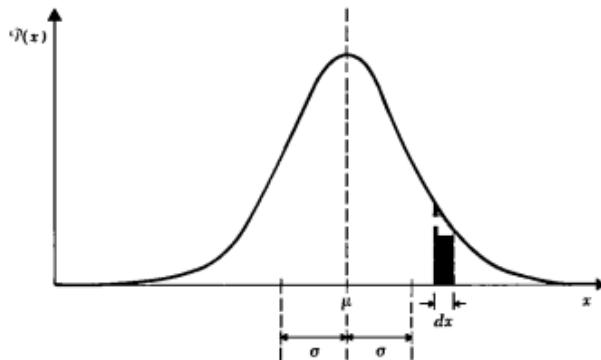


Fig. 1·6·2 The Gaussian distribution. Here $\varphi(x) dx$ is the area under the curve in the interval between x and $x + dx$ and is thus the probability that the variable x lies in this range.

vanishes by symmetry. The second integral is the same as that in (1·6·7), so that one gets

$$\blacktriangleright \quad \bar{x} = \mu \quad (1\cdot6\cdot8)$$

This is a simple consequence of the fact that $\mathcal{P}(x)$ is only a function of $|x - \mu|$ and is thus symmetric about the position $x = \mu$ of its maximum. Hence this point corresponds also to the mean value of x .

The dispersion becomes

$$\begin{aligned}\overline{(x - \mu)^2} &\equiv \int_{-\infty}^{\infty} (x - \mu)^2 \mathcal{P}(x) dx \\ &= \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{\infty} y^2 e^{-y^2/2\sigma^2} dy \\ &= \frac{1}{\sqrt{2\pi} \sigma} \left[\frac{\sqrt{\pi}}{2} (2\sigma^2)^{\frac{3}{2}} \right] \\ &= \sigma^2\end{aligned}$$

where we have used the integral formulas (A·4·6). Thus

$$\blacktriangleright \quad \overline{(\Delta x)^2} = \overline{(x - \mu)^2} = \sigma^2 \quad (1\cdot6\cdot9)$$

Hence σ is simply the root-mean-square deviation of x from the mean of the Gaussian distribution.

By (1·6·5) and (1·6·6), one then obtains for the random walk problem the relations

$$\begin{aligned}\bar{x} &= (p - q)Nl & (1\cdot6\cdot10) \\ \overline{(\Delta x)^2} &= 4Npql^2 & (1\cdot6\cdot11)\end{aligned}$$

These results (derived here for the case of large N) agree, as they must, with the mean values $\bar{x} = \bar{m}l$ and $\overline{(\Delta x)^2} = \overline{(\Delta m)^2}l^2$ already calculated in (1·4·6) and (1·4·12) for the general case of arbitrary N .

GENERAL DISCUSSION OF THE RANDOM WALK

Our discussion of the random walk problem has yielded a great number of important results and introduced many fundamental concepts in probability theory. The approach we have used, based on combinatorial analysis to calculate the probability distribution, has, however, severe limitations. In particular, it is difficult to generalize this approach to other cases, for example, to situations where the length of each step is not the same or where the random walk takes place in more than one dimension. We now turn, therefore, to a discussion of more powerful methods which can readily be generalized and yet possess a basic simplicity and directness of their own.

1 · 7 Probability distributions involving several variables

The statistical description of a situation involving more than one variable requires only straightforward generalizations of the probability arguments applicable to a single variable. Let us then, for simplicity, consider the case of only two variables u and v which can assume the possible values

$$\begin{aligned} u_i &\quad \text{where } i = 1, 2, \dots, M \\ \text{and} \quad v_j &\quad \text{where } j = 1, 2, \dots, N \end{aligned}$$

Let $P(u_i; v_j)$ be the probability that u assumes the value u_i and that v assumes the value v_j .

The probability that the variables u and v assume any of their possible sets of values must be unity; i.e., one has the normalization requirement

$$\sum_{i=1}^M \sum_{j=1}^N P(u_i; v_j) = 1 \quad (1 \cdot 7 \cdot 1)$$

where the summation extends over all possible values of u and all possible values of v .

The probability $P_u(u_i)$ that u assumes the value u_i , irrespective of the value assumed by the variable v , is the sum of the probabilities of all possible situations consistent with the given value of u_i ; i.e.,

$$P_u(u_i) = \sum_{j=1}^N P(u_i; v_j) \quad (1 \cdot 7 \cdot 2)$$

where the summation is over all possible values of v_j . Similarly, the probability $P_v(v_j)$ that v assumes the value v_j , irrespective of the value assumed by u , is

$$P_v(v_j) = \sum_{i=1}^M P(u_i; v_j) \quad (1 \cdot 7 \cdot 3)$$

Each of the probabilities P_u and P_v is, of course, properly normalized. For example, by (1 · 7 · 2) and (1 · 7 · 1) one has

$$\sum_{i=1}^M P_u(u_i) = \sum_{i=1}^M \left[\sum_{j=1}^N P(u_i; v_j) \right] = 1 \quad (1 \cdot 7 \cdot 4)$$

An important special case occurs when the probability that one variable assumes a certain value does not depend on the value assumed by the other variable. The variables are then said to be "statistically independent" or,

“uncorrelated.” The probability $P(u_i, v_j)$ can then be expressed very simply in terms of the probability $P_u(u_i)$ that u assumes the value u_i and the probability $P_v(v_j)$ that v assumes the value v_j . Indeed, in this case [the number of instances in the ensemble where $u = u_i$ and where simultaneously $v = v_j$] is simply obtained by multiplying [the number of instances where $u = u_i$] by [the number of instances where $v = v_j$]; hence

$$P(u_i, v_j) = P_u(u_i)P_v(v_j) \quad (1 \cdot 7 \cdot 5)$$

if u and v are statistically independent.

Let us now mention some properties of mean values. If $F(u, v)$ is any function of u and v , then its mean value is defined by

$$\overline{F(u, v)} \equiv \sum_{i=1}^M \sum_{j=1}^N P(u_i, v_j)F(u_i, v_j) \quad (1 \cdot 7 \cdot 6)$$

Note that if $f(u)$ is a function of u only, it also follows by (1 · 7 · 2) that

$$\overline{f(u)} = \sum_i \sum_j P(u_i, v_j)f(u_i) = \sum_i P_u(u_i)f(u_i) \quad (1 \cdot 7 \cdot 7)$$

If F and G are any functions of u and v , then one has the general result

$$\begin{aligned} \overline{F + G} &\equiv \sum_i \sum_j P(u_i, v_j)[F(u_i, v_j) + G(u_i, v_j)] \\ &= \sum_i \sum_j P(u_i, v_j)F(u_i, v_j) + \sum_i \sum_j P(u_i, v_j)G(u_i, v_j) \end{aligned}$$

or

$$\blacktriangleright \quad \overline{F + G} = \overline{F} + \overline{G} \quad (1 \cdot 7 \cdot 8)$$

i.e., the average of a sum equals simply the sum of the averages.

Given any two functions $f(u)$ and $g(v)$, one can also make a general statement about the mean value of their product if u and v are statistically independent variables. Indeed, one then finds

$$\begin{aligned} \overline{f(u)g(v)} &\equiv \sum_i \sum_j P(u_i, v_j)f(u_i)g(v_j) \\ &= \sum_i \sum_j P_u(u_i)P_v(v_j)f(u_i)g(v_j) \quad \text{by (1 · 7 · 5)} \\ &= \left[\sum_i P_u(u_i)f(u_i) \right] \left[\sum_j P_v(v_j)g(v_j) \right] \end{aligned}$$

Thus

$$\blacktriangleright \quad \overline{f(u)g(v)} = \overline{f(u)} \overline{g(v)} \quad (1 \cdot 7 \cdot 9)$$

i.e., the average of a product equals the product of the averages if u and v are statistically independent. If u and v are statistically not independent, the statement (1 · 7 · 9) is in general not true.

The generalization of the definitions and results of this section to the case of more than two variables is immediate.

1 · 8 Comments on continuous probability distributions

Consider first the case of a single variable u which can assume any value in the continuous range $a_1 < u < a_2$. To give a probability description of such a situation, one can focus attention on any infinitesimal range of the variable between u and $u + du$ and ask for the probability that the variable assumes a value in this range. One expects that this probability is proportional to the magnitude of du if this interval is sufficiently small; i.e., one expects that this probability can be written in the form $\varphi(u) du$, where $\varphi(u)$ is independent of the size of du .* The quantity $\varphi(u)$ is called a "probability density." Note that it must be multiplied by du to yield an actual probability.

It is readily possible to reduce the problem dealing with a continuous variable to an equivalent discrete problem where the number of possible values of the variable becomes countable. It is only necessary to subdivide the accessible range $a_1 < u < a_2$ of the variable into (arbitrarily small) equal intervals of fixed size δu . Each such interval can then be labeled by some index i . The value of u in this interval can be denoted simply by u_i , and the probability of finding the variable in this range by $P(u_i)$. One can then deal with a denumerable set of values of the variable u (each of them corresponding to one of the chosen fixed infinitesimal intervals). It also becomes clear that relations involving probabilities of discrete variables are equally valid for probabilities of continuous variables. For example, the simple properties (1 · 3 · 5) and (1 · 3 · 6) of mean values are also applicable if u is a continuous variable.

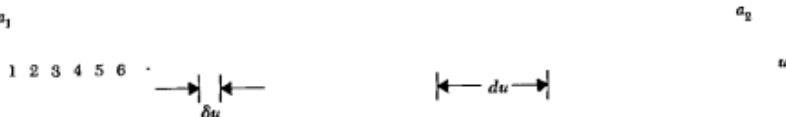


Fig. 1 · 8 · 1 Subdivision of the range $a_1 < u < a_2$ of a continuous variable u into a countable number of infinitesimal intervals δu of fixed size.

To make the connection between the continuous and discrete points of view quite explicit, note that in terms of the original infinitesimal subdivision interval δu ,

$$P(u) = \varphi(u) \delta u$$

Similarly, if one considers any interval between u and $u + du$ which is such that du is macroscopically small although $du \gg \delta u$, then this interval contains $du/\delta u$ possible values of u_i for which the probability $P(u_i)$ has essentially the same value—call it simply $P(u)$. Then the probability $P(u) du$ of

* Indeed, the probability must be expressible as a Taylor's series in powers of du and must vanish as $du \rightarrow 0$. Hence the leading term must be of the form φdu , while terms involving higher powers of du are negligible if du is sufficiently small.

the variable assuming a value between u and $u + du$ should be given by multiplying the probability $P(u)$ for assuming any discrete value in this range by the number $du/\delta u$ of discrete values in this range; i.e., one has properly

$$\varphi(u) du = P(u_i) \frac{du}{\delta u} = \frac{P(u)}{\delta u} du \quad (1 \cdot 8 \cdot 1)$$

Note that the sums involved in calculating normalization conditions or mean values can be written as integrals if the variable is continuous. For example, the normalization condition asserts that the sum of the probabilities over all possible values of the variable must equal unity; in symbols

$$\sum_i P(u_i) = 1 \quad (1 \cdot 8 \cdot 2)$$

But if the variable is continuous, one can first sum over all values of the variable in a range between u and $u + du$, thus obtaining the probability $\varphi(u) du$ that the variable lies in this range, and then complete the sum (1·8·2) by summing (i.e., integrating) over all such possible ranges du . Thus (1·8·2) is equivalent to

$$\int_{a_1}^{a_2} \varphi(u) du = 1 \quad (1 \cdot 8 \cdot 3)$$

which expresses the normalization condition in terms of the probability density $\varphi(u)$. Similarly, one can calculate mean values in terms of $\varphi(u)$. The general definition of the mean value of a function f was given in terms of discrete variables by (1·3·4) as

$$\overline{f(u)} = \sum_i P(u_i) f(u_i) \quad (1 \cdot 8 \cdot 4)$$

In a continuous description one can again sum first over all values between u and $u + du$ (this contributes to the sum an amount $\varphi(u) du f(u)$) and then integrate over all possible ranges du . Thus (1·8·4) is equivalent to the relation

$$\overline{f(u)} = \int_{a_1}^{a_2} \varphi(u) f(u) du \quad (1 \cdot 8 \cdot 5)$$

Remark Note that in some cases it is possible that the probability density $\varphi(u)$ itself becomes infinite for certain values of u . This does not lead to any difficulties as long as the integral $\int_{a_1}^{a_2} \varphi(u) du$, which measures the probability that u assumes any value in some arbitrary range between a_1 and a_2 , always remains finite.

The extension of these comments to probabilities involving several variables is immediate. Consider, for example, the case of two variables u and v which can assume all values in the continuous respective ranges $a_1 < u < a_2$

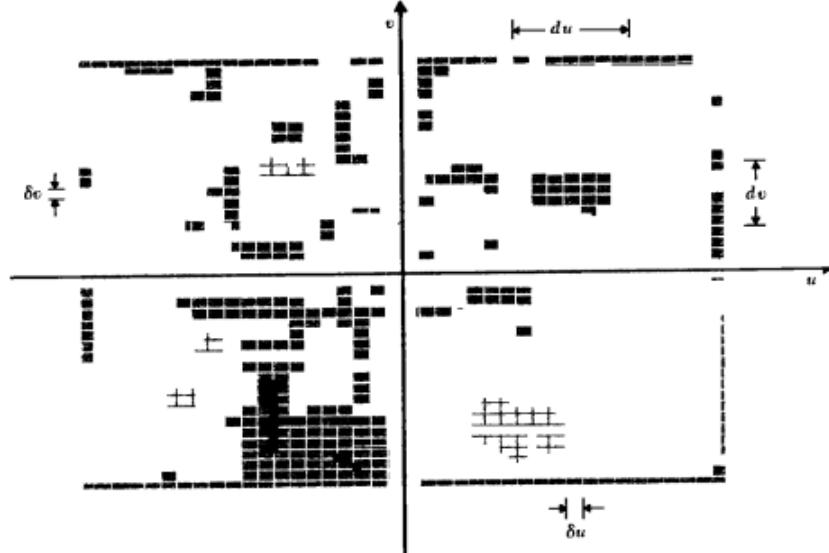


Fig. 1·8·2 Subdivision of the continuous variables u and v into small intervals of magnitude δu and δv .

and $b_1 < v < b_2$. One can then talk of the probability $\Phi(u,v) du dv$ that the variables lie in the ranges between u and $u + du$ and between v and $v + dv$, respectively, where $\Phi(u,v)$ is a probability density independent of the size of du and dv . It is again possible to reduce the problem to an equivalent one involving discrete countable values of the variables. It is only necessary to subdivide the possible values of u into fixed infinitesimal intervals of size δu and labeled by i , and those of v into fixed infinitesimal intervals of size δv and labeled by j . Then one can speak of the probability $P(u_i, v_j)$ that $u_i = u$ and that simultaneously $v = v_j$. Analogously to (1·7·1) one then has the relation

$$\Phi(u,v) du dv = P(u,v) \frac{du}{\delta u} \frac{dv}{\delta v}$$

where the factor multiplying $P(u,v)$ is simply the number of infinitesimal cells of magnitude $\delta u \delta v$ contained in the range lying between u and $u + du$ and between v and $v + dv$.

The normalization condition (1·7·2) can then be written in terms of the probability density $\Phi(u,v)$ as

$$\int_{a_1}^{a_2} \int_{b_1}^{b_2} du dv \Phi(u,v) = 1 \quad (1·8·6)$$

Analogously to (1·7·7) one can also write

$$\overline{F(u,v)} = \int_{a_1}^{a_2} \int_{b_1}^{b_2} du dv \Phi(u,v) F(u,v) \quad (1·8·7)$$

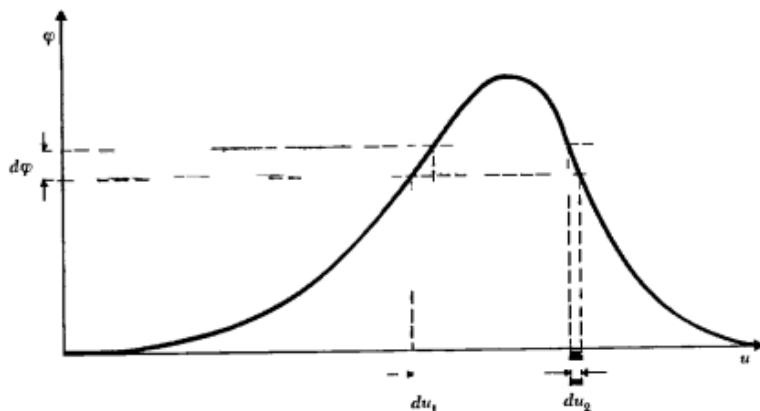


Fig. 1·8·3 Illustration showing a function $\varphi(u)$ which is such that $u(\varphi)$ is a double-valued function of φ . Here the range $d\varphi$ corresponds to u being either in the range du_1 or in the range du_2 .

Since the problem can be formulated in discrete as well as continuous terms, the general properties (1·7·8) and (1·7·9) of mean values remain, of course, valid in the continuous case.

Functions of random variables Consider the case of a single variable u and suppose that $\varphi(u)$ is some continuous function of u . The following question arises quite frequently. If $\Phi(u) du$ is the probability that u lies in the range between u and $u + du$, what is the corresponding probability $W(\varphi) d\varphi$ that φ lies in the range between φ and $\varphi + d\varphi$? Clearly, the latter probability is obtained by adding up the probabilities for all those values u which are such that φ lies in the range between φ and $\varphi + d\varphi$; in symbols

$$W(\varphi) d\varphi = \int_{d\varphi} \Phi(u) du \quad (1\cdot8\cdot8)$$

Here u can be considered a function of φ and the integral extends over all those values of u lying in the range between $u(\varphi)$ and $u(\varphi + d\varphi)$. Thus (1·8·8) becomes simply

$$W(\varphi) d\varphi = \int_{\varphi}^{\varphi+d\varphi} \Phi(u) \left| \frac{du}{d\varphi} \right| d\varphi = \Phi(u) \left| \frac{du}{d\varphi} \right| d\varphi \quad (1\cdot8\cdot9)$$

The last step assumes that u is a single-valued function of φ and follows, since the integral is extended only over an infinitesimal range $d\varphi$. Since $u = u(\varphi)$, the right side of (1·8·9) can, of course, be expressed completely in terms of φ . If $\varphi(u)$ is not a single-valued function of u , then the integral (1·8·8) may consist of several contributions similar to those of (1·8·9) (see Fig. 1·8·3).

Similar arguments can be used to find the probabilities for functions of *several* variables when the probabilities of the variables themselves are known.

Example Suppose that a two-dimensional vector \mathbf{B} of constant length $B = |\mathbf{B}|$ is equally likely to point in any direction specified by the angle θ (see Fig. 1·8·4). The probability $\mathcal{P}(\theta) d\theta$ that this angle lies in the range between θ and $\theta + d\theta$ is then given by the ratio of the angular range $d\theta$ to the total angular range 2π subtended by a full circle; i.e.,

$$\mathcal{P}(\theta) d\theta = \frac{d\theta}{2\pi} \quad (1\cdot8\cdot10)$$

If the vector makes an angle θ with the x axis, its x component is given by

$$B_x = B \cos \theta \quad (1\cdot8\cdot11)$$

What is the probability $W(B_x) dB_x$ that the x component of this vector lies between B_x and $B_x + dB_x$? Clearly, B_x is always such that $-B \leq B_x \leq B$. In this interval an infinitesimal range between B_x and $B_x + dB_x$ corresponds to two possible infinitesimal ranges of $d\theta$ (see Fig. 1·8·4), each of magnitude $d\theta$ connected to dB_x through the relation (1·8·11) so that $dB_x = |B \sin \theta| d\theta$. By virtue of (1·8·10) the probability $W(B_x) dB_x$ is then given by

$$W(B_x) dB_x = 2 \left[\frac{1}{2\pi} \frac{dB_x}{|B \sin \theta|} \right] = \frac{1}{\pi B} \frac{dB_x}{|\sin \theta|}$$

But, by (1·8·11),

$$|\sin \theta| = (1 - \cos^2 \theta)^{\frac{1}{2}} = \left[1 - \left(\frac{B_x}{B} \right)^2 \right]^{\frac{1}{2}}$$

Hence

$$W(B_x) dB_x = \begin{cases} \frac{dB_x}{\pi \sqrt{B^2 - B_x^2}} & \text{for } -B \leq B_x \leq B \\ 0 & \text{otherwise} \end{cases} \quad (1\cdot8\cdot12)$$

The probability density is maximum (indeed infinite) as $|B_x| \rightarrow B$ and is

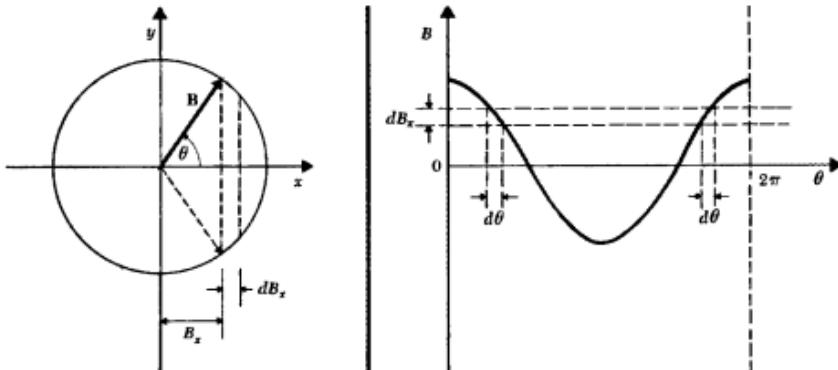


Fig. 1·8·4 Dependence of the x component $B_x = B \cos \theta$ of a two-dimensional vector \mathbf{B} on its polar angle θ .

minimum when $B_x = 0$. This result is apparent from the geometry of Fig. (1·8·4), since a given narrow range of dB_x corresponds to a relatively large range of the angle θ when $B_x \approx B$ and to a very much narrower range of the angle θ when $B_x \approx 0$.

1 · 9 General calculation of mean values for the random walk

The comments of Sec. 1·7 allow one to calculate mean values for very general situations by very simple and transparent methods. Consider a quite general form of the one-dimensional random walk problem. Let s_i denote the displacement (positive or negative) in the i th step

Let $w(s_i) ds_i$ be the probability that the i th displacement lies in the range between s_i and $s_i + ds_i$.

We assume again that this probability is independent of what displacements occur in any other steps. For simplicity we further assume that the probability distribution w is the same for each step i . Nevertheless, it is clear that the situation envisaged here is considerably more general than before, for we no longer necessarily assume a fixed magnitude l of displacement at each step, but a distribution of possible step lengths with relative probability specified by w .

We are interested in the total displacement x after N steps. We can ask for $\Phi(x) dx$, the probability that x lies in the range between x and $x + dx$. We can also ask for mean values (or moments) of x . In this section we show that the calculation of these moments can be achieved very simply without prior knowledge of $\Phi(x)$.



Fig. 1·9·1 Some examples of probability distributions giving, for any one step, the probability $w(s) ds$ that the displacement is between s and $s + ds$.
 (a) A rather general case, displacements to the right being more probable than those to the left. (b) The special case discussed in Sec. 1·2. Here the peaks, centered about $+l$ and $-l$, respectively, are very narrow; the area under the right peak is p , that under the left one is q . (The curves (a) and (b) are not drawn to the same scale; the total area under each should be unity.)

The total displacement x is equal to

$$x = s_1 + s_2 + \dots + s_N = \sum_{i=1}^N s_i \quad (1 \cdot 9 \cdot 1)$$

Taking mean values of both sides,

$$\bar{x} = \overline{\sum_{i=1}^N s_i} = \sum_{i=1}^N \bar{s}_i \quad (1 \cdot 9 \cdot 2)$$

where we have used the property (1 · 7 · 8). But since $w(s_i)$ is the same for each step, independent of i , each mean value \bar{s}_i is the same. Thus (1 · 9 · 2) is simply the sum of N equal terms and becomes

► $\bar{x} = N\bar{s} \quad (1 \cdot 9 \cdot 3)$

where $\bar{s} \equiv \bar{s}_i = \int ds w(s)s \quad (1 \cdot 9 \cdot 4)$

is merely the mean displacement per step.

Next we calculate the dispersion

$$\overline{(\Delta x)^2} \equiv \overline{(x - \bar{x})^2} \quad (1 \cdot 9 \cdot 5)$$

By (1 · 9 · 1) and (1 · 9 · 2) one has

$$x - \bar{x} = \sum_i (s_i - \bar{s})$$

or $\Delta x = \sum_{i=1}^N \Delta s_i \quad (1 \cdot 9 \cdot 6)$

where $\Delta s = s_i - \bar{s} \quad (1 \cdot 9 \cdot 7)$

By squaring (1 · 9 · 6) one obtains

$$(\Delta x)^2 = \left(\sum_{i=1}^N \Delta s_i \right) \left(\sum_{j=1}^N \Delta s_j \right) = \sum_i (\Delta s_i)^2 + \sum_i \sum_{i \neq j} (\Delta s_i)(\Delta s_j) \quad (1 \cdot 9 \cdot 8)$$

Here the first term on the right represents all the square terms, and the second term all the cross terms, originating from the multiplication of the sum by itself. Taking the mean value of (1 · 9 · 8) yields, by virtue of (1 · 7 · 8),

$$\overline{(\Delta x)^2} = \sum_i \overline{(\Delta s_i)^2} + \sum_i \sum_{i \neq j} \overline{\Delta s_i \Delta s_j} \quad (1 \cdot 9 \cdot 9)$$

In the cross terms we make use of the fact that different steps are statistically independent and apply the relation (1 · 7 · 9) to write for $i \neq j$

$$\overline{(\Delta s_i)(\Delta s_j)} = \overline{(\Delta s_i)} \overline{(\Delta s_j)} = 0 \quad (1 \cdot 9 \cdot 10)$$

since $\overline{\Delta s_i} = \bar{s}_i - \bar{s} = 0$

In short, each cross term vanishes on the average, being as often positive as

negative. Thus (1·9·9) reduces simply to a sum of square terms

$$\overline{(\Delta x)^2} = \sum_{i=1}^N \overline{(\Delta s_i)^2} \quad (1\cdot9\cdot11)$$

Of course, none of these square terms can be negative. Since the probability distribution $w(s_i)$ is the same for each step, independent of i , it again follows that $\overline{(\Delta s_i)^2}$ must be the same for each step. Thus the sum in (1·9·11) consists merely of N equal terms and becomes simply

$$\blacktriangleright \quad \overline{(\Delta x)^2} = N \overline{(\Delta s)^2} \quad (1\cdot9\cdot12)$$

where $\overline{(\Delta s)^2} \equiv \overline{(\Delta s_i)^2} = \int ds w(s) (\Delta s)^2 \quad (1\cdot9\cdot13)$

is just the mean square displacement per step.

Despite their great simplicity, the relations (1·9·3) and (1·9·12) are very general and important results. The dispersion $\overline{(\Delta x)^2} = \overline{(x - \bar{x})^2}$ is a measure of the *square* of the width of the distribution of the net displacement about its mean \bar{x} . The square root $\Delta^*x \equiv [\overline{(\Delta x)^2}]^{1/2}$, i.e., the “root-mean-square (rms) deviation from the mean,” thus provides a direct measure of the width of the distribution of the displacement about the mean \bar{x} . The results (1·9·3) and (1·9·12) thus make possible the following interesting statements about the sum (1·9·1) of statistically independent variables. If $\bar{s} \neq 0$ and the number N of these variables (e.g., of displacements) increases, the mean value \bar{x} of their sum tends to increase proportionally to N . The width Δ^*x of the distribution about the mean also increases, but only proportionately to $N^{1/2}$. Hence the *relative* magnitude of the width Δx^* compared to the mean \bar{x} itself *decreases* like $N^{-1/2}$; explicitly, one has by (1·9·3) and (1·9·12), if $\bar{s} \neq 0$,

$$\frac{\Delta^*x}{\bar{x}} = \frac{\Delta^*s}{\bar{s}} \frac{1}{\sqrt{N}}$$

where $\Delta^*s = [\overline{(\Delta s)^2}]^{1/2}$. This means that the percentage deviation of the distribution of values of x about their mean \bar{x} becomes increasingly negligible as the number N becomes large. This is a characteristic feature of statistical distributions involving large numbers.

Example Let us apply the general results (1·9·3) and (1·9·12) of this section to the special case of the random walk with fixed step length l previously discussed in Sec. 1·2. There the probability of a step to the right is p , that of a step to the left is $q = 1 - p$. The mean displacement per step is then given by

$$\bar{s} = pl + q(-l) = (p - q)l = (2p - 1)l \quad (1\cdot9\cdot14)$$

As a check, note that $\bar{s} = 0$ if $p = q$, as required by symmetry.

Also $\overline{s^2} = pl^2 + q(-l)^2 = (p + q)l^2 = l^2$

Hence $\overline{(\Delta s)^2} = \overline{s^2} - \bar{s}^2 = l^2[1 - (2p - 1)^2]$

$$= l^2[1 - 4p^2 + 4p - 1] = 4l^2p(1 - p)$$

or $\overline{(\Delta s)^2} = 4pql^2 \quad (1\cdot9\cdot15)$

Hence the relations (1·9·3) and (1·9·12) yield

$$\begin{aligned}\bar{x} &= (p - q)Nl \\ \frac{\bar{x}}{(\Delta x)^2} &= 4pqNl^2\end{aligned}\quad (1\cdot 9\cdot 16)$$

Since $x = ml$, these agree precisely with the previously computed results (1·4·6) and (1·4·12).

*1·10 Calculation of the probability distribution

For the problem discussed in the last section, the total displacement x in N steps is given by

$$x = \sum_{i=1}^N s_i \quad (1\cdot 10\cdot 1)$$

We now want to find the probability $\Phi(x) dx$ of finding x in the range between x and $x + dx$. Since the steps are statistically independent, the probability of a particular sequence of steps where

$$\begin{aligned}&\text{the 1st displacement lies in the range between } s_1 \text{ and } s_1 + ds_1 \\ &\text{the 2nd displacement lies in the range between } s_2 \text{ and } s_2 + ds_2 \\ &\quad \dots \\ &\text{the } N\text{th displacement lies in the range between } s_N \text{ and } s_N + ds_N\end{aligned}$$

is simply given by the product of the respective probabilities, i.e., by

$$w(s_1) ds_1 \cdot w(s_2) ds_2 \cdots w(s_N) ds_N$$

If we sum this probability over all the possible individual displacements which are consistent with the condition that the total displacement x in (1·10·1) always lies in the range between x and $x + dx$, then we obtain the total probability $\Phi(x) dx$, irrespective of the sequence of steps producing this total displacement. In symbols we can write

$$\Phi(x) dx = \iint \cdots \int_{-\infty}^{\infty} w(s_1) w(s_2) \cdots w(s_N) ds_1 ds_2 \cdots ds_N \quad (1\cdot 10\cdot 2)$$

where the integration is over all possible values of the variables s_i , subject to the restriction that

$$x < \sum_{i=1}^N s_i < x + dx \quad (1\cdot 10\cdot 3)$$

In principle, evaluation of the integral (1·10·2) solves completely the problem of finding $\Phi(x)$.

In practice, the integral (1·10·2) is difficult to evaluate because the condition (1·10·3) makes the limits of integration very awkward; i.e., one faces the complicated geometrical problem of determining over what subspace, consistent with (1·10·3), one has to integrate. A powerful way of handling this kind of problem is to eliminate the geometrical problem by integrating over *all* values of the variables s_i *without* restriction, while shifting the complication introduced by the condition of constraint (1·10·3) to the *integrand*. This can readily be done by multiplying the integrand in (1·10·2) by a factor which is equal to unity when the s_i are such that (1·10·3) is satisfied, but which equals zero otherwise. The Dirac δ function $\delta(x - x_0)$, discussed in Appendix A·7, has precisely the selective property that it vanishes whenever $|x - x_0| > \frac{1}{2}|dx|$, while it becomes infinite like $(dx)^{-1}$ in the infinitesimal range where $|x - x_0| < \frac{1}{2}|dx|$; i.e., $\delta(x - x_0) dx = 1$ in this latter range. Hence (1·10·2) can equally well be written

$$\Phi(x) dx = \iint_{-\infty}^{\infty} \cdots \int w(s_1) w(s_2) \cdots w(s_N) \left[\delta \left(x - \sum_{i=1}^N s_i \right) dx \right] ds_1 ds_2 \cdots ds_N \quad (1 \cdot 10 \cdot 4)$$

where there is now *no* further restriction on the domain of integration. At this point we can use the convenient analytical representation of the δ function in terms of the integral (A·7·14); i.e., we can write

$$\delta(x - \sum s_i) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(\sum s_i - x)} \quad (1 \cdot 10 \cdot 5)$$

Substituting this result in (1·10·4) yields:

$$\begin{aligned} \Phi(x) &= \iint \cdots \int w(s_1) w(s_2) \cdots w(s_N) \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(s_1 + \cdots + s_N - x)} ds_1 ds_2 \cdots ds_N \\ \text{or } \Phi(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ikx} \int_{-\infty}^{\infty} ds_1 w(s_1) e^{iks_1} \cdots \int_{-\infty}^{\infty} ds_N w(s_N) e^{iks_N} \end{aligned} \quad (1 \cdot 10 \cdot 6)$$

where we have interchanged the order of integration and used the multiplicative property of the exponential function. Except for the irrelevant symbol used as variable of integration, each of the last N integrals is identical and equal to

$$\blacktriangleright Q(k) \equiv \int_{-\infty}^{\infty} ds e^{iks} w(s) \quad (1 \cdot 10 \cdot 7)$$

Hence (1·10·6) becomes

$$\blacktriangleright \Phi(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ikx} Q^N(k) \quad (1 \cdot 10 \cdot 8)$$

Thus the evaluation of two simple (Fourier) integrals solves the problem completely.

Example Let us again apply the present results to the case of equal step lengths l discussed in Sec. 1·2. There the probability of a displacement $+l$ is equal to p , that of a displacement $-l$ is equal to $q = 1 - p$; i.e., the corresponding probability density w is given by

$$w(s) = p\delta(s - l) + q\delta(s + l)$$

The quantity (1.10.7) becomes

$$Q(k) = \overline{e^{ikx}} = p e^{ikl} + q e^{-ikl}$$

Using the binomial expansion, one then obtains

$$\begin{aligned} Q^N(k) &= (p e^{ikl} + q e^{-ikl})^N \\ &= \sum_{n=0}^N \frac{N!}{n!(N-n)!} (p e^{ikl})^n (q e^{-ikl})^{N-n} \\ &= \sum_{n=0}^N \frac{N!}{n!(N-n)!} p^n q^{N-n} e^{ikl(2n-N)} \end{aligned}$$

Thus (1·10·8) yields

$$\begin{aligned} \sigma(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ikx} Q^N(k) \\ &= \sum_{n=0}^N \frac{N!}{n!(N-n)!} p^n q^{N-n} \left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik[(2n-N)l-x]} \right\} \\ \text{or } \sigma(x) &= \sum_{n=0}^N \frac{N!}{n!(N-n)!} p^n q^{N-n} \delta[x - (2n - N)l] \quad (1 \cdot 10 \cdot 9) \end{aligned}$$

This says that the probability density $\sigma(x)$ vanishes unless

$$x = (2n - N)l, \quad \text{where } n = 0, 1, 2, \dots, N$$

The probability $P(2n - N)$ of finding the particle at such a position is then given by

$$P(2n - N) = \int_{(2n-N)l-\epsilon}^{(2n-N)l+\epsilon} \sigma(x) dx = \frac{N!}{n!(N-n)!} p^n q^{N-n}$$

where ϵ is some sufficiently small quantity; i.e., P is given by the coefficient of the corresponding δ function in (1·10·9). Thus we regain the result of Sec. 1·2 as a special case of the present more general formulation and without the need to use any combinatorial reasoning.

*1·11 Probability distribution for large N

We consider the integrals (1·10·7) and (1·10·8) which solve the problem of finding $\sigma(x)$ and ask what approximations become appropriate when N is large. The argument used here will be similar to the method detailed in Appendix A·6 to derive Stirling's formula.

The integrand in (1 · 10 · 7) contains the factor e^{iks} , which is an oscillatory function of s and oscillates the more rapidly with increasing magnitude of k . Hence the quantity $Q(k)$ given by the integral in (1 · 10 · 7) tends in general to be increasingly small as k becomes large. (See remark below.) If Q is raised to a large power N , it thus follows that $Q^N(k)$ tends to decrease very rapidly with increasing k . To compute $\Phi(x)$ by Eq. (1 · 10 · 8), a knowledge of $Q^N(k)$ for small values of k is then sufficient for calculating the integral, since for large values of k the contribution of $Q^N(k)$ to this integral is negligibly small. But for small values of k , it should be possible to approximate $Q^N(k)$ by a suitable expansion in powers of k . Since $Q^N(k)$ is a rapidly varying function of k , it is preferable (as in Sec. 1 · 5) to seek the more readily convergent power series expansion of its slowly varying logarithm $\ln Q^N(k)$.

Remark To the extent that $w(s)$ varies slowly over a period of oscillation, the integral $Q(k) = \int ds e^{iks} w(s) \approx 0$. The reason is that in any range $a < s < b$ in which w varies slowly so that $|dw/ds|(b-a) \ll w$, but which contains many oscillations so that $(b-a)k \gg 1$, the integral

$$\int_a^b ds e^{iks} w(s) \approx w(a) \int_a^b ds e^{iks} \approx 0$$

Combining these two inequalities one can say that

$$\int_{-\infty}^{\infty} ds e^{iks} w(s) \approx 0$$

to the extent that k is large enough so that everywhere

$$\left| \frac{dw}{ds} \right| \frac{1}{k} \ll w$$

↓

The actual calculation is straightforward. We want first to compute $Q(k)$ for small values of k . Expanding e^{iks} in Taylor's series, Eq. (1 · 10 · 7) becomes

$$Q(k) = \int_{-\infty}^{\infty} ds w(s) e^{iks} = \int_{-\infty}^{\infty} ds w(s)(1 + iks - \frac{1}{2}k^2s^2 + \dots) \quad (1 \cdot 11 \cdot 1)$$

$$\text{or} \quad Q(k) = 1 + i\bar{s}k - \frac{1}{2}\bar{s}^2k^2. \quad (1 \cdot 11 \cdot 1)$$

$$\text{where} \quad \bar{s}^n \equiv \int_{-\infty}^{\infty} ds w(s)s^n \quad (1 \cdot 11 \cdot 2)$$

is a constant which represents the usual definition of the n th moment of s . Here we assume that $|w(s)| \rightarrow 0$ rapidly enough as $|s| \rightarrow \infty$ so that these moments are finite. Hence (1 · 11 · 1) yields

$$\ln Q^N(k) = N \ln Q(k) = N \ln [1 + i\bar{s}k - \frac{1}{2}\bar{s}^2k^2 \dots] \quad (1 \cdot 11 \cdot 3)$$

Using the Taylor's series expansion valid for $y \ll 1$,

$$\ln(1 + y) = y - \frac{1}{2}y^2 \dots$$

Eq. (1 · 11 · 3) becomes, up to terms quadratic in k ,

$$\begin{aligned}\ln Q^N &= N[i\bar{s}k - \frac{1}{2}s^2k^2 - \frac{1}{2}(i\bar{s}k)^2 \dots] \\ &= N[i\bar{s}k - \frac{1}{2}(s^2 - \bar{s}^2)k^2 \dots] \\ &= N[i\bar{s}k - \frac{1}{2}(\Delta s)^2 k^2 \dots]\end{aligned}$$

where

$$\overline{(\Delta s)^2} \equiv \bar{s}^2 - \bar{s}^2 \quad (1 \cdot 11 \cdot 4)$$

Hence we obtain

$$Q^N(k) = e^{iN\bar{s}k - \frac{1}{2}N(\Delta s)^2 k^2} \quad (1 \cdot 11 \cdot 5)$$

Thus (1 · 10 · 8) becomes

$$\mathcal{P}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{i(N\bar{s}-x)k - \frac{1}{2}N(\Delta s)^2 k^2} \quad (1 \cdot 11 \cdot 6)$$

The integral here is of the following form, where a is real and positive:

$$\begin{aligned}\int_{-\infty}^{\infty} du e^{-au^2+bu} &= \int_{-\infty}^{\infty} du e^{-a[u^2-(b/a)u]} \\ &= \int_{-\infty}^{\infty} du e^{-a(u-b/2a)^2+b^2/4a} \quad \text{by completing the square} \\ &= e^{b^2/4a} \int_{-\infty}^{\infty} dy e^{-ay^2} \quad \text{by putting } y = u - \frac{b}{2a} \\ &= e^{b^2/4a} \sqrt{\frac{\pi}{a}} \quad \text{by (A · 4 · 2)}\end{aligned}$$

Thus

$$\int_{-\infty}^{\infty} du e^{-au^2+bu} = \sqrt{\frac{\pi}{a}} e^{b^2/4a} \quad (1 \cdot 11 \cdot 7)$$

Applying this integral formula to (1 · 11 · 6) we get, with $b = i(N\bar{s} - x)$ and $a = \frac{1}{2}N(\Delta s)^2$, the result

$$\mathcal{P}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \quad (1 \cdot 11 \cdot 8)$$

where

$$\left. \begin{aligned}\mu &\equiv N\bar{s} \\ \sigma^2 &\equiv N(\Delta s)^2\end{aligned}\right\} \quad (1 \cdot 11 \cdot 9)$$

Thus the distribution has the Gaussian form previously encountered in Sec. 1 · 6. Note, however, the extreme generality of this result. *No matter what* the probability distribution $w(s)$ for each step may be, as long as the steps are statistically independent and $w(s)$ falls off rapidly enough as $|s| \rightarrow \infty$, the total displacement x will be distributed according to the Gaussian law if N is sufficiently large. This very important result is the content of the so-called “central limit theorem,” probably the most famous theorem in mathematical probability theory.* The generality of the result also accounts for the fact that so many phenomena in nature (e.g., errors in measurement) obey approximately a Gaussian distribution.

* A proof of the theorem with attention to fine points of mathematical rigor can be found in A. I. Khinchin, “Mathematical Foundations of Statistical Mechanics,” p. 166, Dover Publications, New York, 1949.

We already showed that for the Gaussian distribution (1·6·4)

$$\text{and } \left. \begin{array}{l} \bar{x} = \mu \\ (\Delta x)^2 = \sigma^2 \end{array} \right\}$$

Hence (1·11·9) implies

$$\text{and } \left. \begin{array}{l} \bar{x} = N\bar{s} \\ (\Delta x)^2 = N(\Delta s)^2 \end{array} \right\} \quad (1·11·10)$$

which agree with the results obtained from our general moment calculations (1·9·3) and (1·9·12).

SUGGESTIONS FOR SUPPLEMENTARY READING

Probability theory

- F. Mosteller, R. E. K. Rourke, and G. B. Thomas: "Probability and Statistics," Addison-Wesley Publishing Company, Reading, Mass., 1961. (An elementary introduction.)
- W. Feller: "An Introduction to Probability Theory and its Applications," 2d ed., John Wiley & Sons, Inc., New York, 1959.
- H. Cramer: "The Elements of Probability Theory," John Wiley & Sons, Inc., New York, 1955.

Random walk problem

- S. Chandrasekhar: "Stochastic Problems in Physics and Astronomy, *Rev. Mod. Phys.*, vol. 15, pp. 1-89 (1943). This article is also reprinted in M. Wax, "Selected Papers on Noise and Stochastic Processes," Dover Publications, New York, 1954.
- R. B. Lindsay: "Introduction to Physical Statistics," chap. 2, John Wiley & Sons, Inc., New York, 1941. (An elementary discussion of the random walk and related physical problems.)

PROBLEMS

- 1.1 What is the probability of throwing a total of 6 points or less with three dice?
- 1.2 Consider a game in which six true dice are rolled. Find the probability of obtaining
- (a) exactly one ace
 - (b) at least one ace
 - (c) exactly two aces
- 1.3 A number is chosen at random between 0 and 1. What is the probability that exactly 5 of its first 10 decimal places consist of digits less than 5?
- 1.4 A drunk starts out from a lamppost in the middle of a street, taking steps of equal length either to the right or to the left with equal probability. What is the probability that the man will again be at the lamppost after taking N steps

- (a) if N is even?
 (b) if N is odd?
- 1.5** In the game of Russian roulette (*not* recommended by the author), one inserts a single cartridge into the drum of a revolver, leaving the other five chambers of the drum empty. One then spins the drum, aims at one's head, and pulls the trigger.
- What is the probability of being still alive after playing the game N times?
 - What is the probability of surviving $(N - 1)$ turns in this game and then being shot the N th time one pulls the trigger?
 - What is the mean number of times a player gets the opportunity of pulling the trigger in this macabre game?
- 1.6** Consider the random walk problem with $p = q$ and let $m = n_1 - n_2$ denote the net displacement to the right. After a total of N steps, calculate the following mean values: \bar{m} , $\bar{m^2}$, $\bar{m^3}$, and $\bar{m^4}$.
- 1.7** Derive the binomial distribution in the following algebraic way, which does not involve any explicit combinatorial analysis. One is again interested in finding the probability $W(n)$ of n successes out of a total of N independent trials. Let $w_1 \equiv p$ denote the probability of a success, $w_2 = 1 - p = q$ the corresponding probability of a failure. Then $W(n)$ can be obtained by writing

$$W(n) = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^2 \cdots \sum_{m=1}^2 w_i w_j w_k \cdots w_m \quad (1)$$

Here each term contains N factors and is the probability of a particular combination of successes and failures. The sum over all combinations is then to be taken only over those terms involving w_1 exactly n times, i.e., only over those terms involving w_1^n .

By rearranging the sum (1), show that the unrestricted sum can be written in the form

$$W(n) = (w_1 + w_2)^N .$$

Expanding this by the binomial theorem, show that the sum of all terms in (1) involving w_1^n , i.e., the desired probability $W(n)$, is then simply given by the one binomial expansion term which involves w_1^n .

- 1.8** Two drunks start out together at the origin, each having equal probability of making a step to the left or right along the x axis. Find the probability that they meet again after N steps. It is to be understood that the men make their steps simultaneously. (It may be helpful to consider their relative motion.)
- 1.9** The probability $W(n)$ that an event characterized by a probability p occurs n times in N trials was shown to be given by the binomial distribution

$$W(n) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n} \quad (1)$$

Consider a situation where the probability p is small ($p \ll 1$) and where one is interested in the case $n \ll N$. (Note that if N is large, $W(n)$ becomes very small if $n \rightarrow N$ because of the smallness of the factor p^n when $p \ll 1$. Hence $W(n)$ is indeed only appreciable when $n \ll N$.) Several approximations can then be made to reduce (1) to simpler form.

- (a) Using the result $\ln(1 - p) \approx -p$, show that $(1 - p)^{N-n} \approx e^{-Np}$.
 (b) Show that $N!/(N - n)! \approx N^n$.
 (c) Hence show that (1) reduces to

$$W(n) = \frac{\lambda^n}{n!} e^{-\lambda} \quad (2)$$

where $\lambda = Np$ is the mean number of events. The distribution (2) is called the "Poisson distribution."

- 1.10** Consider the Poisson distribution of the preceding problem.

- (a) Show that it is properly normalized in the sense that $\sum_{n=0}^N W_n = 1$.

(The sum can be extended to infinity to an excellent approximation, since W_n is negligibly small when $n \gtrsim N$.)

- (b) Use the Poisson distribution to calculate \bar{n} .

- (c) Use the Poisson distribution to calculate $(\Delta n)^2 \equiv \overline{(n - \bar{n})^2}$.

- 1.11** Assume that typographical errors committed by a typesetter occur completely at random. Suppose that a book of 600 pages contains 600 such errors. Use the Poisson distribution to calculate the probability

- (a) that a page contains no errors

- (b) that a page contains at least three errors

- 1.12** Consider the α particles emitted by a radioactive source during some time interval t . One can imagine this time interval to be subdivided into many small intervals of length Δt . Since the α particles are emitted at random times, the probability of a radioactive disintegration occurring during any such time Δt is completely independent of whatever disintegrations occur at other times. Furthermore, Δt can be imagined to be chosen small enough so that the probability of more than one disintegration occurring in a time Δt is negligibly small. This means that there is some probability p of one disintegration occurring during a time Δt (with $p \ll 1$, since Δt was chosen small enough) and probability $1 - p$ of no disintegration occurring during this time. Each such time interval Δt can then be regarded as an independent trial, there being a total of $N = t/\Delta t$ such trials during a time t .

- (a) Show that the probability W_n of n disintegrations occurring in a time t is given by a Poisson distribution.

- (b) Suppose that the strength of the radioactive source is such that the mean number of disintegrations per minute is 24. What is the probability of obtaining n counts in a time interval of 10 seconds? Obtain numerical values for all integral values of n from 0 to 8.

- 1.13** A metal is evaporated in vacuum from a hot filament. The resultant metal atoms are incident upon a quartz plate some distance away and form there a thin metallic film. This quartz plate is maintained at a low temperature so that any metal atom incident upon it sticks at its place of impact without further migration. The metal atoms can be assumed equally likely to impinge upon any element of area of the plate.

- If one considers an element of substrate area of size b^2 (where b is the metal atom diameter), show that the number of metal atoms piled up on this area should be distributed approximately according to a Poisson distribution. Suppose that one evaporates enough metal to form a film of mean thickness corresponding to 6 atomic layers. What fraction of the substrate area is then not

covered by metal at all? What fraction is covered, respectively, by metal layers 3 atoms thick and 6 atoms thick?

- 1.14** A penny is tossed 400 times. Find the probability of getting 215 heads. (Suggestion: use the Gaussian approximation.)
- 1.15** A set of telephone lines is to be installed so as to connect town *A* to town *B*. The town *A* has 2000 telephones. If each of the telephone users of *A* were to be guaranteed instant access to make calls to *B*, 2000 telephone lines would be needed. This would be rather extravagant. Suppose that during the busiest hour of the day each subscriber in *A* requires, on the average, a telephone connection to *B* for two minutes, and that these telephone calls are made at random. Find the minimum number *M* of telephone lines to *B* which must be installed so that at most only 1 percent of the callers of town *A* will fail to have immediate access to a telephone line to *B*. (Suggestion: approximate the distribution by a Gaussian distribution to facilitate the arithmetic.)
- 1.16** Consider a gas of N_0 noninteracting molecules enclosed in a container of volume V_0 . Focus attention on any subvolume V of this container and denote by N the number of molecules located within this subvolume. Each molecule is equally likely to be located anywhere within the container; hence the probability that a given molecule is located within the subvolume V is simply equal to V/V_0 .
- (a) What is the mean number \bar{N} of molecules located within V ? Express your answer in terms N_0 , V_0 , and V .
- (b) Find the relative dispersion $(\bar{N} - \bar{N})^2/\bar{N}^2$ in the number of molecules located within V . Express your answer in terms of \bar{N} , V , and V_0 .
- (c) What does the answer to part (b) become when $V \ll V_0$?
- (d) What value should the dispersion $(\bar{N} - \bar{N})^2$ assume when $V \rightarrow V_0$? Does the answer to part (b) agree with this expectation?
- 1.17** Suppose that in the preceding problem the volume V under consideration is such that $0 \ll V/V_0 \ll 1$. What is the probability that the number of molecules in this volume is between N and $N + dN$?
- 1.18** A molecule in a gas moves equal distances l between collisions with equal probability in any direction. After a total of N such displacements, what is the mean square displacement \bar{R}^2 of the molecule from its starting point?
- 1.19** A battery of total emf V is connected to a resistor R ; as a result an amount of power $P = V^2/R$ is dissipated in this resistor. The battery itself consists of N individual cells connected in series so that V is just equal to the sum of the emf's of all these cells. The battery is old, however, so that not all cells are in perfect condition. Thus there is only a probability p that the emf of any individual cell has its normal value v ; and a probability $1 - p$ that the emf of any individual cell is zero because the cell has become internally shorted. The individual cells are statistically independent of each other. Under these conditions, calculate the *mean* power \bar{P} dissipated in the resistor, expressing the result in terms of N , v , and p .
- 1.20** Consider N similar antennas emitting linearly polarized electromagnetic radiation of wavelength λ and velocity c . The antennas are located along the x axis at a separation λ from each other. An observer is located on the x axis at a great distance from the antennas. When a *single* antenna radiates, the observer measures an *intensity* (i.e., mean-square electric-field amplitude) equal to I .
- (a) If all the antennas are driven in phase by the same generator of frequency $\nu = c/\lambda$, what is the total intensity measured by the observer?

(b) If the antennas all radiate at the same frequency $\nu = c/\lambda$ but with completely random phases, what is the mean intensity measured by the observer? (Hint: Represent the amplitudes by vectors, and deduce the observed intensity from the resultant amplitude.)

- 1.21** Radar signals have recently been reflected from the planet Venus. Suppose that in such an experiment a pulse of electromagnetic radiation of duration τ is sent from the earth toward Venus. A time t later (which corresponds to the time necessary for light to go from the earth to Venus and back again) the receiving antenna on the earth is turned on for a time τ . The returning echo ought then to register on the recording meter, placed at the output of the electronic equipment following the receiving antenna, as a very faint signal of definite amplitude a_s . But a fluctuating random signal (due to the inevitable fluctuations in the radiation field in outer space and due to current fluctuations always existing in the sensitive receiving equipment itself) also registers as a signal of amplitude a_n on the recording meter. This meter thus registers a total amplitude $a = a_s + a_n$.

Although $\bar{a}_n = 0$ on the average, since a_n is as likely to be positive as negative, there is considerable probability that a_n attains values considerably in excess of a_s ; i.e., the root-mean-square amplitude $(\bar{a}_n^2)^{\frac{1}{2}}$ can be considerably greater than the signal a_s of interest. Suppose that $(\bar{a}_n^2)^{\frac{1}{2}} = 1000 a_s$. Then the fluctuating signal a_n constitutes a background of "noise" which makes observation of the desired echo signal essentially impossible.

On the other hand, suppose that N such radar pulses are sent out in succession and that the total amplitudes a picked up at the recording equipment after each pulse are all added together before being displayed on the recording meter. The resulting amplitude must then have the form $A = A_s + \bar{A}_n$, where A_n represents the resultant noise amplitude (with $\bar{A}_n = 0$) and $\bar{A} = A_s$ represents the resultant echo-signal amplitude. How many pulses must be sent out before $(\bar{A}_n^2)^{\frac{1}{2}} = A_s$, so that the echo signal becomes detectable?

- 1.22** Consider the random walk problem in one dimension, the probability of a displacement between s and $s + ds$ being

$$w(s) ds = (2\pi\sigma^2)^{-\frac{1}{2}} e^{-(s-l)^2/2\sigma^2}$$

After N steps,

(a) What is the mean displacement \bar{x} from the origin?

(b) What is the dispersion $(x - \bar{x})^2$?

- 1.23** Consider the random walk problem for a particle in one dimension. Assume that in each step its displacement is always positive and equally likely to be anywhere in the range between $l - b$ and $l + b$ where $b < l$. After N steps, what is

(a) the mean displacement \bar{x} ?

(b) the dispersion $(x - \bar{x})^2$?

- 1.24** (a) A particle is equally likely to lie anywhere on the circumference of a circle. Consider as the z axis any straight line in the plane of the circle and passing through its center. Denote by θ the angle between this z axis and the straight line connecting the center of the circle to the particle. What is the probability that this angle lies between θ and $\theta + d\theta$?

(b) A particle is equally likely to lie anywhere on the surface of a sphere. Consider any line through the center of this sphere as the z axis. Denote by θ the angle between this z axis and the straight line connecting the center of the

sphere to the particle. What is the probability that this angle lies between θ and $\theta + d\theta$?

- 1.25** Consider a polycrystalline sample of $\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$ in an external magnetic field \mathbf{B} in the z direction. The internal magnetic field (in the z direction) produced at the position of a given proton in the H_2O molecule by the neighboring proton is given by $(\mu/a^3)(3 \cos^3 \theta - 1)$ if the spin of this neighboring proton points along the applied field; it is given by $-(\mu/a^3)(3 \cos^2 \theta - 1)$ if this neighboring spin points in a direction opposite to the applied field. Here μ is the magnetic moment of the proton and a is the distance between the two protons, while θ denotes the angle between the line joining the protons and the z axis. In this sample of randomly oriented crystals the neighboring proton is equally likely to be located anywhere on the sphere of radius a surrounding the given proton.

(a) What is the probability $W(b) db$ that the internal field b lies between b and $b + db$ if the neighboring proton spin is parallel to \mathbf{B} ?

(b) What is this probability $W(b) db$ if the neighboring proton spin is equally likely to be parallel or antiparallel to \mathbf{B} ? Draw a sketch of $W(b)$ as a function of b .

(In a nuclear magnetic resonance experiment the frequency at which energy is absorbed from a radio-frequency magnetic field is proportional to the local magnetic field existing at the position of a proton. The answer to part (b) gives, therefore, the shape of the absorption line observed in the experiment.)

- ***1.26** Consider the random walk problem in one dimension and suppose that the probability of a single displacement between s and $s + ds$ is given by

$$w(s) ds = \frac{1}{\pi} \frac{b}{s^2 + b^2} ds$$

Calculate the probability $\Phi(x) dx$ that the total displacement after N steps lies between x and $x + dx$. Does $\Phi(x)$ become Gaussian when N becomes large? If not, does this violate the central limit theorem of Sec. 1.11?

- ***1.27** Consider a very general one-dimensional random walk, where the probability that the i th displacement lies between s_i and $s_i + ds_i$ is given by $w_i(s_i) ds_i$. Here the probability density w_i characterizing each step may be different and thus dependent on i . It is still true, however, that different displacements are statistically independent, i.e., w_i for any one step does not depend on the displacements performed by the particle in any other step. Use arguments similar to those of Sec. 1.11 to show that when the number N of displacements becomes large, the probability $\Phi(x) dx$ that the total displacement lies between x and $x + dx$ will still tend to approach the Gaussian form with a mean value $\bar{x} = \Sigma \bar{s}_i$ and a dispersion $\overline{(\Delta x)^2} = \Sigma \overline{(\Delta s_i)^2}$. This result constitutes a very general form of the central limit theorem.

- ***1.28** Consider the random walk of a particle in three dimensions and let $w(s) d^3s$ denote the probability that its displacement s lies in the range between s and $s + ds$ (i.e., that s_x lies between s_x and $s_x + ds_x$, s_y between s_y and $s_y + ds_y$, and s_z between s_z and $s_z + ds_z$). Let $\Phi(r) d^3r$ denote the probability that the total displacement r of the particle after N steps lies in the range between r and $r + dr$. By generalizing the argument of Sec. 1.10 to three dimensions, show that

$$\Phi(r) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} d^3k e^{-ik \cdot r} Q^N(k)$$

where

$$Q(k) = \int_{-\infty}^{\infty} d^3s e^{ik \cdot s} w(s)$$

- *1.29 (a) Using an appropriate Dirac-delta function, find the probability density $w(s)$ for displacements of uniform length l , but in any random direction of three-dimensional space. (Hint: Remember that the function $w(s)$ must be such that $\iiint w(s) ds = 1$ when integrated over all space.)
- (b) Use the result of part (a) to calculate $Q(\mathbf{k})$. (Perform the integration in spherical coordinates.)
- (c) Using this value of $Q(\mathbf{k})$, compute $\mathcal{P}(\mathbf{r})$ for $N = 3$, thus solving the random walk problem in three dimensions for the case of three steps.

Statistical description of systems of particles

2

NOW THAT we are familiar with some elementary statistical ideas, we are ready to turn our attention to the main subject of this book, the discussion of systems consisting of very many particles. In analyzing such systems, we shall attempt to combine some statistical ideas with our knowledge of the laws of mechanics applicable to the particles. This approach forms the basis of the subject of "statistical mechanics" and is quite similar to that which would be used in discussing a game of chance. To make the analogy explicit consider, for example, a system consisting of 10 dice which are thrown from a cup onto a table in a gambling experiment. The essential ingredients necessary for an analysis of this situation are the following:

1. *Specification of the state of the system:* One needs a detailed method for describing the outcome of each experiment. For example, in this case a specification of the state of the system after the throw requires a statement as to which face is uppermost for each of the 10 dice.

2. *Statistical ensemble:* In principle it may be true that the problem is deterministic in the following sense: if we really knew the initial positions and orientations as well as corresponding velocities of all the dice in the cup at the beginning of the experiment, and if we knew just how the cup is manipulated in the act of throwing the dice, then we could indeed predict the outcome of the experiment by applying the laws of classical mechanics and solving the resulting differential equations. But we do not have any such detailed information available to us. Hence we proceed to describe the experiment in terms of probabilities. That is, instead of focusing our attention on a single experiment, we consider an ensemble consisting of many such experiments, all carried out under similar conditions where 10 dice are thrown from the cup. The outcome of each such experiment will, in general, be different. But we can ask for the *probability* of occurrence of a particular outcome, i.e., we can determine the fraction of cases (in this set of similar experiments) which are characterized by a particular final state of the dice. This procedure shows how the prob-

ability is determined experimentally. Our theoretical aim is to predict this probability on the basis of some fundamental postulates.

3. Basic postulate about *a priori* probabilities: To make theoretical progress, we must introduce some basic postulates. Our knowledge of the physical situation leads us to expect that there is nothing in the laws of mechanics which, for regular dice of uniform density, would result in the preferred appearance uppermost of any one face of a die compared to any other face. Hence we may introduce the postulate that *a priori* (i.e., based on our prior notions as yet unverified by actual observations) the probabilities are equal that any of the six faces of a die land uppermost. This postulate is eminently reasonable and certainly does not contradict any of the laws of mechanics. That the postulate is actually valid can only be decided by making theoretical predictions based on this postulate and checking that these predictions are confirmed by experimental observations. To the extent that such predictions are repeatedly verified, the validity of this postulate can be accepted with increasing confidence.

4. Probability calculations: Once the basic postulate has been adopted, the theory of probability allows the theoretical calculation of the probability of the outcome for any experiment with these dice.

In studying systems consisting of a large number of particles, our considerations will be similar to those used in formulating the preceding problem of several dice.

STATISTICAL FORMULATION OF THE MECHANICAL PROBLEM

2 · 1 Specification of the state of a system

Consider any system of particles, no matter how complicated (e.g., an assembly of weakly interacting harmonic oscillators, a gas, a liquid, an automobile). We know that the particles in any such system (i.e., the electrons, atoms, or molecules composing the system) can be described in terms of the laws of quantum mechanics. Specifically, the system can then be described by a wave function $\psi(q_1, \dots, q_f)$ which is a function of some set of f coordinates (including possible spin variables) required to characterize the system. The number f is the "number of degrees of freedom" of the system. A particular quantum state of the system is then specified by giving the values of some set of f quantum numbers. This description is complete since, if ψ is thus specified at any time t , the equations of motion of quantum mechanics allow prediction of ψ at any other time.

Example 1 Consider a system consisting of a single particle, considered fixed in position, but having a spin $\frac{1}{2}$ (i.e., intrinsic spin angular momentum $\frac{1}{2}\hbar$). In a quantum-mechanical description the state of this particle is specified by the

projection m of its spin along some fixed axis (which we shall choose to call the z axis). The quantum number m can then assume the two values $m = \frac{1}{2}$ or $m = -\frac{1}{2}$; i.e., roughly speaking, one can say that the spin can point either "up" or "down" with respect to the z axis.

—

Example 2 Consider a system consisting of N particles considered fixed in position, but each having spin $\frac{1}{2}$. Here N may be large, say of the order of Avogadro's number $N_a = 6 \times 10^{23}$. The quantum number m of each particle can then assume the two values $\frac{1}{2}$ or $-\frac{1}{2}$. The state of the entire system is then specified by stating the values of the N quantum numbers m_1, \dots, m_N which specify the orientation of the spin of each particle.

Example 3 Consider a system consisting of a one-dimensional simple harmonic oscillator whose position coordinate is x . The possible quantum states of this oscillator can be specified by a quantum number n such that the energy of the oscillator can be expressed as

$$E_n = (n + \frac{1}{2})\hbar\omega$$

where ω is the classical angular frequency of oscillation. Here the quantum number n can assume any integral value $n = 0, 1, 2, \dots$.

Example 4 Consider a system consisting of N weakly interacting one-dimensional simple harmonic oscillators. The quantum state of this system can be specified by the set of numbers n_1, \dots, n_N , where the quantum number n_i refers to the i th oscillator and can assume any value $0, 1, 2, \dots$

Example 5 Consider a system consisting of a single particle (without spin) confined within a rectangular box (so that the particle's coordinates lie within the ranges $0 \leq x \leq L_x$, $0 \leq y \leq L_y$, $0 \leq z \leq L_z$), but otherwise subject to no forces. The wave function ψ of the particle (of mass m) must then satisfy the Schrödinger equation

$$-\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \psi = E\psi \quad (2 \cdot 1 \cdot 1)$$

and, to guarantee confinement of the particle within the box, ψ must vanish at the walls. The wave function having these properties has the form

$$\psi = \sin\left(\pi \frac{n_x x}{L_x}\right) \sin\left(\pi \frac{n_y y}{L_y}\right) \sin\left(\pi \frac{n_z z}{L_z}\right) \quad (2 \cdot 1 \cdot 2)$$

This satisfies (2 · 1 · 1), provided that the energy E of the particle is related to n_x, n_y, n_z by

$$E = \frac{\hbar^2}{2m} \pi^2 \left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2} \right) \quad (2 \cdot 1 \cdot 3)$$

Also $\psi = 0$ properly when $x = 0$ or $x = L_x$, when $y = 0$ or $y = L_y$, and when $z = 0$ or $z = L_z$, provided that the three numbers n_x, n_y, n_z assume any discrete integral values. The state of the particle is then specified by stating the values assumed by these three quantum numbers n_x, n_y, n_z , while Eq. (2 · 1 · 3) gives the corresponding value of the quantized energy.

Comments on the classical description Atoms and molecules are properly described in terms of quantum mechanics. Throughout this book our theoretical discussion of systems of many such particles will, therefore, be based consistently on quantum ideas. A description in terms of classical mechanics, although inadequate in general, may nevertheless sometimes be a useful approximation. It is, therefore, worth making a few comments about the specification of the state of a system in classical mechanics.

Let us start with a very simple case—a single particle in one dimension. This system can be completely described in terms of its position coordinate q and its corresponding momentum p . (This specification is complete, since the laws of classical mechanics are such that a knowledge of q and p at any one time permits prediction of the values of q and p at any other time.) It is possible to represent the situation geometrically by drawing cartesian axes labeled by q and p , as shown in Fig. 2·1·1. Specification of q and p is then equivalent to specifying a point in this two-dimensional space (commonly called "phase space"). As the coordinate and momentum of the particle change in time, this representative point moves through this phase space.

In order to describe the situation in terms where the possible states of the particle are countable, it is convenient to subdivide the ranges of the variables q and p into arbitrarily small discrete intervals. For example, one can choose fixed intervals of size δq for the subdivision of q , and fixed intervals of size δp for the subdivision of p . Phase space is then subdivided into small cells of equal size and of two-dimensional volume (i.e., area)

$$\delta q \delta p = h_0$$

where h_0 is some small constant having the dimensions of angular momentum. The state of the system can then be specified by stating that its coordinate lies

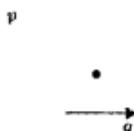


Fig. 2·1·1 Classical phase space for a single particle in one dimension.

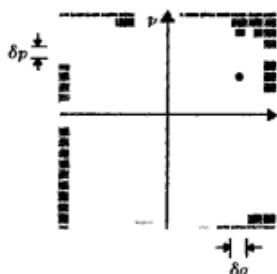


Fig. 2·1·2 The phase of Fig. 2·1·1 is here shown subdivided into equal cells of volume $\delta q \delta p = h_0$.

in some interval between q and $q + \delta q$ and that its momentum lies in some interval between p and $p + \delta p$, i.e., by stating that the representative point (q, p) lies in a particular cell of phase space. The specification of the state of the system clearly becomes more precise as one decreases the size chosen for the cells into which phase space has been divided, i.e., as one decreases the magnitude chosen for h_0 . Of course, h_0 can be chosen arbitrarily small in this classical description.

It may be remarked that quantum theory imposes a limitation on the accuracy with which a simultaneous specification of a coordinate q and its corresponding momentum p is possible. Indeed, this limitation is expressed by the Heisenberg uncertainty principle, which states that the uncertainties δq and δp in these two quantities are such that $\delta q \delta p \gtrsim \hbar$, where \hbar is Planck's constant (divided by 2π). Thus subdivision of the phase space into cells of volume less than \hbar is really physically meaningless; i.e., a choice of $h_0 < \hbar$ would lead to a specification of the system more precise than is allowed by quantum theory.

The generalization of the above remarks to an arbitrarily complex system is immediate. Such a system can be described by some set of f coordinates q_1, q_2, \dots, q_f and f corresponding momenta p_1, p_2, \dots, p_f , i.e., by a total of $2f$ parameters. The number f of independent coordinates needed for the description of the system is called the "number of degrees of freedom" of the system. (For example, if we are dealing with a system of N point particles, then each particle is characterized by three position coordinates so that $f = 3N$.) The set of numbers $\{q_1, \dots, q_f, p_1, \dots, p_f\}$ can again be regarded as a "point" in a "phase space" of $2f$ dimensions in which each cartesian-coordinate axis is labeled by one of the coordinates or momenta. (Except for the fact that this space is not so readily visualized by our provincial three-dimensional minds, this space is completely analogous to the two-dimensional diagram of Fig. 2-1-1). Once again this space can be subdivided into little cells (of volume $\delta q_1 \cdots \delta q_f \delta p_1 \cdots \delta p_f = h_0^f$ if one always chooses the interval of subdivision δq_k of the k th coordinate and δp_k of the k th momentum such that $\delta q_k \delta p_k = h_0$). The state of the system can then again be specified by stating in which particular range, or cell in phase space, the coordinates $q_1, \dots, q_f, p_1, \dots, p_f$ of the system can be found.

Summary The microscopic state, or "microstate," of a system of particles can be simply specified in the following way:

Enumerate in some convenient order, and label with some index r ($r = 1, 2, 3, \dots$), all the possible quantum states of the system. The state of the system is then described by specifying the particular state r in which the system is found.

If it is desired to use the approximation of classical mechanics, the situation is quite analogous. After the phase space for the system has been subdivided into suitable cells of equal size, one can enumerate these cells in some

convenient order and label them with some index r ($r = 1, 2, 3, \dots$). The state of the system is then described by specifying the particular cell r in which the representative point of the system is located.

The quantum mechanical and classical descriptions are thus very similar, a cell in phase space in the classical discussion being analogous to a quantum state in the quantum-mechanical discussion.

2 · 2 Statistical ensemble

In principle the problem of a system consisting of many particles is completely deterministic in the sense that a complete specification of the quantum state ψ of the system at any one time allows calculation of all physical quantities, as well as prediction of the state ψ of the system at all other times. (Similarly, in classical mechanics, complete specification of the state of the system by all of its coordinates q and momenta p at any one time allows calculation of all physical quantities, as well as prediction of the coordinates and momenta at all other times.) But in general we neither have available to us, nor are we interested in, such a complete specification of the system. Hence we proceed to a discussion of the system in terms of probability concepts. For this purpose we consider *not* an isolated instance of a single system, but instead imagine attention focused on an ensemble consisting of a very large number of identical systems, all prepared subject to whatever conditions are specified as known. The systems in this ensemble will, in general, be in different states and will, therefore, also be characterized by different macroscopic parameters (e.g., by different values of pressure or magnetic moment). But we can ask for the probability of occurrence of a particular value of such a parameter, i.e., we can determine the fraction of cases in the ensemble when the parameter assumes this particular value. The aim of theory will be to predict the probability of occurrence in the ensemble of various values of such a parameter on the basis of some basic postulates.

Example Consider a system of three fixed particles, each having spin $\frac{1}{2}$ so that each spin can point either up or down (i.e., along or opposite some direction chosen as the z axis). Each particle has a magnetic moment along the z axis of μ when it points up, and $-\mu$ when it points down. The system is placed in an external magnetic field H pointing along this z axis.

The state of the particle i can be specified by its magnetic quantum number m_i , which can assume the two values $m_i = \pm \frac{1}{2}$. The state of the whole system is specified by giving the values of the three quantum numbers m_1, m_2, m_3 . A particle has energy $-\mu H$ when its spin points up, and energy μH when its spin points down.

We list in the table below all the possible states of the system. We also list some parameters, such as total magnetic moment and total energy, which characterize the system as a whole. (For the sake of brevity $m = \frac{1}{2}$ is denoted simply by +, and $m = -\frac{1}{2}$ by -.)

<i>State index r</i>	<i>Quantum numbers</i> m_1, m_2, m_3	<i>Total magnetic moment</i>	<i>Total energy</i>
1	+++	3μ	$-3\mu H$
2	++-	μ	$-\mu H$
3	+ - +	μ	$-\mu H$
4	- + +	μ	$-\mu H$
5	+ - -	$-\mu$	μH
6	- + -	$-\mu$	μH
7	- - +	$-\mu$	μH
8	- - -	-3μ	$3\mu H$

One usually has available some partial knowledge about the system under consideration. (For example, one might know the total energy and the volume of a gas.) The system can then only be in any of its states which are compatible with the available information about the system. These states will be called the "states accessible to the system." In a statistical description the representative ensemble thus contains only systems all of which are consistent with the specified available knowledge about the system; i.e., the systems in the ensemble must all be distributed over the various accessible states.

Example Suppose that in the previous example of a system consisting of three spins the total energy of the system is known to be equal to $-\mu H$. If this is the only information available, then the system can be in only one of the following three states:

$$(++) \quad (+-) \quad (-+)$$

Of course, we do not know in which of these states the system may actually be, nor do we necessarily know the relative probability of finding the system in any one of these states.

2 · 3 Basic postulates

In order to make theoretical progress, it is necessary to introduce some postulate about the relative probability of finding a system in any of its accessible states. Suppose that the system under consideration is *isolated* and thus cannot exchange energy with its surroundings. The laws of mechanics then imply that the total energy of the system is conserved. Thus it is known

that the system must always be characterized by this value of the energy and that the states accessible to the system must all have this energy. But there are usually a great many states of this kind, and the system can be in any one of them. What can one say about the relative probability of finding the system in any such state?

One can hope to make some general statements in the simple case where the isolated system is in *equilibrium*. Such an equilibrium situation is characterized by the fact that the probability of finding the system in any one state is independent of time (i.e., the representative ensemble is the same irrespective of time). All macroscopic parameters describing the isolated system are then also time-independent. When one considers such an isolated system in equilibrium, the only information one has available about the system is that it must be in one of its accessible states consistent with the constant value of its energy. But there is nothing in the laws of mechanics which would lead one to expect that the system should be found more frequently in one of its accessible states rather than in another. Hence it seems eminently reasonable to *assume* that the system is equally likely to be found in any one of its accessible states. Indeed, one can show explicitly from the laws of mechanics that if one considers a representative ensemble of such isolated systems where these systems are distributed uniformly (i.e., with equal probability) over all their accessible states at *any one time*, then they will remain uniformly distributed over these states *forever*.* This fact shows that such a uniform distribution of systems in the ensemble over their accessible states corresponds indeed to a possible equilibrium situation which does not change in time. It also suggests that there is nothing intrinsic in the laws of mechanics which favors some states at the expense of others, because there exists no tendency to destroy the uniform distribution by populating some states preferentially while depleting other states.

The foregoing considerations suggest that all accessible states of an isolated system have intrinsically the same probability of being occupied by this system. One is thus led to introduce the following fundamental *postulate* of equal a priori probabilities:

- An isolated system in equilibrium is equally likely to be in any of its accessible states.

The same postulate is made in classical mechanics where state refers to a cell in phase space. That is, if phase space is subdivided into small cells of equal size, then an isolated system in equilibrium is equally likely to be in any of its accessible cells.†

This fundamental postulate is eminently reasonable and certainly does not contradict any of the laws of mechanics. Whether the postulate is actually

* This is a consequence of what is called "Liouville's theorem." A proof of this theorem in classical mechanics is given in Appendix A-13 and discussed more fully in R. C. Tolman, "The Principles of Statistical Mechanics," chap. 3, Oxford University Press, Oxford, 1938. A discussion of this theorem in quantum mechanics can be found in chap. 9 of the same book.

† Further comments about the postulate can be found at the end of this section.

valid can, of course, only be decided by making theoretical predictions based on it and by checking whether these predictions are confirmed by experimental observations. A large body of calculations based on this postulate have indeed yielded results in very good agreement with observations. The validity of this postulate can therefore be accepted with great confidence as the basis of our theory.

We illustrate this postulate with a few simple examples.

Example 1 In the previous example of a system of three spins, assume that the system is isolated. Its total energy is then known to have some constant value; suppose that it is known to be equal to $-\mu H$. As already mentioned, the system can then be in any of the following three states

$$(+ + -) \quad (+ - +) \quad (- + +)$$

The postulate asserts that when the system is in equilibrium it is equally likely to be found in any of these three states.

Note, incidentally, that it is *not* true that a given spin is equally likely to point up or down, i.e., to be in any of its two possible states. (There is, of course, no paradox here, since a given spin is not an isolated system, but interacts with the other two spins.) Indeed, it is seen that in the present example it is twice as probable that a given spin points up (is in a state of lower energy) than that it points down (is in a state of higher energy).

Example 2 An example more representative of situations encountered in practice is that of a macroscopic system consisting of N magnetic atoms, where N is of the order of Avogadro's number. If these atoms have spin $\frac{1}{2}$ and are placed in an external magnetic field, the situation is, of course, completely analogous to that of the preceding case of only three spins. But now there exists, in general, an extremely large number of possible states of the system for each specified value of its total energy.

Example 3 Consider a one-dimensional harmonic oscillator of mass m and spring constant κ , and let us discuss it in terms of classical mechanics. Denote the displacement coordinate of the oscillator by x and its linear momentum by p . Phase space is then two dimensional. The energy E of the oscillator is given by

$$E = \frac{p^2}{2m} + \frac{1}{2} \kappa x^2 \quad (2 \cdot 3 \cdot 1)$$

where the first term on the right is its kinetic, the second term its potential energy. For a constant energy E , Eq. (2 3 1) describes an ellipse in phase space, i.e., in the px plane. Suppose one knows that the energy of the oscillator lies in the small range* between E and $E + \delta E$. Then there are still

* The energy of a system can never be physically known to infinite precision (in quantum physics not even in principle, unless one spends an infinite amount of time in the measurement), and working with an infinitely sharply defined energy can also lead to unnecessary conceptual difficulties of a purely mathematical kind.

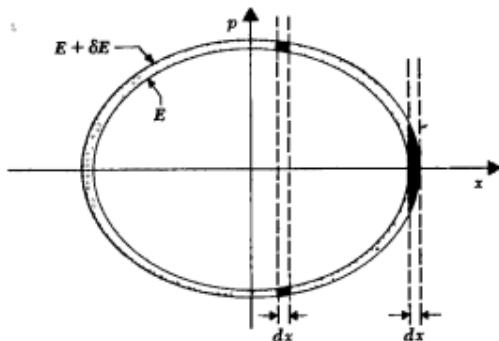


Fig. 2·3·1 Classical phase space for a one-dimensional harmonic oscillator with energy between E and $E + \delta E$. The accessible region of phase space consists of the area lying between the two ellipses.

many cells in phase space contained between the two ellipses corresponding to the respective energies E and $E + \delta E$, i.e., many different corresponding sets of values of x and p are possible for the oscillators in a representative ensemble. If the only information available about the oscillator is that it is in equilibrium with an energy in the specified range, then our statistical postulate asserts that it is equally probable that the oscillator has values of x and p lying within any one of these cells.

Another way of looking at the situation is the following. The time dependence of x and p for the oscillator is, by elementary mechanics, of the form

$$\begin{aligned}x &= A \cos(\omega t + \varphi) \\p &= m\dot{x} = -mA\omega \sin(\omega t + \varphi)\end{aligned}$$

where $\omega = \sqrt{k/m}$, while A and φ are constants. By (2·3·1) the total energy is then

$$E = \frac{m\omega^2}{2} A^2 \sin^2(\omega t + \varphi) + \frac{k}{2} A^2 \cos^2(\omega t + \varphi) = \frac{1}{2} m\omega^2 A^2$$

This is indeed equal to a constant, and the above relation determines the amplitude A in terms of E . But the phase angle φ is still quite arbitrary, depending on unknown initial conditions, and can assume any value in the range $0 < \varphi < 2\pi$. This gives rise to the many possible sets of values of x and p which correspond to the same energy.

Note that a given interval dx corresponds to a larger number of cells (i.e., a larger area) lying between the two ellipses when $x \approx A$ than when $x \approx 0$. Hence it is more probable that an oscillator in the ensemble is found with its position x close to A than close to 0. This result is, of course, also obvious from the fact that near the extremes of its position, where $x \approx A$, the oscillator has small velocity; hence it spends a longer time there than near $x \approx 0$, where it moves rapidly.

The approach to equilibrium Consider a situation where it is known that an isolated system is *not* equally likely to be found in any of the states accessible to it. Our fundamental postulate asserts that this situation cannot be one where equilibrium prevails. Thus one expects the situation to change with time. This means that in the representative statistical ensemble the distribution of systems over the accessible states will change in time; correspondingly, the mean values of various macroscopic parameters describing the system will also change.

Before discussing this nonequilibrium situation in greater detail, it is worth making a few comments about the nature of the states used in our theory to describe an isolated system of many particles. These states are *not* rigorously exact quantum states of the perfectly isolated system with all interactions between particles taken into account.* It would be prohibitively complicated to attempt any such utterly precise description; nor does one have available sufficiently detailed information about a macroscopic system to make such a precise description of any experimental interest. Instead one describes the system in terms of some complete set of approximate quantum states which take into account substantially all of its predominant dynamical features without being rigorously exact. When the system is known to be in such a state at any one time, it will not remain in this state indefinitely. Instead, there exists a finite probability that the system will at some later time be found in some of the other approximate states accessible to it, the transitions to these other states being caused by the presence of small residual interactions between the particles (interactions not taken into account in defining the approximate quantum states of the system).

Suppose then that at some initial time t the system is known to be in some subset of the states actually accessible to it. There are no restrictions which would prevent the system from being found in *any* of its accessible states at some later time since all these states satisfy the conservation of energy and are consistent with the other constraints to which the system is known to be subject; nor is there anything in the laws of mechanics which would make any of these states intrinsically preferable to any other one. It is therefore exceedingly unlikely that the system remains indefinitely in the restricted subset of states in which it finds itself at the initial time t . Instead, the system will in the course of time always make transitions between all its various accessible states as a result of small interactions between its constituent particles. What then is the probability of finding the system in any of these states at some much later time?[†]

To see what happens it is only necessary to consider a statistical ensemble

* If the system were known to be in such an exact eigenstate at any one time, it would remain in this state forever.

† In principle, one could ask more detailed questions about subtle correlations existing between states of the system, i.e., questions about the quantum mechanical phases as well as amplitudes of the relevant wave functions. But it is generally meaningless to seek such a precise description in a theory where completely detailed information about any system is neither available nor of interest.

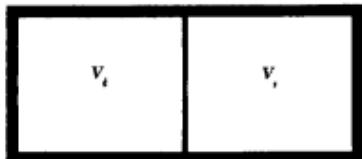
of such systems. Suppose that these systems are initially distributed over their accessible states in some arbitrary way; e.g., that they are only found in some particular subset of these accessible states. The systems in the ensemble will then constantly continue making transitions between the various accessible states, each system passing ultimately through practically all the states in which it can possibly be found. One expects that the net effect of these constant transitions will be analogous to the effect of repeated shufflings of a deck of cards. In the latter case, if one keeps on shuffling long enough, the cards get so mixed up that each one is equally likely to occupy any position in the deck irrespective of how the deck was arranged initially. Similarly, in the case of the ensemble of systems, one expects that ultimately the systems will become randomly (e.g., uniformly) distributed over all their accessible states. The distribution of systems over these states then remains uniform, i.e., it corresponds to a final time-independent equilibrium situation. In other words, one expects that, no matter what the initial conditions may be, an isolated system left to itself will ultimately attain a final equilibrium situation in which it is equally likely to be found in any one of its accessible states. One can look upon the expectation discussed in the preceding sentences as a highly plausible basic hypothesis very well confirmed by experience. From a more fundamental point of view this hypothesis can be regarded as a consequence of the so-called "*H* theorem," which can be established on the basis of the laws of mechanics and certain approximations inherent in a statistical description.*

Example 1 Consider again the very simple example of an isolated system of three spins $\frac{1}{2}$ in a large external magnetic field H . The approximate quantum states of the system can be labeled by the orientation of each spin with respect to this field ("up" or "down"). Suppose that this system has been prepared in such a way that it is known to be in the state $(+ + -)$ at some initial time; the system is then left to itself. Small interactions exist between the spins because the magnetic moment of one spin produces a small field H_m ($H_m \ll H$) with which the moment of some other spin can interact. These interactions between the magnetic moments of the spins bring about transitions in which one spin flips from the "up" direction to the "down" direction while some other spin does the reverse; of course, such a mutual spin-flip leaves the total energy of the system unchanged. The net result is that, after a sufficiently long time, the system will be found with equal probability in any of its three accessible states $(+ + -)$, $(+ - +)$, and $(- + +)$.

Example 2 Another vivid example is shown in Fig. 2·3·2, in which a gas of molecules is originally confined to the left half of a box, the right half being empty. Suppose now that the partition is removed at some initial time t . Immediately after this act, the molecules are certainly not distributed with equal probability over all their accessible states, since the molecules are all localized in the left half of the box whereas the right half, although now per-

* The interested reader can find a discussion of the *H* theorem in Appendix A·12.

Fig. 2·3·2 A system consisting of a box divided by a partition into two equal parts, each of volume V_1 . The left side is filled with gas; the right side is empty.



factly accessible, is empty. But it is clearly fantastically improbable that this situation will prevail for any length of time. Indeed, as a result of collisions with the walls and with each other, the molecules will very quickly redistribute themselves over the entire volume of the box. The final equilibrium situation, where the density of molecules is uniform throughout the entire box, is thus attained quite rapidly.

Note that the preceding comments say nothing about how long one has to wait before the ultimate equilibrium situation is reached. If a system is initially not in equilibrium, the time necessary to attain equilibrium, (the so-called "relaxation time") might be shorter than a microsecond or longer than a century. It all depends on the detailed nature of the interactions between the particles of the particular system and on the resultant rate at which transitions actually occur between the accessible states of this system. The problem of calculating the *rate* of approaching equilibrium is a difficult one. On the other hand, one knows that isolated systems *do* tend to approach equilibrium if one waits long enough. The task of calculating the properties of systems in such *time-independent* situations is then quite straightforward (in principle), since it requires only arguments based on the fundamental statistical postulate of equal a priori probabilities.

***Remark on classical phase space** Phase space is defined in terms of generalized coordinates and *momenta* because it is in terms of these variables that Liouville's theorem holds. In cartesian coordinates it is usually true that $p_i = mv_i$, and phase space could therefore have been defined equally well in terms of coordinates and *velocities*. But more generally, for example in the presence of a magnetic field, the relation between p_i and v_i is more complicated.

***Remark on the fundamental postulate in quantum mechanics** The probability P_r that a quantum-mechanical system is in a state r (which is an eigenstate of the Hamiltonian) is given by $P_r = |a_r|^2$, where a_r is the complex "probability amplitude" which characterizes the state r of the system. Strictly speaking, the fundamental postulate asserts that in equilibrium the probabilities P_r are equal for all accessible states and that the corresponding amplitudes a_r have random phase factors.*

* R. C. Tolman, "The Principles of Statistical Mechanics," Oxford University Press, Oxford, 1938, pp. 349-356.

2 · 4 Probability calculations

The postulate of equal a priori probabilities is fundamental to all statistical mechanics and allows a complete discussion of the properties of systems in equilibrium. In principle, the calculations are very simple. For purposes of illustration, consider a system in equilibrium which is isolated so that its total energy is known to have a constant value in some range between E and $E + \delta E$. To make statistical predictions, we focus attention on an ensemble of such systems, all of which satisfy the condition that their total energy lies in this energy range. Let $\Omega(E)$ denote the *total* number of states of the system in this range. Suppose that there are among these states a certain number $\Omega(E; y_k)$ of states for which some parameter y of the system assumes the value y_k . The parameter might be the magnetic moment of the system, or the pressure exerted by the system, etc. (We label the possible values which y may assume by the index k ; if the possible values which y can assume are continuous instead of discrete, we think of successive values of k as corresponding to values of y which differ by infinitesimal amounts.) Our fundamental postulate tells us that among the states accessible to the system, i.e., among the $\Omega(E)$ states which satisfy the condition that the energy of the system lies in the specified range, all states are equally likely to occur in the ensemble. Hence we can simply write for the probability $P(y_k)$ that the parameter y of the system assumes the value y_k

$$P(y_k) = \frac{\Omega(E; y_k)}{\Omega(E)} \quad (2 \cdot 4 \cdot 1)$$

Also, to calculate the *mean* value of the parameter y for this system, we simply take the average over the systems in the ensemble; i.e.,

$$\bar{y}_k = \frac{\sum_k \Omega(E; y_k) y_k}{\Omega(E)} \quad (2 \cdot 4 \cdot 2)$$

Here the summation over k denotes a sum over all possible values which the parameter y can assume.

Calculations of this kind are in principle quite straightforward. It is true that purely mathematical difficulties of computation may be encountered unless one is dealing with very simple systems. The reason is that, although it is quite easy to count states when there are no restrictions, it may be a formidable problem to pick out only those particular $\Omega(E)$ states which satisfy the condition that they have an energy near some specified value E . Mathematical complications of this sort are, however, not prohibitive and there are methods for overcoming them quite readily.

- Example** Let us illustrate these general comments in the case of the extremely simple example of the system consisting of three spins in equilibrium in a magnetic field H . If the total energy of this system is known to be $-\mu H$,

then the system is equally likely to be in any of the three states

$$(++) \quad (+-+) \quad (-++)$$

Focus attention on one of these spins, say the first. What is the probability P_+ that this spin points up? Since there are two cases where it points up, one has

$$P_+ = \frac{2}{3}$$

What is the mean magnetic moment μ_s (in the $+z$ direction) of such a spin? Since the probability of occurrence of each state of the entire system is $\frac{1}{8}$, one has simply

$$\mu_s = \frac{1}{8}\mu + \frac{1}{8}\mu + \frac{1}{8}(-\mu) = \frac{1}{8}\mu$$

2 · 5 Behavior of the density of states

A macroscopic system is one which has very many degrees of freedom (e.g., a copper block, a bottle of wine, etc.). Denote the energy of the system by E . Subdivide the energy scale into equal small ranges of magnitude δE , the magnitude of δE determining the precision within which one chooses to measure the energy of the system. For a macroscopic system, even a physically very small interval δE contains many possible states of the system. We shall denote by $\Omega(E)$ the number of states whose energy lies between E and $E + \delta E$.

The number of states $\Omega(E)$ depends on the magnitude δE chosen as the subdivision interval in a given discussion. Suppose that δE , while being large compared to the spacing between the possible energy levels of the system, is macroscopically sufficiently small. Then $\Omega(E)$ must be proportional* to δE , i.e., one can write

$$\Omega(E) = \omega(E) \delta E \quad (2 \cdot 5 \cdot 1)$$

where $\omega(E)$ is independent of the size of δE . Thus $\omega(E)$ is a characteristic property of the system which measures the number of states per unit energy range, i.e., the "density of states." Since all statistical calculations involve the counting of states, it is worth examining how sensitively $\Omega(E)$ (or equivalently $\omega(E)$) depends on the energy E of a macroscopic system.

We are not interested in any exact results, but rather in a rough estimate adequate to reveal the essential behavior of Ω as a function of E . A simple argument can then be given along the following lines. Consider a system of f degrees of freedom so that f quantum numbers are required to specify each of its possible states. Let E be the energy of the system measured from its lowest possible energy (i.e., measured from the energy of its quantum-mechanical ground state) and let $\Phi(E)$ denote the total number of possible quantum

* The number of states $\Omega(E)$ must vanish when $\delta E \rightarrow 0$ and must be expressible as a Taylor's series in powers of δE . When δE is sufficiently small, all terms involving higher powers of δE are negligibly small and one is left with an expression of the form (2.5.1).

states of the system which are characterized by energies less than E . Clearly $\Phi(E)$ increases as the energy E increases. Let us examine how *rapidly* $\Phi(E)$ increases.

Consider first *one* typical degree of freedom of the system. Denote by $\Phi_1(\epsilon)$ the *total* number of possible values which can be assumed by the quantum number associated with this particular degree of freedom when it contributes to the system an amount of energy ϵ or less. Again $\Phi_1(\epsilon)$ must clearly increase as ϵ increases (the smallest value of Φ_1 being unity when ϵ has its lowest possible value). If ϵ is not too small, Φ_1 is of the order of $\epsilon/\Delta\epsilon$ where $\Delta\epsilon$ denotes the mean spacing between the possible quantized energies associated with a typical degree of freedom and may itself depend on the magnitude of ϵ . But without getting involved in insignificant details, one can say that Φ_1 ought to increase roughly proportionally to ϵ ; or in symbols that

$$\Phi_1 \propto \epsilon^\alpha \quad (\alpha \approx 1) \quad (2 \cdot 5 \cdot 2)$$

where α is some number of the order of unity.

Let us now return to the whole system having an energy E and described by f quantum numbers $\{s_1, s_2, \dots, s_f\}$. Then the energy ϵ per degree of freedom is of the order of

$$\epsilon \approx \frac{E}{f} \quad (2 \cdot 5 \cdot 3)$$

and, corresponding to this amount of energy or less, there are roughly $\Phi_1(\epsilon)$ possible values which can be assumed by the quantum number describing this degree of freedom. Corresponding to a total energy of E or less of the entire system, there are then approximately $\Phi_1(\epsilon)$ possible values which can be assumed by the quantum number s_1 associated with the first degree of freedom, approximately $\Phi_1(\epsilon)$ possible values which can be assumed by s_2 , approximately $\Phi_1(\epsilon)$ possible values which can be assumed by s_3 , etc. Hence the total number of $\Phi(E)$ of possible sets of values of the f quantum numbers is approximately given by

$$\Phi(E) \approx [\Phi_1(\epsilon)]^f, \quad \text{where } \epsilon = \frac{E}{f} \quad (2 \cdot 5 \cdot 4)$$

This gives the total number of states of the system when it has energy E or less. The number of states $\Omega(E)$ in the range between E and $E + \delta E$ is then

$$\Omega(E) = \Phi(E + \delta E) - \Phi(E) = \frac{\partial \Phi}{\partial E} \delta E \quad (2 \cdot 5 \cdot 5)$$

$$\text{Thus } \Omega(E) \approx f \Phi_1^{f-1} \frac{\partial \Phi_1}{\partial \epsilon} \frac{1}{f} \delta E = \Phi_1^{f-1} \frac{\partial \Phi_1}{\partial \epsilon} \delta E \quad (2 \cdot 5 \cdot 6)$$

When the energy E of the system increases, the number of states Φ_1 per degree of freedom increases slowly and, by (2 · 5 · 3), roughly proportionately to $\epsilon = E/f$. But when one is dealing with a macroscopic system, f is very large—of the order of Avogadro's number—so that $f \approx 10^{24}$. Since the exponent in (2 · 5 · 6) is so very large, it follows that the number of possible states $\Omega(E)$ accessible to the entire system is an *extremely* rapidly increasing function of

the energy E of the system. This is a general characteristic of the number of states $\Omega(E)$, or equivalently of the density of states $\omega = \Omega/\delta E$, of all ordinary macroscopic systems.

The relation (2.5.6) allows one to make some statements about orders of magnitude. Thus it implies that

$$\ln \Omega = (f - 1) \ln \Phi_1 + \ln \left(\frac{\partial \Phi_1}{\partial \epsilon} \delta E \right) \quad (2.5.7)$$

We recall that δE is supposed to be large compared to the spacing between the energy levels of the system. The quantity $(\partial \Phi_1 / \partial \epsilon) \delta E$ is thus of the order of unity in the widest sense; i.e., it is certainly not greater than f nor much less than f^{-1} . Hence its logarithm is certainly of the order of unity in a strict sense (i.e., it lies between $\ln f$ and $-\ln f$ or between 55 and -55, if $f \approx 10^{24}$). On the other hand, the first term of (2.5.7) is of the order of f itself (if the energy ϵ of the system is not so very close to its ground-state energy that $\Phi_1 = 1$ for all degrees of freedom) and is thus fantastically larger than the second term which is only of order $\ln f$. (That is, $\ln f < < f$, if f is very large.) Thus to an excellent approximation, (2.5.7) becomes

$$\ln \Omega \approx f \ln \Phi_1 \quad (2.5.8)$$

and

►
$$\ln \Omega \approx \Theta(f) \quad \text{if } E > 0 \quad (2.5.9)$$

That is, $\ln \Omega$ is of the order of f if the energy of the system is not too close to the energy of its ground state. Furthermore, it follows by (2.5.8) and (2.5.2) that

►
$$\Omega \approx \Phi_1^f \propto E^f \quad (2.5.10)$$

This last relation is intended to be only an order-of-magnitude relation showing roughly how rapidly $\Omega(E)$ varies with the energy E of the system. Thus we have put $\alpha = 1$ for simplicity, since we are not particularly interested in whether the exponent in (2.5.10) should be f , $\frac{1}{2}f$, or any other number of the order of f .

Special case: ideal gas in the classical limit Consider the case of a gas of N identical molecules enclosed in a container of volume V . The energy of this system can be written

$$E = K + U + E_{\text{int}} \quad (2.5.11)$$

Here K denotes the total kinetic energy of translation of the molecules. If the momentum of the center of mass of the i th molecule is denoted by \mathbf{p}_i , then K depends only on these momenta and is given by

$$K = K(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N) = \frac{1}{2m} \sum_{i=1}^N \mathbf{p}_i^2 \quad (2.5.12)$$

The quantity $U = U(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ represents the potential energy of

mutual interaction between the molecules. It depends on the relative separations between the molecules, i.e., on their center-of-mass positions \mathbf{r}_i .

Finally, if the molecules are not monatomic, the atoms of each molecule can also rotate and vibrate relative to its center of mass. Let Q_1, Q_2, \dots, Q_M and P_1, P_2, \dots, P_M denote the coordinates and momenta describing this intramolecular motion. Then E_{int} represents the *total* energy of the system due to such intramolecular motion; it depends only on the internal coordinates Q_i and internal momenta P_i of all the molecules. Of course, if the molecules are monatomic, $E_{\text{int}} = 0$.

A particularly simple case is that where the mutual energy of interaction between the molecules is negligibly small. Then $U \approx 0$ and the molecules are said to form an “ideal gas.” This situation can be achieved physically in the limit where the concentration N/V of the molecules is made sufficiently small, for then the mean separation between molecules becomes so large that their mutual interaction becomes negligibly small.

What is $\Omega(E)$ for such an ideal gas? Let us consider the situation in the classical limit, i.e., under circumstances where the energy E of the gas is much greater than its ground-state energy so that all quantum numbers are large. A description in terms of classical mechanics is then expected to be a good approximation. The number of states $\Omega(E)$ lying between the energies E and $E + \delta E$ is then equal to the number of cells in phase space contained between these energies; i.e., it is proportional to the volume of phase space contained therein. In symbols,

$$\Omega(E) \propto \int_E^{E+\delta E} \int d^3\mathbf{r}_1 \cdots d^3\mathbf{r}_N d^3\mathbf{p}_1 \cdots d^3\mathbf{p}_N dQ_1 \cdots dQ_M dP_1 \cdots dP_M \quad (2 \cdot 5 \cdot 13)$$

Here the integrand is simply the element of volume of phase space where we have used the abbreviations

$$\begin{aligned} d^3\mathbf{r}_i &\equiv dx_i dy_i dz_i \\ d^3\mathbf{p}_i &= dp_{ix} dp_{iy} dp_{iz} \end{aligned}$$

to express three-dimensional volume elements in terms of the three cartesian components of the respective position and momentum vectors. The integration extends over all coordinates and momenta which are such that the total energy given by (2 · 5 · 11) lies in the range between E and $E + \delta E$.

Since $U = 0$ for an ideal gas, the expression E in (2 · 5 · 11) is independent of the center-of-mass positions \mathbf{r}_i of the molecules.* Hence the integration over the position vectors \mathbf{r}_i can be performed immediately. Since each integral over \mathbf{r}_i extends over the volume V of the container, $\int d^3\mathbf{r}_i = V$. But there are N such integrals. Hence (2 · 5 · 13) becomes simply

$$\blacktriangleright \quad \Omega(E) \propto V^N \chi(E) \quad (2 \cdot 5 \cdot 14)$$

* This is true as long as each molecule remains within the container. Of course, the walls of the container serve to confine the molecules within its volume by making $E \rightarrow \infty$ whenever a molecule tends to penetrate into a wall.

$$\text{where } \chi(E) \propto \int_{-E}^{E+\delta E} \int d^3 p_1 \cdots d^3 p_N dQ_1 \cdots dQ_M dP_1 \cdots dP_M \quad (2 \cdot 5 \cdot 15)$$

is independent of V , since neither K nor E_{int} in (2·5·11) depends on the coordinates r_i , so that the integral (2·5·15) does not depend on the volume of the container. The relation (2·5·14) expresses a physically reasonable result for noninteracting molecules. It makes the obvious assertion that if the kinetic energy of each molecule is kept fixed while the volume of the box is doubled, twice as many states become available to each molecule; the number of states accessible to the N molecules is then simply increased by a factor $2 \times 2 \times 2 \times 2 \cdots = 2^N$.

Consider now the particularly simple case where the molecules are monoatomic so that $E_{\text{int}} = 0$ and no intramolecular coordinates Q_i and P_i appear in the problem. Then (2·5·11) reduces simply to the kinetic energy and becomes

$$2mE = \sum_{i=1}^N \sum_{\alpha=1}^3 p_{i\alpha}^2 \quad (2 \cdot 5 \cdot 16)$$

where the sum contains the square of each momentum component $p_{i\alpha}$ of each particle (since $p_i^2 = p_{i1}^2 + p_{i2}^2 + p_{i3}^2$, denoting x, y, z components by 1, 2, 3, respectively). The sum in (2·5·16) thus contains $3N = f$ square terms. For $E = \text{constant}$, Eq. (2·5·16) then describes, in the f -dimensional space of the momentum components, a sphere of radius $R(E) = (2mE)^{\frac{1}{2}}$. Hence $\Omega(E)$, or $\chi(E)$ in (2·5·15), is proportional to the volume of phase space contained in the spherical shell lying between the sphere of radius $R(E)$ and that of slightly larger radius $R(E + \delta E)$ (see Fig. 2·5·1). But the volume of a sphere in f dimensions is proportional to R^f , since it is essentially obtained (just as the volume of a cube in f dimensions) by multiplying f linear dimensions by each other. Thus the total number of state $\Phi(E)$ of energy less than E is proportional

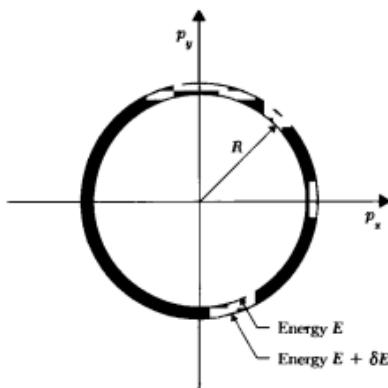


Fig. 2·5·1 Illustration in two dimensions of the “sphere” in momentum space for a single particle (of mass m) moving in two dimensions. Here $(2m)^{-1}(p_x^2 + p_y^2) = E$, the energy of the particle. The radius of the shell is $R = (2mE)^{\frac{1}{2}}$.

to this volume, i.e.,

$$\Phi(E) \propto R^f = (2mE)^{f/2} \quad (2 \cdot 5 \cdot 17)$$

The number of states $\Omega(E)$ lying in the spherical *shell* between energies E and $E + \delta E$ is then given by (2 · 5 · 5), so that

$$\Omega(E) \propto E^{(f/2)-1} \propto E^{(3N/2)-1} \quad (2 \cdot 5 \cdot 18)$$

which is properly proportional to R^{f-1} , i.e., to the *area* of the sphere in phase space. Combining this result with (2 · 5 · 14) one obtains for the classical *monatomic* ideal gas

► $\Omega(E) = BV^N E^{3N/2} \quad (2 \cdot 5 \cdot 19)$

where B is a constant independent of V and E , and where we have neglected 1 compared to N . Note again that since N is of the order of Avogadro's number and thus very large, $\Omega(E)$ is an extremely rapidly increasing function of the energy E of the system.

INTERACTION BETWEEN MACROSCOPIC SYSTEMS

2 · 6 *Thermal interaction*

In describing a macroscopic system it is, in general, possible to specify some macroscopically measurable independent parameters x_1, x_2, \dots, x_n which are known to affect the equations of motion (i.e., to appear in the Hamiltonian) of this system. These parameters are known as the "external parameters" of the system. Examples of such parameters are the applied magnetic or electric fields in which the system is located, or the volume V of the system (e.g., the volume V of the container confining a gas).* The energy levels of the system depend then, of course, on the values of the external parameters. If a particular quantum state r of the system is characterized by an energy E_r , one can thus write the functional relation

$$E_r = E_r(x_1, x_2, \dots, x_n) \quad (2 \cdot 6 \cdot 1)$$

The "macroscopic state," or "macrostate," of the system is defined by specifying the external parameters of the system and any other conditions to which the system is subject. For example, if one deals with an isolated system, the *macrostate* of the system might be specified by stating the values of the external parameters of the system (e.g., the value of the volume of the system) and the value of its constant total energy. The representative ensem-

* The volume V enters the equations of motion because the walls of the container are represented by a potential energy term U which depends on the position coordinates of the particles in such a way that $U \rightarrow \infty$ whenever the position coordinate of a molecule lies outside the available volume, i.e., inside the wall itself. For example, in the case of a single particle, Eq. (2 · 1 · 3) shows explicitly that its energy levels depend on the dimensions of the container; i.e., for a given quantum state, $E \propto V^{-\frac{1}{3}}$ if the volume V of the container is changed without change of shape.

ble for the system is prepared in accordance with the specification of this macrostate; e.g., all systems in the ensemble are characterized by the given values of the external parameters and of the total energy. Of course, corresponding to this given *macrostate*, the system can be in any one of a very large number of possible *microstates* (i.e., quantum states).

Let us now consider two macroscopic systems A and A' which can interact with each other so that they can exchange energy. (Their total energy remains constant, of course, since the combined system $A + A'$ consisting of A and A' is isolated.) In a macroscopic description it is useful to distinguish between two types of possible interactions between such systems. In one case all the external parameters remain fixed so that the possible energy levels of the systems do not change; in the other case the external parameters are changed and some of the energy levels are thereby shifted. We shall discuss these types of interaction in greater detail.

The first kind of interaction is that where the external parameters of the system remain unchanged. This represents the case of purely "thermal interaction."

Example As a trivial illustration, suppose that a bottle of beer is removed from a refrigerator and placed in the trunk of a car, where it remains for a while. No external parameters are changed; i.e., neither the volume of the bottle nor that of the air in the trunk is changed. But energy is transferred from the air in the trunk to the beer and results in a change of the latter's properties (e.g., the beer tastes less good).

As a result of the purely thermal interaction, energy is transferred from one system to the other. In a statistical description where one focuses attention on an ensemble of similar systems ($A + A'$) in interaction (see Fig. 2·6·1), the energy of every A system (or every A' system) does not change by precisely the same amount. One can, however, describe the situation conveniently in terms of the change in *mean* energy of each of the systems. The mean energy transferred from one system to the other as a result of purely thermal interaction is called "heat." More precisely, the change $\Delta\bar{E}$ of the mean energy of system A is called the "heat Q absorbed" by this system; i.e., $Q = \Delta\bar{E}$. This heat can, of course, be negative as well as positive; the quantity $(-Q)$ is called the "heat given off" by the system. Since the combined energy of $(A + A')$ is unchanged, it follows that

$$\Delta\bar{E} + \Delta\bar{E}' = 0 \quad (2 \cdot 6 \cdot 2)$$

where $\Delta\bar{E}$ denotes the change of mean energy of A and $\Delta\bar{E}'$ that of A' . In terms of the definition of heat, one can write correspondingly

$$Q + Q' = 0 \quad \text{or} \quad Q = -Q' \quad (2 \cdot 6 \cdot 3)$$

This merely expresses the conservation of energy by the statement that the

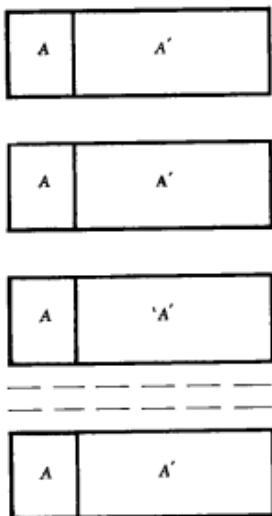


Fig. 2·6·1 Diagram illustrating schematically a representative statistical ensemble of similar isolated systems $A^{(0)}$, each consisting of two systems A and A' in interaction with each other.

heat absorbed by one system must be equal to the heat given off by the other system.

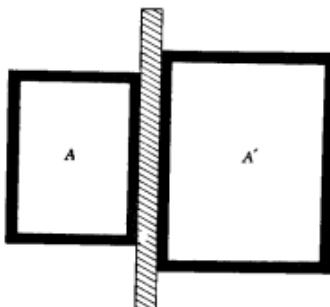
Since the external parameters do not change in a purely thermal interaction, the energy levels of neither system are in any way affected. The change of mean energy of a system comes about because the interaction results in a change in the relative number of systems in the ensemble which are distributed over the fixed energy levels (see Fig. 2·7·3a and b).

2 · 7 Mechanical interaction

A system which cannot interact thermally with any other system is said to be "thermally isolated," (or "thermally insulated"). It is easy to prevent thermal interaction between any two systems by keeping them spatially sufficiently separated, or by surrounding them with "thermally insulating" (sometimes also called "adiabatic") envelopes. These names are applied to an envelope provided that it has the following defining property: if it separates *any* two systems A and A' whose external parameters are fixed and each of which is initially in internal equilibrium, then these systems will remain in their respective equilibrium macrostates indefinitely (see Fig. 2·7·1). This definition implies physically that the envelope is such that no energy transfer is possible through it. (In practice, envelopes made of asbestos or fiberglass might approximate adiabatic envelopes reasonably well.)

When two systems are thermally insulated, they are still capable of interacting with each other through changes in their respective external parameters. This represents the second kind of simple macroscopic interaction, the

Fig. 2·7·1 Two systems A and A' , each consisting of a gas in a container of fixed volume, are separated by a partition. If the partition is adiabatic, each system can independently remain in equilibrium for any value of its mean pressure. If the partition is not adiabatic, the gas pressures will, in general, change in time until they attain mutually compatible values in the final equilibrium situation.



case of purely "mechanical interaction." The systems are then said to exchange energy by doing "macroscopic work" on each other.

Example Consider the situation shown in Fig. 2·7·2 in which a gas is enclosed in a vertical cylinder by a piston of weight w , the piston being thermally insulated from the gas. Initially the piston is clamped in position at a height s_i . When the piston is released, it oscillates for a while and finally comes to rest at a greater height s_f . Let A denote the system consisting of the gas and cylinder, and A' the system consisting of the piston (including the weight) and the earth. Here the interaction involves changes in the external parameters of the system, i.e., a change in the volume of the gas and in the height of the piston. In this process the gas does a net amount of work in lifting the weight.

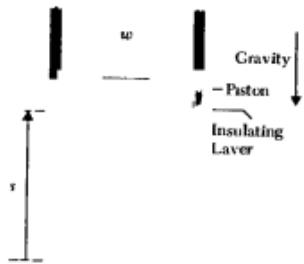


Fig. 2·7·2 A gas contained in a cylinder closed by a piston of weight w . A layer of thermally insulating material (of negligible weight) is attached to the bottom of the piston to separate it from the gas.

In a statistical description one again focuses attention on an ensemble of similar systems ($A + A'$) in interaction. Not every system in the ensemble has its energy changed by exactly the same amount as a result of the change of external parameters, but one can again describe the situation in terms of the change in *mean* energy of the systems. Consider, for example, system A . If the change in its mean energy due to the change of external parameters is denoted by $\Delta_x \bar{E}$, then the "macroscopic work" W done *on* the system is defined as

$$W = \Delta_x \bar{E} \quad (2\cdot7\cdot1)$$

The macroscopic work W done *by* the system is the negative of this and is thus

defined as

$$W \equiv -\nabla \equiv -\Delta_x \bar{E} \quad (2 \cdot 7 \cdot 2)$$

Whenever we shall use the term "work" without further qualifications, we shall be referring to the macroscopic work just defined. The conservation of energy (2 · 6 · 2) is, of course, still valid and can be written in the form

$$W + W' = 0 \quad \text{or} \quad W = -W' \quad (2 \cdot 7 \cdot 3)$$

i.e., the work done by one system must be equal to the work done on the other system.

The mechanical interaction between systems involves changes in the external parameters and hence results in changes of the energy levels of the systems. Note that, even if the energies E_r of different quantum states are originally equal, a change of external parameters usually shifts these energy levels by different amounts for different states r . In general, the change in mean energy of a system depends on how the external parameters are changed and on how rapidly they are changed. When these parameters are changed in some arbitrary way, the energy levels of the possible states of the system change; in addition, transitions are produced between various states of the system. (Thus, if the system is initially in a particular state, it will in general be distributed over many of its states after the parameter change.) Thus the situation may be quite complicated while the parameters are changed and shortly thereafter, even when equilibrium conditions prevail initially and finally.

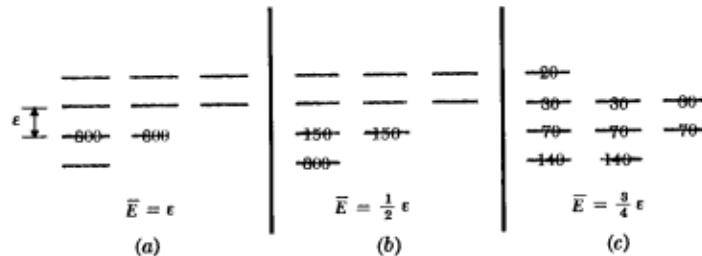


Fig. 2·7·3 Schematic illustration of heat and work. The diagram shows the energy levels (separated by an amount ϵ) of a hypothetical system which can be in any of nine possible states. There are 600 systems in the statistical ensemble and the numbers indicate the number of these systems in each state. (a) Initial equilibrium situation. (b) Final equilibrium situation after the system has given off heat $\frac{1}{2}\epsilon$ to some other system. (c) Final equilibrium situation after the system in (a) has (in some arbitrary way) done work $\frac{1}{4}\epsilon$ on some other system. (The very small numbers used for simplicity in this illustration are, of course, not representative of real macroscopic systems.)

Example The complications can be illustrated by the previous example of Fig. 2 · 7 · 2, or perhaps even more directly by the example of Fig. 2 · 7 · 4. Sup-

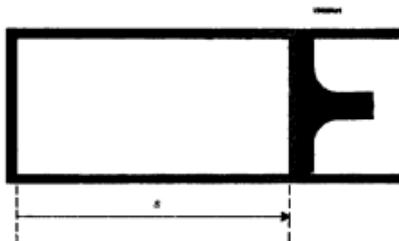


Fig. 2·7·4 A system consisting of a fluid (gas or liquid) contained in a cylinder closed off by a movable piston. The one external parameter is the distance s , which is the distance of the piston from the end wall of the cylinder.

pose that this system is initially in equilibrium, the piston being clamped at a distance s_i from the end wall of the cylinder. The system is then equally likely to be in any of its possible states compatible with the initial value $s = s_i$ and with the initial energy E_i of the system. Suppose that some external device now rapidly moves the piston to a new position $s = s_f$, thus compressing the fluid. In this process the external device does work and the mean energy of the system is increased by some amount $\Delta_x \bar{E}$. But all kinds of pressure nonuniformities and turbulence are also set up in the fluid; during this time the system is *not* equally likely to be in any of its accessible states. Of course, if one keeps s at the value s_f and waits long enough, a new equilibrium situation will again be reached, where there is equal probability that the system is in any of its states compatible with the new value $s = s_f$ and the new mean energy $E_f + \Delta_x \bar{E}$.

Macroscopic work is, nevertheless, a quantity which can be readily measured experimentally. Suppose that in the mechanical interaction between two systems A and A' at least one of them, say A' , is a relatively simple system whose change in mean energy can readily be computed from a change in its external parameters by using considerations based on mechanics. For example, one may know that A' exerts a measurable mean force on A and that the change of external parameters corresponds simply to a definite displacement of the center of mass of A' . Then the mean work W' done by A' on A is immediately obtained as the product of a mean force multiplied by the corresponding displacement; by (2·7·3), the work done by A is then given by $W = -W'$.

Example 1 Consider the previously mentioned illustration of Fig. 2·7·2, where the piston is initially at a height s_i and finally comes to rest at a height s_f . Here the center of mass of the piston is simply displaced by a net amount $(s_f - s_i)$, and one can neglect any change in the internal energy of motion of the molecules in the piston relative to its center of mass. Then the entire change in the energy of the system A' , consisting of the piston and the earth,

is due to the change of potential energy $w(s_f - s_i)$ of the center of mass of the piston in the gravitational field of the earth. Hence it follows that in the process here contemplated the system A, consisting of gas and cylinder, does an amount of work $W = w(s_f - s_i)$ on the system A' .

Example 2 In Fig. 2·7·5 the falling weight w is connected through a string to a paddle wheel, which is thus made to rotate and to churn the liquid in which it is immersed. Suppose that the weight descends with uniform speed a distance s . Then the energy of the system A' , consisting of the weight and the earth, is decreased by an amount ws ; this is then also the work done on the system A consisting of the paddle wheel and the liquid.

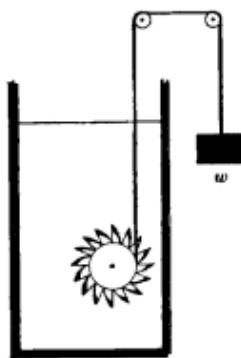


Fig. 2·7·5 *A system consisting of a vessel containing a liquid and a paddle wheel. The falling weight can perform work on the system by rotating the paddle wheel.*

Example 3 Figure 2·7·6 illustrates a similar situation where a battery of emf \mathcal{V} is connected electrically to a resistor immersed in a liquid. When a charge q flows through the circuit, the energy stored in the battery decreases by an amount $q\mathcal{V}$. Hence the battery does an amount of work $q\mathcal{V}$ on the system A consisting of the resistor and the liquid.

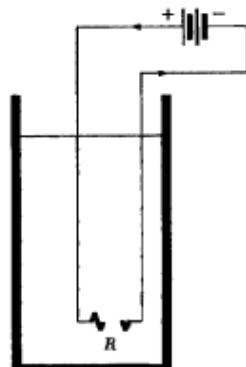


Fig. 2·7·6 *A system consisting of a resistor immersed in a liquid. The battery can perform electrical work on the system by sending current through the resistor.*

2 · 8 General interaction

In the most general case of interaction between two systems their external parameters do *not* remain fixed and the systems are *not* thermally insulated. As a result of such a general interaction the mean energy of a system is changed by some amount $\Delta \bar{E}$, but not all of this change is due to the change of its external parameters. Let $\Delta_x \bar{E} = W$ denote the increase of its mean energy calculable from the change of external parameters (i.e., due to the macroscopic work W done *on* the system). Then the total change in mean energy of the system can be written in the form

$$\Delta \bar{E} = \Delta_x \bar{E} + Q = W + Q \quad (2 \cdot 8 \cdot 1)$$

where the quantity Q thus introduced is simply a measure of the mean energy change *not* due to the change of external parameters. In short, (2 · 8 · 1) defines the quantity Q by the relation

$$Q \equiv \Delta \bar{E} - W = \Delta \bar{E} + W \quad (2 \cdot 8 \cdot 2)$$

where $W = -W$ is the work done *by* the system. The relation (2 · 8 · 2) constitutes the general *definition* of the heat absorbed by a system. When the external parameters are kept fixed, (2 · 8 · 2) reduces, of course, to the definition already introduced in Sec. 2 · 6 for the case of purely thermal interaction.

The relation (2 · 8 · 1) simply splits the total mean energy change into a part W due to mechanical interaction and a part Q due to thermal interaction. One of the fundamental aims of our study will be to gain a better understanding of the relationship between thermal and mechanical interactions. This is the reason for the name "thermodynamics" applied to the classical discipline dealing with such questions.

Note that, by virtue of (2 · 8 · 1), both heat and work have the dimensions of energy and are thus measured in units of ergs or joules.

Example Consider Fig. 2 · 8 · 1 where two gases A and A' are contained in a cylinder and separated by a movable piston.

a. Suppose first that the piston is clamped in a fixed position and that it is thermally insulating. Then the gases A and A' do not interact.

b. If the piston is *not* insulating but is clamped in position, energy will in general flow from one gas to the other (although no macroscopic work gets



Fig. 2 · 8 · 1 Two gases A and A' separated by a piston.

done) and the pressures of the gases will change as a result. This is an example of purely thermal interaction.

c. If the piston is insulating but free to move, then it will in general move so that the volumes and pressures of the gases change, one gas doing mechanical work on the other. This is an example of purely mechanical interaction.

d. Finally, if the piston is noninsulating and free to move, both thermal and mechanical interaction can take place between the two gases A and A' .

If one contemplates infinitesimal changes, the small increment of mean energy resulting from the interaction can be written as the differential $d\bar{E}$. The infinitesimal amount of work done by the system in the process will be denoted by dW ; similarly, the infinitesimal amount of heat absorbed by the system in the process will be denoted by dQ . In terms of the above notation, the definition (2·8·2) becomes for an infinitesimal process

$$dQ \equiv d\bar{E} + dW \quad (2\cdot8\cdot3)$$

Remark The special symbol dW is introduced, instead of W , merely as a convenient notation to emphasize that the work itself is infinitesimal. It does not designate any difference between works. Indeed, the work done is a quantity referring to the interaction *process* itself. Thus it makes no sense to talk of the work in the system before and after the process, or of the difference between these. Similar comments apply to dQ , which again denotes just the infinitesimal amount of heat absorbed in the process, *not* any meaningless difference between heats.

2 · 9 Quasi-static processes

In the last few sections we have considered quite general processes whereby systems can interact with each other. An important, and much simpler, special case is that where a system A interacts with some other system in a process (involving the performance of work, exchange of heat, or both) which is carried out so slowly that A remains arbitrarily close to equilibrium at all stages of the process. Such a process is said to be "quasi-static" for the system A . Just *how* slowly one must proceed to keep a situation quasi-static depends on the time τ (the "relaxation time") that the system requires to attain equilibrium if it is suddenly disturbed. To be slow enough to be quasi-static implies that one proceeds slowly compared to the time τ . For example, if the gas in Fig. 2·7·4 returns to equilibrium within a time $\tau \approx 10^{-8}$ seconds after the distance s is suddenly halved, then a process wherein the piston is moved so as to halve the volume of the gas in 0.1 second can be considered quasi-static to a good approximation.

If the external parameters of a system have values x_1, \dots, x_n , then the energy of the system in a definite quantum state r has some value

$$E_r = E_r(x_1, \dots, x_n) \quad (2 \cdot 9 \cdot 1)$$

When the values of the external parameters are changed, the energy of this state r changes in accordance with the functional relation (2·9·1). In particular, when the parameters are changed by infinitesimal amounts so that $x_\alpha \rightarrow x_\alpha + dx_\alpha$ for each α , then (2·9·1) gives for the corresponding change in energy

$$dE_r = \sum_{\alpha=1}^n \frac{\partial E_r}{\partial x_\alpha} dx_\alpha \quad (2 \cdot 9 \cdot 2)$$

The work dW done by the system when it remains in this particular state r is then defined as

$$dW_r \equiv -dE_r = \Sigma X_{\alpha,r} dx_\alpha \quad (2 \cdot 9 \cdot 3)$$

where we have introduced the definition

$$X_{\alpha,r} \equiv -\frac{\partial E_r}{\partial x_\alpha} \quad (2 \cdot 9 \cdot 4)$$

This is called the "generalized force" (conjugate to the external parameter x_α) in the state r . Note that if x_α denotes a distance, then X_α is simply an ordinary force.

Consider now the statistical description where one focuses attention on an ensemble of similar systems. When the external parameters of the system are changed quasi-statically, then the generalized forces $X_{\alpha,r}$ have at any time well-defined mean values; these are calculable from the distribution of systems in the ensemble corresponding to the equilibrium situation consistent with the values of these external parameters at that time. (For example, if the system is thermally isolated, then the systems in the ensemble are at any time equally likely to be in any of their accessible states which are compatible with the values of the external parameters at that time.) The macroscopic work dW resulting from an infinitesimal quasi-static change of external parameters is then obtained by calculating the decrease in mean energy resulting from this parameter change. Calculating the mean value of (2·9·3) averaged over all accessible states r then gives

$$dW = \sum_{\alpha=1}^n \bar{X}_\alpha dx_\alpha \quad (2 \cdot 9 \cdot 5)$$

where

$$\bar{X}_\alpha \equiv -\frac{\overline{\partial E_r}}{\partial x_\alpha} \quad (2 \cdot 9 \cdot 6)$$

is the *mean* generalized force conjugate to x_α . Here the mean values are to be calculated with the equilibrium distribution of systems in the ensemble corresponding to the external parameter values x_α . The macroscopic work W resulting from a *finite* quasi-static change of external parameters can then be obtained by integration.

Remark If one were dealing with an isolated system in a state r which is an exact stationary quantum state of the entire Hamiltonian (including all interactions between particles), then this system would remain in this state of energy E_r , indefinitely when the external parameters are kept fixed; it would also remain in this state (its energy E_r , varying in accordance with (2·9·1)) when the external parameters are changed infinitely slowly. Thus no transitions to other states would occur in the course of time. But in a statistical description one does not deal with such precisely defined situations. Instead, one contemplates a system which can be in any one of a large number of accessible quantum states which are not exact stationary quantum states of the entire Hamiltonian (including all the interactions), so that transitions between these states do occur. Indeed, if one waits long enough, these transitions bring about the final equilibrium situation where the system, if isolated, is equally likely to be found in any of its accessible states. When the external parameters of the system are changed quasi-statically, a given system in the ensemble does then not always remain in the same state. Instead, there occurs a continual redistribution of systems over their accessible states so as to maintain always a distribution consistent with an equilibrium situation, i.e., a uniform distribution over all accessible states in an ensemble of isolated systems (see Fig. 2·9·1).

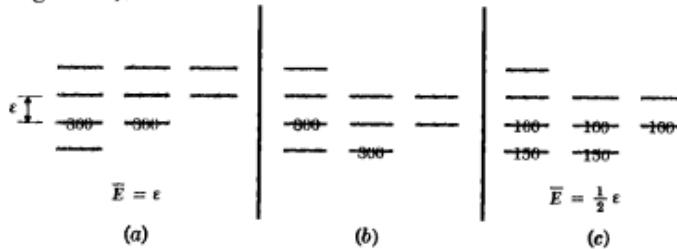


Fig. 2·9·1 Schematic illustration of quasi-static work done by the thermally isolated system of Fig. 2·7·3 as a result of the same change of external parameter as that shown in part (c) of that figure. The diagram shows again the energy levels of this system, and the numbers indicate the number of systems present in each state in the ensemble. (a) Initial equilibrium situation. (b) Hypothetical situation which would prevail after the quasi-static external parameter change if each system remained in its former state. (c) Actual final equilibrium situation resulting from the quasi-static external parameter change, work $\frac{1}{2}\epsilon$ having been done by the system.

2 · 10 Quasi-static work done by pressure

As an important example of quasi-static work, consider the case in which there is only one external parameter of significance, the volume V of the system. Then the work done in changing the volume from V to $V + dV$ can be calcu-

lated from elementary mechanics as the product of a force multiplied by a displacement. Suppose that the system under consideration (see Fig. 2·10·4) is contained in a cylinder. If the system is in state r , let its pressure on the piston of area A be denoted by p_r . The force exerted by the system on the piston is then $p_r A$. The volume of the system is specified by the distance s of the piston from the end wall of the cylinder; thus $V = As$. If the distance s is now changed very slowly by an amount ds , the system remains in the state r and performs an amount of work

$$dW_r = (p_r A) ds = p_r (A ds) = p_r dV \quad (2 \cdot 10 \cdot 1)$$

Since $dW_r = -dE_r$, it follows from this that

$$p_r = -\frac{\partial E_r}{\partial V} \quad (2 \cdot 10 \cdot 2)$$

Thus p_r is the generalized force conjugate to the volume V .

If the volume of the system is changed quasi-statically, the system remains always in internal equilibrium so that its pressure has a well-defined mean value \bar{p} . The macroscopic work done by the system in a quasi-static change of volume is then, by (2·10·1); related to the mean pressure by the relation*

$$dW = \bar{p} dV \quad (2 \cdot 10 \cdot 3)$$

Remark The expression (2·10·3) for the work done is much more general than the derivation based on the simple cylinder would indicate. To show this, consider an arbitrary slow expansion of the system from the volume enclosed by the solid boundary to that enclosed by the dotted boundary in Fig. 2·10·1. If the mean pressure is \bar{p} , the mean force on an element of area dA is $\bar{p} dA$ in the direction of the normal n . If the displacement of this element of area is by an amount ds in the direction making an angle θ with the normal,

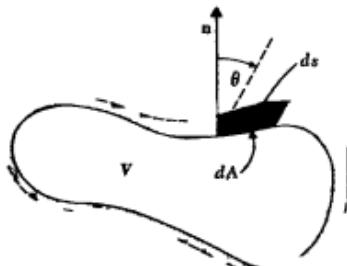


Fig. 2·10·1 Arbitrary expansion of a system of volume V .

* When a system is in one definite state r , the work dW_r and the corresponding pressure p_r may, in general, depend on just how the volume is changed. (For example, if the system is in the shape of a rectangular parallelepiped, the work dW_r may depend on which wall is moved and the force per unit area on different walls may be different.) But after averaging over all the states r , the macroscopic work and mean pressure \bar{p} become insensitive to the precise mode of deformation contemplated for the volume.

then the work done by the pressure on this area is $(\bar{p} dA) ds \cos \theta = \bar{p} dv$, where $dv = (dA ds \cos \theta)$ is the volume of the parallelepiped swept out by the area element dA in its motion through ds . Summing over all the elements of area of the boundary surface gives then for the total work

$$dW = \Sigma \bar{p} dv = \bar{p} \Sigma dv = \bar{p} dV$$

where $dV = \Sigma dv$ is the sum of all the little volumes swept out, i.e., the increase in volume of the total system. Thus one regains (2·10·3).

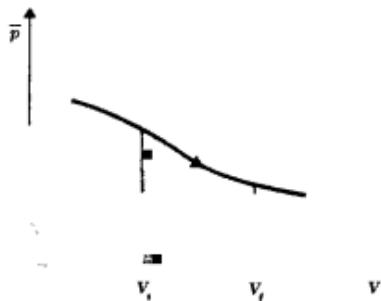


Fig. 2·10·2 Dependence of the mean pressure \bar{p} on the volume V of a system. The shaded area under the curve represents the work done by the system when its volume changes quasi-statically from V_i to V_f .

Suppose that a quasi-static process is carried out in which the volume is changed from V_i to V_f . For example, this process might be carried out in such a way that for all volumes $V_i \leq V \leq V_f$ the mean pressure $\bar{p} = \bar{p}(V)$ assumes the values indicated by the curve of Fig. 2·10·2. In this process the macroscopic work done by the system is given by

$$W_{if} = \int_{V_i}^{V_f} dW = \int_{V_i}^{V_f} \bar{p} dV \quad (2 \cdot 10 \cdot 4)$$

Note that this integral represents geometrically just the shaded area contained below the curve of Fig. 2·10·2.

2 · 11 Exact and “inexact” differentials

The expression (2·8·3) relates the differential $d\bar{E}$ of the energy to the infinitesimal quantities dW and dQ . It is instructive to examine these infinitesimals more closely.

Consider the purely mathematical problem where $F(x,y)$ is some function of the two independent variables x and y . This means that the value of F is determined when the values of x and y are specified. If one goes to a neighboring point corresponding to $x + dx$ and $y + dy$, the function F changes by an amount

$$dF = F(x + dx, y + dy) - F(x, y) \quad (2 \cdot 11 \cdot 1)$$

This can also be written in the form

$$dF = A(x,y) dx + B(x,y) dy \quad (2 \cdot 11 \cdot 2)$$

where $A = \partial F / \partial x$ and $B = \partial F / \partial y$. Clearly dF in (2.11.1) is simply the infinitesimal difference between two adjacent values of the function F . The infinitesimal quantity dF is here just an ordinary differential; it is also called an "exact differential" to distinguish it from other kinds of infinitesimal quantities to be discussed presently. Note that if one goes from an initial point i corresponding to (x_i, y_i) to a final point f corresponding to (x_f, y_f) , the corresponding change in F is simply given by

$$\Delta F = F_f - F_i = \int_i^f dF = \int_i^f (A dx + B dy) \quad (2 \cdot 11 \cdot 3)$$

Since the difference on the left side depends only on the initial and final points, the integral on the right can only depend on these end points; it thus *cannot* depend on the path along which it is evaluated in going from the initial point i to the final point f .

On the other hand, not every infinitesimal quantity is an exact differential. Consider, for example, the infinitesimal quantity

$$A'(x,y) dx + B'(x,y) dy \equiv dG \quad (2 \cdot 11 \cdot 4)$$

where A' and B' are some functions of x and y , and where dG has been introduced merely as an abbreviation for the expression on the left side. Although dG is certainly an infinitesimal quantity, it does *not* follow that it is necessarily an exact differential; i.e., it is in general *not* true that there exists some function $G = G(x,y)$ whose value is determined when x and y are given, and which is such that $dG = G(x+dx, y+dy) - G(x,y)$ is equal to the expression (2.11.4). Equivalently, it is in general *not* true that, if one sums (i.e., integrates) the infinitesimal quantities dG in going from the point i to the point f along a certain path, that the integral

$$\int_i^f dG = \int_i^f (A' dx + B' dy) \quad (2 \cdot 11 \cdot 5)$$

is independent of the particular path used. When an infinitesimal quantity is not an exact differential it is called an "inexact differential."

Example Consider the infinitesimal quantity

$$dG = \alpha dx + \beta \frac{x}{y} dy = \alpha dx + \beta x d(\ln y)$$

where α and β are constants. Let i denote the initial point $(1,1)$ and f the final point $(2,2)$. In Fig. 2.11.1 one can then, for example, calculate the integral of dG along the path $i \rightarrow a \rightarrow f$ passing through the point a with coordinates $(2,1)$; this gives

$$\int_{iaf} dG = \alpha + 2\beta \ln 2$$

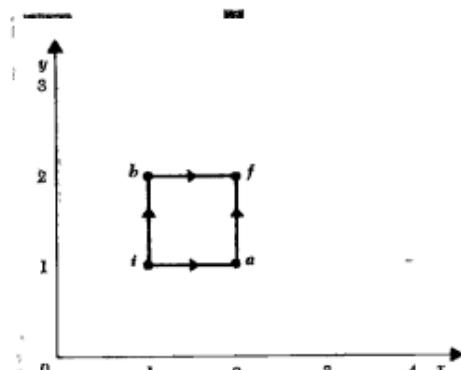


Fig. 2·11·1 Alternative paths connecting the points i and f in the xy plane.

Alternatively, one can calculate it along the path $i \rightarrow b \rightarrow f$ passing through the point b with coordinates $(1,2)$; this gives

$$\int_{ibf} dG = \beta \ln 2 + \alpha$$

Thus the integrals are different and the quantity dG is not an exact differential.

On the other hand, the infinitesimal quantity

$$dF = \frac{dG}{x} = \frac{\alpha}{x} dx + \frac{\beta}{y} dy$$

is an exact differential of the well-defined function $F = \alpha \ln x + \beta \ln y$. The integral of dF from i to f is thus always equal to

$$\int_i^f dF = \int_i^f \frac{dG}{x} = (\alpha + \beta) \ln 2$$

irrespective of the path chosen to go from i to f .

After this purely mathematical illustration, let us return to the physical situation of interest. The macrostate of a macroscopic system can be specified by the values of its external parameters (e.g., of its volume V) and of its mean energy \bar{E} ; other quantities, such as its mean pressure \bar{p} , are then determined. Alternatively, one can choose the external parameters and the pressure \bar{p} as the independent variables describing the macrostate; the mean energy \bar{E} is then determined. Quantities such as $d\bar{p}$ or $d\bar{E}$ are thus infinitesimal differences between well-defined quantities, i.e., they are just ordinary (i.e., exact) differentials. For example, $d\bar{E} = \bar{E}_f - \bar{E}_i$ is simply the difference between the well-defined mean energy \bar{E}_f of the system in a final macrostate f and its well-defined mean energy \bar{E}_i in an initial macrostate i when these two states are only infinitesimally different. It also follows that, if the system is taken from any initial macrostate i to any final macrostate f , its mean energy change is simply

given by

$$\Delta \bar{E} = \bar{E}_f - \bar{E}_i = \int_i' d\bar{E} \quad (2 \cdot 11 \cdot 6)$$

But since a quantity like \bar{E} is just a function of the macrostate under consideration, \bar{E}_i and \bar{E}_f depend only on the particular initial and final macrostates; thus the integral $\int_i' d\bar{E}$ over all the energy increments gained in the process depends *only* on the initial and final macrostates. In particular, therefore, the integral does *not* depend on what particular process is chosen to go from i to f in evaluating the integral.

On the other hand, consider the infinitesimal work dW done by the system in going from some initial macrostate i to some neighboring final macrostate f . In general $dW = \sum \bar{X}_a dx_a$ is *not* the difference between two numbers referring to the two neighboring macrostates, but is merely an infinitesimal quantity characteristic of the *process of going from state i to state f* . (It is meaningless to talk of the work *in* a given state; one can only talk of the work done in going *from* one state *to* another state.) The work dW is then in general an *inexact* differential. The total work done by the system in going from any macrostate i to some other macrostate f can be written as

$$W_{if} = \int_i' dW \quad (2 \cdot 11 \cdot 7)$$

where the integral represents simply the sum of the infinitesimal amounts of work dW performed at each stage of the process. But, in general, the value of the integral *does* depend on the particular process which is used in going from macrostate i to macrostate f .

Example Consider a system, e.g., a gas, whose volume V is the only relevant external parameter (see Fig. 2·7·4). Assume that the system is brought quasi-statically from its initial macrostate of volume V_i to its final macrostate of volume V_f . (During this process the system may be allowed to exchange heat with some other system.) We can describe the particular process used by specifying the mean pressure $\bar{p}(V)$ of the system for all values assumed by its volume in the course of the process. This functional relation can be repre-

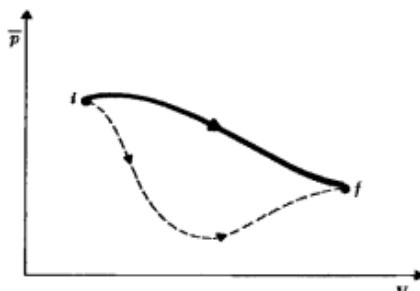


Fig. 2·11·2 Relation between mean pressure \bar{p} and volume V for two different quasi-static processes.

sented by the curve of Fig. 2·11·2 and the corresponding work is given by (2·10 4), i.e., by the area under the curve. If two different processes are used in going from i to f , described respectively by the solid and dotted p versus V curves in Fig. 2·11·2, then the areas under these two curves will be different. Thus the work W_{if} done by the system certainly depends on the particular process used in going from i to f .

In going from macrostate i to macrostate f the change $\Delta\bar{E}$ does *not* depend on the process, while the work W in general *does*. Hence it follows by (2·8·2) that the heat Q in general also *does* depend on the process used. Thus dQ denotes just an infinitesimal amount of heat absorbed during a process; like dW , it is, in general, *not* an exact differential.

Of course, if the system is thermally insulated so that $Q = 0$, Eq. (2·8·2) implies that

$$W_{if} = -\Delta\bar{E} \quad (2\cdot11\cdot8)$$

Then the work done depends *only* on the energy difference between initial and final macrostates and *is* independent of the process. Thus we have a result which is sometimes referred to as the "first law of thermodynamics":

If a *thermally isolated* system is brought from some initial to some final macrostate, the work done by the system is independent of the process used. } (2·11·9)

Remark This statement is an expression of conservation of energy and is subject to direct experimental verification. For example, one may conceive of the following type of experiment. A thermally insulated cylinder is closed by a piston. The cylinder contains a system which consists of a liquid in which there is immersed a small paddle wheel which can be rotated from outside by a falling weight. Work can be done on this system by either (a) moving the piston, or (b) rotating the paddle wheel. The respective amounts of work can be measured in terms of mechanical quantities by knowing (a) the mean pressure on, and displacement of, the piston, and (b) the distance

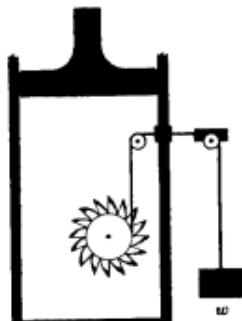


Fig. 2·11·3 A thermally insulated system on which work can be done in various ways.

by which the known weight descends. By doing such work the system can be brought from its initial macrostate of volume V_i and pressure \bar{p}_i to a final state of volume V_f and pressure \bar{p}_f . But this can be done in many ways: e.g., by rotating the paddle wheel first and then moving the piston the required amount; or by moving the piston first and then rotating the paddle wheel through the requisite number of revolutions; or by performing these two types of work alternately in smaller amounts. The statement (2.11.9) asserts that if the *total* work performed in each such procedure is measured, the result is always the same.*

Similarly, it follows that if the external parameters of a system are kept fixed so that it does no work, then $dW = 0$ and (2.8.3) reduces to

$$dQ = d\bar{E}$$

so that dQ becomes an exact differential. The amount of heat Q absorbed in going from one macrostate to another is then independent of the process used and depends only on the mean energy difference between them.

SUGGESTIONS FOR SUPPLEMENTARY READING

Statistical formulation

R. C. Tolman: "The Principles of Statistical Mechanics," chaps. 3 and 9, Oxford University Press, Oxford, 1938. (This book is a classic in the field of statistical mechanics and is entirely devoted to a careful exposition of fundamental ideas. The chapters cited discuss ensembles of systems and the fundamental statistical postulate in classical and quantum mechanics, respectively.)

Work and heat—macroscopic discussion

M. W. Zemansky: "Heat and Thermodynamics," 4th ed, chaps. 3 and 4, McGraw-Hill Book Company, New York, 1957.
 H. B. Callen: "Thermodynamics," secs. 1.1–1.7, John Wiley & Sons, Inc., New York, 1960. (The analogy mentioned on pp. 19 and 20 is particularly instructive.)

PROBLEMS

- 2.1** A particle of mass m is free to move in one dimension. Denote its position coordinate by x and its momentum by p . Suppose that this particle is confined within a box so as to be located between $x = 0$ and $x = L$, and suppose that its energy is known to lie between E and $E + \delta E$. Draw the classical phase space

* Paddle wheels such as this were historically used by Joule in the last century to establish the equivalence of heat and mechanical energy. In the experiment just mentioned we might equally well replace the paddle wheel by an electric resistor on which electrical work can be done by sending through it a known electric current.

of this particle, indicating the regions of this space which are accessible to the particle.

- 2.2** Consider a system consisting of two weakly interacting particles, each of mass m and free to move in one dimension. Denote the respective position coordinates of the two particles by x_1 and x_2 , their respective momenta by p_1 and p_2 . The particles are confined within a box with end walls located at $x = 0$ and $x = L$. The total energy of the system is known to lie between E and $E + \delta E$. Since it is difficult to draw a four-dimensional phase space, draw separately the part of the phase space involving x_1 and x_2 and that involving p_1 and p_2 . Indicate on these diagrams the regions of phase space accessible to the system.

- 2.3** Consider an ensemble of classical one-dimensional harmonic oscillators.

(a) Let the displacement x of an oscillator as a function of time t be given by $x = A \cos(\omega t + \varphi)$. Assume that the phase angle φ is equally likely to assume any value in its range $0 < \varphi < 2\pi$. The probability $w(\varphi) d\varphi$ that φ lies in the range between φ and $\varphi + d\varphi$ is then simply $w(\varphi) d\varphi = (2\pi)^{-1} d\varphi$. For any fixed time t , find the probability $P(x) dx$ that x lies between x and $x + dx$ by summing $w(\varphi) d\varphi$ over all angles φ for which x lies in this range. Express $P(x)$ in terms of A and x .

(b) Consider the classical phase space for such an ensemble of oscillators, their energy being known to lie in the small range between E and $E + \delta E$. Calculate $P(x) dx$ by taking the ratio of that volume of phase space lying in this energy range *and* in the range between x and $x + dx$ to the total volume of phase space lying in the energy range between E and $E + \delta E$ (see Fig. 2·3·1). Express $P(x)$ in terms of E and x . By relating E to the amplitude A , show that the result is the same as that obtained in part (a).

- 2.4** Consider an isolated system consisting of a large number N of very weakly interacting localized particles of spin $\frac{1}{2}$. Each particle has a magnetic moment μ which can point either parallel or antiparallel to an applied field H . The energy E of the system is then $E = -(n_1 - n_2)\mu H$, where n_1 is the number of spins aligned parallel to H and n_2 the number of spins aligned antiparallel to H .

(a) Consider the energy range between E and $E + \delta E$ where δE is very small compared to E but is microscopically large so that $\delta E \gg \mu H$. What is the total number of states $\Omega(E)$ lying in this energy range?

(b) Write down an expression for $\ln \Omega(E)$ as a function of E . Simplify this expression by applying Stirling's formula in its simplest form (A·6·2).

(c) Assume that the energy E is in a region where $\Omega(E)$ is appreciable, i.e., that it is not close to the extreme possible values $\pm N\mu H$ which it can assume. In this case apply a Gaussian approximation to part (a) to obtain a simple expression for $\Omega(E)$ as a function of E .

- 2.5** Consider the infinitesimal quantity

$$A dx + B dy \equiv dF$$

where A and B are both functions of x and y .

(a) Suppose that dF is an exact differential so that $F = F(x, y)$. Show that A and B must then satisfy the condition

$$\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x}$$

(b) If dF is an exact differential, show that the integral $\int dF$ evaluated along *any* closed path in the xy plane must vanish.

- 2.6 Consider the infinitesimal quantity

$$(x^2 - y) dx + x dy \equiv dF \quad (1)$$

(a) Is this an exact differential?

- (b) Evaluate the integral $\int dF$ between the points (1,1) and (2,2) of Fig. 2-11-1 along the straight-line paths connecting the following points:

$$(1,1) \rightarrow (1,2) \rightarrow (2,2)$$

$$(1,1) \rightarrow (2,1) \rightarrow (2,2)$$

$$(1,1) \rightarrow (2,2)$$

(c) Suppose that both sides of (1) are divided by x^2 . This yields the quantity $dG = dF/x^2$. Is dG an exact differential?

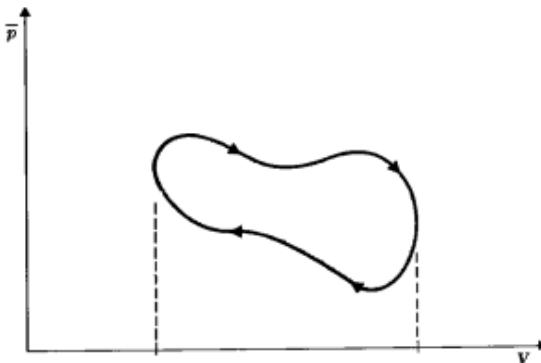
(d) Evaluate the integral $\int dG$ along the three paths of part (b).

- 2.7 Consider a particle confined within a box in the shape of a rectangular parallelepiped of edges L_x , L_y , and L_z . The possible energy levels of this particle are then given by (2-1-3).

(a) Suppose that the particle is in a given state specified by particular values of the three integers n_x , n_y , and n_z . By considering how the energy of this state must change when the length L_x of the box is changed quasistatically by a small amount dL_x , show that the force exerted by the particle in this state on a wall perpendicular to the x axis is given by $F_x = -\partial E / \partial L_x$.

(b) Calculate explicitly the force per unit area (or pressure) on this wall. By averaging over all possible states, find an expression for the mean pressure on this wall. (Exploit the property that the average values $\overline{n_x^2} = \overline{n_y^2} = \overline{n_z^2}$ must all be equal by symmetry.) Show that this mean pressure can be very simply expressed in terms of the mean energy \bar{E} of the particle and the volume $V = L_x L_y L_z$ of the box.

- 2.8 A system undergoes a quasi-static process which appears in a diagram of mean pressure \bar{p} versus volume V as a closed curve. (See diagram. Such a process is called "cyclic" since the system ends up in a final macrostate which is identical to its initial macrostate.) Show that the work done by the system is given by the area contained within the closed curve.



- 2.9 The tension in a wire is increased quasi-statically from F_1 to F_2 . If the wire has length L , cross-sectional area A , and Young's modulus Y , calculate the work done.

- 2.10** The mean pressure \bar{p} of a thermally insulated amount of gas varies with its volume V according to the relation

$$\bar{p}V^\gamma = K$$

where γ and K are constants. Find the work done by this gas in a quasistatic process from a macrostate with pressure \bar{p}_i and volume V_i , to one with pressure \bar{p}_f and volume V_f . Express your answer in terms of \bar{p}_i , V_i , \bar{p}_f , V_f , and γ .

- 2.11** In a quasi-static process $A \rightarrow B$ (see diagram) in which no heat is exchanged with the environment, the mean pressure \bar{p} of a certain amount of gas is found to change with its volume V according to the relation

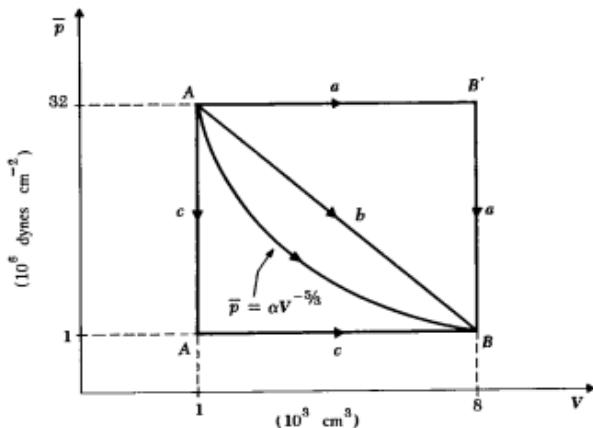
$$\bar{p} = \alpha V^{-\frac{5}{3}}$$

where α is a constant. Find the quasi-static work done and the net heat absorbed by this system in each of the following three processes, all of which take the system from macrostate A to macrostate B .

(a) The system is expanded from its original to its final volume, heat being added to maintain the pressure constant. The volume is then kept constant, and heat is extracted to reduce the pressure to 10^6 dynes cm^{-2} .

(b) The volume is increased and heat is supplied to cause the pressure to decrease linearly with the volume.

(c) The two steps of process (a) are performed in the opposite order.



Statistical thermodynamics

THE FUNDAMENTAL statistical postulate of equal a priori probabilities can be used as the basis of the entire theory of systems in equilibrium. In addition, the hypothesis mentioned at the end of Sec. 2·3 (and based on the assumed validity of the H theorem) also makes a statement about isolated systems not in equilibrium, asserting that these tend to approach ultimate equilibrium situations (characterized by the uniform statistical distribution over accessible states which is demanded by the fundamental postulate).

In this chapter we shall show how these basic statements lead to some very general conclusions concerning all macroscopic systems. The important results and relationships thus established constitute the basic framework of the discipline of "equilibrium statistical mechanics" or, as it is sometimes called, "statistical thermodynamics." Indeed, the major portion of this book will deal with systems in equilibrium and will therefore be an elaboration of the fundamental ideas developed in this chapter.

IRREVERSIBILITY AND THE ATTAINMENT OF EQUILIBRIUM

3 · 1 *Equilibrium conditions and constraints*

Consider an isolated system whose energy is specified to lie in a narrow range. As usual, we denote by Ω the number of states accessible to this system. By our fundamental postulate we know that, in equilibrium, such a system is equally likely to be found in any one of these states.

We recall briefly what we mean by "accessible states." There are in general some specified conditions which the system is known to satisfy. These act as constraints which limit the number of states in which the system can possibly be found without violating these conditions. The accessible states are then all the states consistent with these constraints.

The constraints can be described more quantitatively by specifying the

values of some parameters* y_1, y_2, \dots, y_n which characterize the system on a macroscopic scale. The number of states accessible to the system depends then on the values of these parameters; i.e., one can write the functional relation

$$\Omega = \Omega(y_1, \dots, y_n)$$

for the number of states accessible to the system when each parameter labeled by α lies in the range between y_α and $y_\alpha + \delta y_\alpha$. For example, a parameter y_α might denote the volume or the energy of some subsystem. We give some concrete illustrations.

Example 1 Consider the system shown in Fig. 2.3.2 where a box is divided by a partition into two equal parts, each of volume V_s . The left half of the box is filled with gas, while the right one is empty. Here the partition acts as a constraint which specifies that only those states of the system are accessible for which the coordinates of all the molecules lie in the left half of the box. In other words, the volume V accessible to the gas is a parameter which has the prescribed value $V = V_s$.

Example 2 Consider a system $A^{(0)}$ consisting of two subsystems A and A' separated by a fixed thermally insulating partition (see Fig. 2·7·1). This partition acts as a constraint which specifies that no energy can be exchanged between A and A' . Hence only those states of $A^{(0)}$ are accessible which have the property that the energy of A remains constant at some specified value $E = E_s$, while that of A' remains constant at some other specified value $E' = E'_s$.

Example 3 Consider the system of Fig. 2·8·1 where a thermally insulated piston separates two gases A and A' . If the piston is clamped in position, then this piston acts as a constraint which specifies that only those states of the total system are accessible which are such that the A molecules lie within a given fixed volume V_s , while the A' molecules lie within a given fixed volume V'_s .

Suppose that the initial situation with the given constraints is one of equilibrium where the isolated system is equally likely to be found in any of its Ω_s accessible states. Consider that some of the constraints are now removed. Then all the states formerly accessible to the system still remain accessible to it; but many more additional states will, in general, also become accessible. A removal of constraints can then only result in increasing, or possibly leaving unchanged, the number of states accessible to the system. Denoting the final number of accessible states by Ω_f , one can write

$$\Omega_f \geq \Omega_s \quad (3 \cdot 1 \cdot 1)$$

* These are not necessarily *external* parameters.

Focus attention on a representative ensemble of systems similar to the one under consideration and suppose that, when the constraints are removed, $\Omega_f > \Omega_i$. Immediately after the constraints are removed, the systems in the ensemble will not be in any of the states from which they were previously excluded. But the systems occupy then only a fraction

$$P_i = \frac{\Omega_i}{\Omega_f} \quad (3 \cdot 1 \cdot 2)$$

of the Ω_f states now accessible to them. This is *not* an equilibrium situation. Indeed, our fundamental postulate asserts that in the final equilibrium situation consistent with the absence of constraints, it is equally likely that each of the Ω_f states be occupied by the systems. If $\Omega_f \gg \Omega_i$, the particular situation where the systems are distributed only over the Ω_i original states becomes thus a very unlikely one; to be precise, its probability of occurrence is given by (3.1.2). In accordance with the hypothesis discussed at the end of Sec. 2.3, there is then a pronounced tendency for the situation to change in time until the much more probable final equilibrium situation is reached where the systems in the ensemble are distributed equally over all the possible Ω_f states.

Let us illustrate these statements with the examples mentioned previously.

Example 1 Suppose that the partition in Fig. 2.3.2 is removed. There is now no longer any constraint preventing a molecule from occupying the right half of the box. It is therefore exceedingly improbable that all the molecules will remain concentrated in the left half. Instead, they move about until they become randomly distributed throughout the entire box. In this final equilibrium situation each molecule is then equally likely to be found anywhere inside the box.

Suppose that this final equilibrium situation in the absence of the partition has been attained. What then is the probability P_i of encountering a situation where all the molecules are again concentrated in the left half of the box? The probability of finding one given molecule in the left half of the box is $\frac{1}{2}$. Hence the probability P_i of simultaneously finding all N molecules in the left half of the box is obtained by simply multiplying the respective probabilities of each molecule being in the left half, i.e.,

$$P_i = \left(\frac{1}{2}\right)^N$$

When N is of the order of Avogadro's number, so that $N \approx 6 \times 10^{23}$, this probability is *fantastically* small; i.e.,

$$P_i \approx 10^{-2 \times 10^{23}}$$

Example 2 Imagine that the partition in Fig. 2.7.1 is made thermally conducting. This removes the former constraint because the systems A and A' are now free to exchange energy with each other. The number of states accessible to the combined system $A^{(0)} = A + A'$ will, in general, be much greater if A adjusts its energy to some new value (and if A' correspondingly

adjusts its energy so as to keep the energy of the isolated total system $A + A'$ unchanged). Hence the most probable final equilibrium situation results when such an adjustment has taken place by virtue of heat transfer between the two systems.

Example 3 Imagine that the piston in Fig. 2·8·1 is unclamped so that it is free to move. Then the number of states accessible to the combined system $A + A'$ is, in general, much increased if the volumes of A and A' assume new values significantly different from their original ones. Thus a much more probable final equilibrium situation for the combined system $A + A'$ is attained if the piston moves so as to bring the volumes of A and A' to these new values.* As one would expect (and as we shall prove later), this final equilibrium situation corresponds to one where the mean pressures of both gases are equal so that the piston is in mechanical equilibrium.

This discussion can be phrased in terms of the relevant parameters y_1, \dots, y_n of the system. Suppose that a constraint is removed; for example, one of the parameters (call it simply y), which originally had the value $y = y_i$, is now allowed to vary. Since all the states accessible to the system are a priori equally likely, the equilibrium probability distribution $P(y)$ of finding the system in the range between y and $y + \delta y$ is proportional to the number of states accessible to the system when the parameter lies in this range; i.e.,

$$P(y) \propto \Omega(y) \quad (3 \cdot 1 \cdot 3)$$

This probability implies an occurrence of possible values of y , which is, in general, extremely different from the original situation where all systems in the ensemble were characterized by the value $y = y_i$ (see Fig. 3·1·1). In the absence of constraints the value $y = y_i$ of the parameter represents, therefore, a very improbable configuration. Hence the situation tends to change in time until the uniform equilibrium distribution of systems over

* The piston may oscillate back and forth several times before settling down in its final equilibrium position.

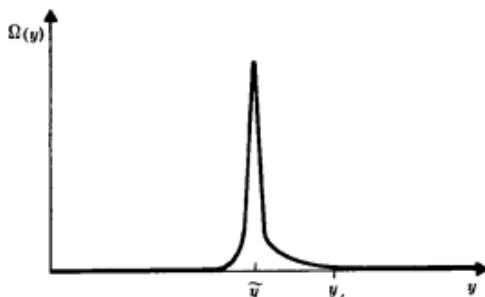


Fig. 3·1·1 Schematic diagram showing the number of states $\Omega(y)$ accessible to a system as a function of a parameter y . The initial value of this parameter is denoted by y_i .

accessible states is attained, i.e., until the various values of y occur with respective probabilities given by (3·1·3). Usually $\Omega(y)$ has a very pronounced maximum at some value \tilde{y} . In that case practically all systems in the final equilibrium situation will have values corresponding to the most probable situation where y is very close to \tilde{y} . Hence, if initially $y_i \neq \tilde{y}$, the parameter y will change after the constraint is removed until it attains values close to \tilde{y} where Ω is maximum. This discussion can be summarized by the following statement:

If some constraints of an isolated system are removed, the parameters of the system tend to readjust themselves in such a way that $\Omega(y_1, \dots, y_n)$ approaches a maximum. In symbols

$$\Omega(y_1, \dots, y_n) \rightarrow \text{maximum} \quad (3 \cdot 1 \cdot 4)$$

3 · 2 *Reversible and irreversible processes*

Suppose that the final equilibrium situation has been reached so that the systems in the ensemble are uniformly distributed over the Ω_f accessible final states. If the constraints are now simply restored, the systems in the ensemble will still occupy these Ω_f states with equal probability. Thus, if $\Omega_f > \Omega_i$, simply restoring the constraints does *not* restore the initial situation. Once the systems are randomly distributed over the Ω_f states, simply imposing or reimposing a constraint cannot cause the systems to move spontaneously out of some of their possible states so as to occupy a more restricted class of states. Nor could the removal of any other constraints make the situation any better; it could only lead to a situation where even more states become accessible to the system so that the system could be found in them as well.

Consider an isolated system (i.e., one which cannot exchange energy in the form of heat or work with any other system) and suppose that some process occurs in which the system goes from some initial situation to some final situation. If the final situation is such that the imposition or removal of constraints of this *isolated* system cannot restore the initial situation, then the process is said to be "irreversible." On the other hand, if it is such that the imposition or removal of constraints *can* restore the initial situation, then the process is said to be "reversible."

In terms of these definitions, the original removal of the constraints in the case where $\Omega_f > \Omega_i$ can be said to be an irreversible process. Of course, it is possible to encounter the special case where the original removal of the constraints does not change the number of accessible states so that $\Omega_f = \Omega_i$. Then the system, originally in equilibrium and equally likely to be in any of its Ω_i states, will simply remain distributed with equal probability over these states. The equilibrium of the system is then completely undisturbed so that this special process is reversible.

We again illustrate these comments with the previous examples:

Example 1 Once the molecules are in equilibrium and uniformly distributed throughout the box, the simple act of replacing the partition does not change the essential situation. The molecules still remain uniformly distributed throughout the box. The original removal of the partition thus constitutes an irreversible process.

This does *not* mean that the original situation of this system can *never* be restored. It can, provided that the system is *not* kept isolated but is allowed to interact with other systems. For example, in the present situation one can take a thin piston, which is originally coincident with the right wall of the box. One can now use some outside device A' (e.g., a falling weight) to move the piston to the center of the box, thus doing work on the gas in order to recompress it into the left half against the pressure exerted by the gas. The volume of the gas has now been restored to its original value V , and the right half of the box is empty as before. But the energy of the gas is greater than originally because of the work done on it during recompression. To restore the energy of the gas to its previous value, one can now let the gas give off just the right amount of heat by bringing it into thermal contact with some suitable system A'' . The gas has thus been restored to its original situation, its volume and energy being the same as initially.

Of course, the *isolated* system $A^{(0)}$ consisting of the gas *and* the systems A' and A'' has *not* been restored to its original situation, since the systems A' and A'' have been changed in the process. The process is still irreversible for the entire system $A^{(0)}$. Indeed, in releasing the weight to move the piston and in eliminating the thermal insulation to allow heat exchange with A'' , we have removed constraints of $A^{(0)}$ and increased the number of states accessible to this isolated system.

Example 2 Suppose that thermal interaction between A and A' has taken place and that the systems are in equilibrium. Simply making the partition again thermally insulating does not change the new energies of A and A' . One cannot restore the system $A + A'$ to its original situation by making heat flow in a direction opposite to the original direction of spontaneous heat transfer (unless one introduces interaction with suitable outside systems). The original heat transfer is thus an irreversible process.

Of course, a special case may arise where the initial energies of the systems A and A' are such that making the partition originally thermally conducting does *not* increase the number of states accessible to the combined system $A + A'$. Then no net energy exchange takes place between A and A' , and the process is reversible.

Example 3 This again is, in general, an irreversible process. Simply clamping the piston in its new position, so that it is again not free to move, does not restore the initial volumes of the gases.

The discussion of this section can be summarized by the statement that if some constraints of an isolated system in equilibrium are removed, the num-

ber of states accessible to the system can only increase or remain the same, i.e., $\Omega_f \geq \Omega_i$.

If $\Omega_f = \Omega_i$, then the systems in the representative ensemble are already distributed with equal probability over all their accessible states. The system remains, therefore, always in equilibrium and the process is *reversible*.

If $\Omega_f > \Omega_i$, then the distribution of systems over the possible states in the representative ensemble is a very improbable one. The system will therefore tend to change in time until the most probable final equilibrium situation of uniform distribution of systems over accessible states is reached. Equilibrium does not prevail at all stages of the process and the process is *irreversible*.

Remarks on significant time scales Note that we have nowhere made any statements about the *rate* of a process, i.e., about the relaxation time τ required by a system to reach the final equilibrium situation. An answer to this kind of question could only be obtained by a *detailed* analysis of the interactions between particles, since these interactions are responsible for bringing about the change of the system in time which results in the attainment of the final equilibrium state. The beauty of our general probability arguments is precisely the fact that they yield information about equilibrium situations, *without* the necessity of getting involved in the difficult detailed analysis of interactions between the very many particles of a system.

The general probability arguments of statistical mechanics are thus basically restricted to a consideration of equilibrium situations which do not change in time. But this limitation is not quite as severe as it might appear at first sight. The important parameter is really the time t_{exp} of experimental interest compared to the significant relaxation times τ of the system under consideration. There are really three cases which may arise.

1. $\tau \ll t_{\text{exp}}$: In this case the system comes to equilibrium very quickly compared to times of experimental interest. Hence probability arguments concerning the resulting equilibrium situation are certainly applicable.

2. $\tau \gg t_{\text{exp}}$: This is the opposite limit where equilibrium is achieved very slowly compared to experimental times. Here the situation would not be changed significantly if one imagined constraints to be introduced which would prevent the system from ever reaching equilibrium at all. But in the presence of these constraints the system would be in equilibrium; hence it can again be treated by general probability arguments.

Example 1 Imagine that in Fig. 2·7·1 the partition has a very small thermal conductivity so that the amount of energy transferred between A and A' during a time t_{exp} of experimental interest is very small. Then the situation would be substantially the same if the partition were made thermally insulating; in that case both A and A' can separately be considered to be in thermal equilibrium and can be discussed accordingly.

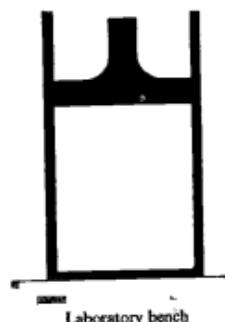


Fig. 3·2·1 Experimental setup in which a gas is contained in a cylinder closed by a piston which is free to oscillate. The gas is in thermal contact with the laboratory bench.

Example 2 As a second example, consider the situation of Fig. 3·2·1. Here a gas is contained in a cylinder closed by a movable piston, and the whole apparatus is sitting on a bench in the laboratory. When the piston is pushed down and then released, it will oscillate with a period t_{osc} about its equilibrium position. There are two significant relaxation times in the problem. If the piston is suddenly displaced, it takes a time τ_{th} before the gas will again come to thermal equilibrium with the laboratory bench by exchanging heat with it; it also takes a time τ_{int} before the gas of molecules will regain internal thermal equilibrium so that it is again uniformly distributed over all its accessible states. Ordinarily $\tau_{\text{int}} \ll \tau_{\text{th}}$. If the time of experimental interest (which is here the period of oscillation t_{osc}) is such that

$$\tau_{\text{int}} \ll t_{\text{osc}} \ll \tau_{\text{th}}$$

one can treat the problem to good approximation by considering the gas to be always in internal equilibrium in a macrostate corresponding to the instantaneous position of the piston, and by considering the walls of the cylinder to be thermally insulating.

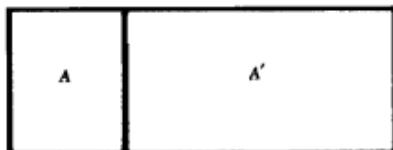
3. $\tau \approx t_{\text{exp}}$: In this case the time required to reach equilibrium is comparable to times of experimental significance. The statistical distribution of the system over its accessible states is then not uniform and keeps on changing during the time under consideration. One is then faced with a difficult problem which cannot be reduced to a discussion of equilibrium situations.

THERMAL INTERACTION BETWEEN MACROSCOPIC SYSTEMS

3 · 3 Distribution of energy between systems in equilibrium

Let us now discuss in greater detail the thermal interaction between two macroscopic systems A and A' . We shall denote the respective energies of these systems by E and E' . For convenience, we imagine these energy scales to be subdivided into equal small intervals of respective magnitudes δE and $\delta E'$; then we shall denote by $\Omega(E)$ the number of states of A in the range

Fig. 3·3·1 Two macroscopic systems A and A' in thermal interaction with each other.



between E and $E + \delta E$, and by $\Omega'(E')$ the number of states of A' in the range between E' and $E' + \delta E'$.

We assume that the systems are not thermally insulated from each other so that they are free to exchange energy. (The external parameters of the systems are supposed to remain fixed; thus the energy transfer is in the form of heat.) The combined system $A^{(0)} = A + A'$ is isolated and its total energy $E^{(0)}$ is therefore constant. The energy of each system separately is, however, not fixed, since it can exchange energy with the other system. We assume always, when speaking of thermal contact between two systems, that the interaction between the systems is weak so that their energies are simply additive. Thus we can write

$$E + E' = E^{(0)} = \text{constant} \quad (3\cdot3\cdot1)$$

Remark The Hamiltonian (or energy) \mathfrak{H} of the combined system can always be written in the form

$$\mathfrak{H} = \mathfrak{H}_A + \mathfrak{H}_{A'} + \mathfrak{H}^{\text{(int)}}$$

where \mathfrak{H} depends only on the variables describing A , \mathfrak{H}' only on the variables describing A' , and the interaction term $\mathfrak{H}^{\text{(int)}}$ on the variables of both systems.* This last term $\mathfrak{H}^{\text{(int)}}$ cannot be zero, because then the two systems would not interact at all and would have no way of exchanging energy and thus coming to equilibrium with each other. But the assumption of *weak* interaction is that $\mathfrak{H}^{\text{(int)}}$, although finite, is negligibly small compared to \mathfrak{H}_A and $\mathfrak{H}_{A'}$.

Suppose that the systems A and A' are in equilibrium with each other, and focus attention on a representative ensemble such as that shown in Fig. 2·6·1. Then the energy of A can assume a large range of possible values, but these values occur by no means with equal probability. Indeed, suppose that A has an energy E (i.e., more precisely an energy between E and $E + \delta E$). Then the corresponding energy of A' is by (3·3·1) known to be

$$E' = E^{(0)} - E \quad (3\cdot3\cdot2)$$

The number of states accessible to the entire system $A^{(0)}$ can thus be regarded as a function of a single parameter, the energy E of system A . Let us denote

* For example, for two particles moving in one dimension

$$\mathfrak{H} = \frac{p^2}{2m} + \frac{p'^2}{2m'} + U(x, x')$$

where the first terms describe their kinetic energies and the last one describes their potential energy of mutual interaction which depends on their positions x and x' .

by $\Omega^{(0)}(E)$ the number of states accessible to $A^{(0)}$ when A has an energy between E and $E + \delta E$. Our fundamental postulate asserts that in equilibrium $A^{(0)}$ must be equally likely to be found in any one of its states. Hence it follows that the probability $P(E)$ of finding this combined system in a configuration where A has an energy between E and $E + \delta E$ is simply proportional to the number of states $\Omega^{(0)}(E)$ accessible to the total system $A^{(0)}$ under these circumstances. In symbols this can be written

$$P(E) = C\Omega^{(0)}(E) \quad (3 \cdot 3 \cdot 3)$$

where C is a constant of proportionality independent of E .

More explicitly, this probability could also be written as

$$P(E) = \frac{\Omega^{(0)}(E)}{\Omega^{(0)}_{\text{tot}}}$$

where $\Omega^{(0)}_{\text{tot}}$ denotes the *total* number of states accessible to $A^{(0)}$. Of course, $\Omega^{(0)}_{\text{tot}}$ can be obtained by summing $\Omega^{(0)}(E)$ over all possible energies E of the system A . Similarly, the constant C in (3·3·3) can be determined by the normalization requirement that the probability $P(E)$ summed over all possible energies of A must yield unity. Thus

$$C^{-1} = \Omega^{(0)}_{\text{tot}} = \sum_E \Omega^{(0)}(E)$$

But when A has an energy E it can be in any one of its $\Omega(E)$ possible states. At the same time A' must then have an energy $E' = E^{(0)} - E$ so that it can be in any one of its $\Omega'(E') = \Omega'(E^{(0)} - E)$ possible states. Since every possible state of A can be combined with every possible state of A' to give a different state of the total system $A^{(0)}$, it follows that the number of distinct states accessible to $A^{(0)}$ when A has energy E is simply given by the product

$$\Omega^{(0)}(E) = \Omega(E)\Omega'(E^{(0)} - E) \quad (3 \cdot 3 \cdot 4)$$

Correspondingly, the probability (3·3·3) of system A having an energy near E is simply given by

$$P(E) = C\Omega(E)\Omega'(E^{(0)} - E) \quad (3 \cdot 3 \cdot 5)$$

Illustrative example with very small numbers Consider the two systems A and A' having the characteristics illustrated in Fig. 3·3·2. Suppose that the total energy $E^{(0)}$ of both systems is, in the arbitrary units used, equal to 15. One possible situation, for example, would be that $E = 4$ and $E' = 11$. In this case A could be any one of its two possible states and A' in any one of its 40 states. There are then a total of $\Omega^{(0)} = 2 \times 40 = 80$ different possible states for the combined system $A + A'$. Let us enumerate

systematically some of the conceivable situations in a table when the total energy of the system is $E^{(0)} = 15$.

Suppose $E = \dots$; then $E' = \dots$ Here $\Omega(E) = \dots$, and $\Omega'(E') = \dots$. Hence $\Omega^{(0)}(E) = \dots$

4	11	2	40	80
5	10	5	26	130
6	9	10	16	160
7	8	17	8	136
8	7	25	3	75

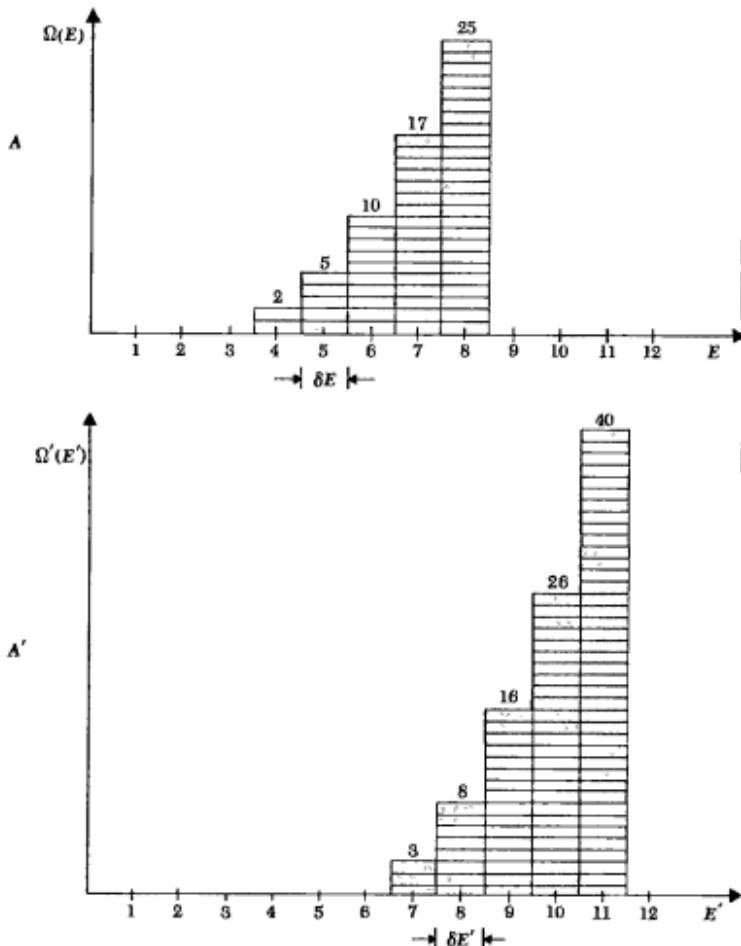


Fig. 3.3.2 Graph showing, in the case of two special very small systems A and A' , the number of states $\Omega(E)$ accessible to A when its energy is E and the number of states $\Omega'(E')$ accessible to A' when its energy is E' . (The energies are measured in terms of an arbitrary unit.)

Note that it would be most probable in the ensemble to find the combined system in a state where A has energy $E = 6$ and A' has energy $E' = 9$. This situation would be likely to occur twice as frequently as the situation where $E = 4$ and $E' = 11$.

Let us now investigate the dependence of $P(E)$ on the energy E . Since A and A' are both systems of very many degrees of freedom, we know by (2 · 5 · 10) that both $\Omega(E)$ and $\Omega'(E')$ are extremely rapidly increasing functions of their respective arguments. Hence it follows that if one considers the expression (3 · 3 · 5) as a function of increasing energy E , the factor $\Omega(E)$ increases extremely rapidly while the factor $\Omega'(E^{(0)} - E)$ decreases extremely rapidly. The result is that the product of these two factors, i.e., the probability $P(E)$, exhibits an extremely sharp maximum for some particular value \bar{E} of the energy E . Thus the dependence of $P(E)$ on E must show the general behavior illustrated in Fig. 3 · 3 · 3 where the width Δ^*E of the region where $P(E)$ has appreciable magnitude is such that $\Delta^*E \ll \bar{E}$.

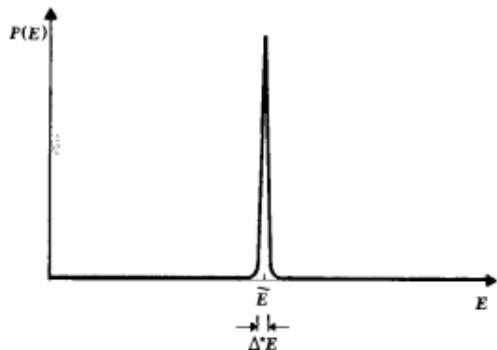


Fig. 3 · 3 · 3 Schematic illustration of the functional dependence of the probability $P(E)$ on the energy E .

Remark More explicitly, if the number of states exhibits the behavior discussed in (2 · 5 · 10) so that $\Omega \propto E^f$ and $\Omega' \propto E'^g$, then (3 · 3 · 5) gives

$$\ln P \approx f \ln E + g \ln (E^{(0)} - E) + \text{constant}$$

Thus $\ln P$ exhibits a unique maximum as a function of E , and this maximum of the logarithm corresponds to a very pronounced maximum of P itself. Except for the fact that this maximum is enormously sharper for these macroscopic systems where Ω and Ω' are such rapidly varying functions of energy, the situation is analogous to the simple example discussed above. We shall postpone until Sec. 3 · 6 a more quantitative estimate of the width Δ^*E of the maximum in the case of macroscopic systems.

To locate the position of the maximum of $P(E)$, or equivalently, of the

maximum of its logarithm, we need to find the value $E = \tilde{E}$, where*

$$\frac{\partial \ln P}{\partial E} = \frac{1}{P} \frac{\partial P}{\partial E} = 0 \quad (3 \cdot 3 \cdot 6)$$

But by (3 · 3 · 5)

$$\ln P(E) = \ln C + \ln \Omega(E) + \ln \Omega'(E') \quad (3 \cdot 3 \cdot 7)$$

where $E' = E^{(0)} - E$. Hence (3 · 3 · 6) becomes

$$\frac{\partial \ln \Omega(E)}{\partial E} + \frac{\partial \ln \Omega'(E')}{\partial E'} (-1) = 0$$

or

► $\beta(\tilde{E}) = \beta'(\tilde{E}')$ (3 · 3 · 8)

where \tilde{E} and \tilde{E}' denote the corresponding energies of A and A' at the maximum, and where we have introduced the definition

► $\beta(E) \equiv \frac{\partial \ln \Omega}{\partial E}$ (3 · 3 · 9)

with a corresponding definition for β' . The relation (3 · 3 · 8) is the equation which determines the value \tilde{E} where $P(E)$ is maximum.

By its definition, the parameter β has the dimensions of a reciprocal energy. It is convenient to introduce a dimensionless parameter T defined by writing

$$kT \equiv \frac{1}{\beta} \quad (3 \cdot 3 \cdot 10)$$

where k is some positive constant having the dimensions of energy and whose magnitude can be chosen in some convenient arbitrary way. The parameter T is then by (3 · 3 · 9) defined as

$$\frac{1}{T} = \frac{\partial S}{\partial E} \quad (3 \cdot 3 \cdot 11)$$

where we have introduced the definition

► $S \equiv k \ln \Omega$ (3 · 3 · 12)

This quantity S is given the name of "entropy." The condition of maximum probability $P(E)$ is then, by (3 · 3 · 7), expressible as the condition that the total entropy

$$S + S' = \text{maximum} \quad (3 \cdot 3 \cdot 13)$$

The condition that this occurs can, by (3 · 3 · 8), be written as

$$T = T' \quad (3 \cdot 3 \cdot 14)$$

* We write this as a *partial* derivative to emphasize that all external parameters of the system are considered to remain unchanged in this discussion. The reason that it is somewhat more convenient to work with $\ln P$ instead of P itself is that the logarithm is a much more slowly varying function of the energy E , and that it involves the numbers Ω and Ω' as a simple sum rather than as a product.

Remark Note that the number Ω of accessible states in the energy range δE (and hence also the entropy $S = k \ln \Omega$ of (3·3·12)) depends on the size δE chosen as the fixed small energy-subdivision interval in a given discussion. This dependence of S is, however, utterly negligible for a macroscopic system and does not at all affect the parameter β .

Indeed, by (2·5·1), $\Omega(E)$ is simply proportional to δE ; i.e., $\Omega(E) = \omega(E) \delta E$, where ω is the density of states which is independent of δE . Since δE is a fixed interval independent of E , it follows by (3·3·10) that

$$\beta = \frac{\partial}{\partial E} (\ln \omega + \ln \delta E) = \frac{\partial \ln \omega}{\partial E} \quad (3 \cdot 3 \cdot 15)$$

which is independent of δE . Furthermore, suppose that one had chosen instead of δE a different energy subdivision interval δ^*E . The corresponding number of states $\Omega^*(E)$ in the range between E and $E + \delta^*E$ would then be given by

$$\Omega^*(E) = \frac{\Omega(E)}{\delta E} \delta^*E$$

The corresponding entropy defined by (3·3·12) would then be

$$S^* = k \ln \Omega^* = S + k \ln \frac{\delta^*E}{\delta E} \quad (3 \cdot 3 \cdot 16)$$

Now by (2·5·9), $S = k \ln \Omega$ is of the order of kf , where f is the number of degrees of freedom of the system. Imagine then an extreme situation where the interval δE^* would be chosen to be so fantastically different from δE as to differ from it by a factor as large as f (e.g., by as much as 10^{24}). Then the second term on the right side of (3·3·16) would at most be of the order of $k \ln f$. But when f is a large number, $\ln f \ll f$. (For example, if $f = 10^{24}$, $\ln f = 55$, which is certainly utterly negligible compared to f itself.) The last term in (3·3·16) is therefore completely negligible compared to S , so that one has to excellent approximation

$$S^* = S$$

The value of the entropy $S = k \ln \Omega$ calculated by (3·3·12) is thus essentially independent of the interval δE chosen for the subdivision of the energy scale.

3 · 4 The approach to thermal equilibrium

We pointed out already that the maximum exhibited by $P(E)$ at the energy $E = \bar{E}$ is extremely sharp. There is, therefore, an overwhelmingly large probability that, in an equilibrium situation where A and A' are in thermal contact, the system A has an energy E very close to \bar{E} , while the system A' has correspondingly an energy very close to $E^{(0)} - \bar{E} \equiv \bar{E}'$. The respective mean energies of the systems in thermal contact must therefore also be equal to these energies; i.e., when the systems are in thermal contact,

$$\bar{E} = \bar{E} \quad \text{and} \quad \bar{E}' = \bar{E}' \quad (3 \cdot 4 \cdot 1)$$

Consider then the situation where A and A' are initially separately in equilibrium and isolated from each other, their respective energies being very close to E_i and $E_{i'}$. (Their respective mean energies are accordingly $\bar{E}_i = E_i$ and $\bar{E}_{i'} = E_{i'}$.) The systems A and A' are now placed in thermal contact so that they are free to exchange energy with each other. The resulting situation is then an extremely improbable one, unless it happens that the systems initially have energies very close to \bar{E} and \bar{E}' , respectively. The situation will therefore tend to change in time until the systems attain final mean energies \bar{E}_f and $\bar{E}_{f'}$ which are such that

$$\bar{E}_f = \bar{E} \quad \text{and} \quad \bar{E}_{f'} = \bar{E}' \quad (3 \cdot 4 \cdot 2)$$

so that the probability $P(E)$ becomes maximum. By (3.3.8) the β parameters of the systems are then equal, i.e.,

$$\beta_f = \beta_{f'} \quad (3 \cdot 4 \cdot 3)$$

where $\beta_f \equiv \beta(\bar{E}_f)$ and $\beta_{f'} \equiv \beta(\bar{E}_{f'})$

The final probability is maximum and thus never less than the original one. By virtue of (3.3.7) this statement can be expressed in terms of the definition (3.3.12) of the entropy as

$$S(\bar{E}_f) + S'(\bar{E}_{f'}) \geq S(\bar{E}_i) + S(\bar{E}_{i'}) \quad (3 \cdot 4 \cdot 4)$$

When A and A' exchange energy in attaining the final equilibrium, their total energy is, of course, always conserved. Thus

$$\bar{E}_f + \bar{E}_{f'} = \bar{E}_i + \bar{E}_{i'} \quad (3 \cdot 4 \cdot 5)$$

Let us denote the entropy changes of the systems by

$$\begin{aligned} \Delta S &\equiv S_f - S_i \equiv S(\bar{E}_f) - S(\bar{E}_i) \\ \Delta S' &\equiv S_{f'} - S_{i'} \equiv S(\bar{E}_{f'}) - S(\bar{E}_{i'}) \end{aligned} \quad (3 \cdot 4 \cdot 6)$$

Then the condition (3.4.4) can be written more compactly as

$$\blacktriangleright \quad \Delta S + \Delta S' \geq 0 \quad (3 \cdot 4 \cdot 7)$$

Similarly, the mean energy changes are, by definition, simply the respective heats absorbed by the two systems. Thus

$$\begin{aligned} Q &\equiv \bar{E}_f - \bar{E}_i \\ Q' &\equiv \bar{E}_{f'} - \bar{E}_{i'} \end{aligned} \quad (3 \cdot 4 \cdot 8)$$

The conservation of energy (3.4.5) can then be written more compactly as

$$\blacktriangleright \quad Q + Q' = 0 \quad (3 \cdot 4 \cdot 9)$$

Hence $Q' = -Q$, so that if Q is positive then Q' is negative, and vice versa. A negative heat absorbed is simply a heat given off, and (3.4.9) expresses the obvious fact that the heat absorbed by one system must be equal to the heat given off by the other system.

By *definition*, we shall call the system which absorbs heat the “colder” system and the system which gives off heat the “warmer,” or “hotter,” system.

There are thus basically two cases which can arise:

1. The initial energies of the system may be such that $\beta_i = \beta'_i$, where $\beta_i = \beta(\bar{E}_i)$ and $\beta'_i = \beta(\bar{E}'_i)$. Then $\bar{E}_i = \bar{E}'_i$, and the condition of maximum probability (or entropy) is already fulfilled. (Equation (3·4·7) becomes an equality.) The systems remain, therefore, in equilibrium. There is then also no net exchange of energy (i.e., of heat) between the systems.

2. More generally, the initial energies of the systems are such that $\beta_i \neq \beta'_i$. Then $\bar{E}_i \neq \bar{E}'_i$, and the systems are in a very improbable nonequilibrium situation. (Equation (3·4·7) is an inequality.) This situation will therefore change in time. Transfer of heat between the systems takes place until the condition of maximum probability (or entropy) is achieved where $\bar{E}_f = \bar{E}$ and where $\beta_f' = \beta_f$.

3 · 5 Temperature

In the preceding section we saw that the parameter β (or equivalently, $T = (k\beta)^{-1}$) has the following two properties:

1. If two systems separately in equilibrium are characterized by the *same* value of the parameter, then the systems will remain in equilibrium when brought into thermal contact with each other.

2. If the systems are characterized by *different* values of the parameter, then they will *not* remain in equilibrium when brought into thermal contact with each other.

In addition, suppose we have three systems *A*, *B*, and *C*. We know that if *A* and *C* remain in equilibrium when brought into thermal contact, then $\beta_A = \beta_C$. Similarly, we know that if *B* and *C* also remain in equilibrium when brought into thermal contact, then $\beta_B = \beta_C$. But then we can conclude that $\beta_A = \beta_B$, so that systems *A* and *B* will also remain in equilibrium when brought into thermal contact. We thus arrive at the following statement, sometimes known as the “zeroth law of thermodynamics”:

If two systems are in thermal equilibrium with a third system, then they must be in thermal equilibrium with each other. (3·5·1)

This property makes possible the use of test systems, called “thermometers,” which allow measurements to decide whether any two systems will or will not remain in equilibrium when brought into thermal contact with each other. Such a thermometer is any macroscopic system *M* chosen in accordance with the following two specifications:

1. Among the many macroscopic parameters characterizing the system *M*, select one (call it ϑ) which varies by appreciable amounts when *M* is brought

into thermal contact with the various systems to be tested. All the other macroscopic parameters of M are held fixed. The parameter ϑ , which is allowed to vary, is called the "thermometric parameter" of M .

2. The system M is chosen to be much smaller (i.e., to have many fewer degrees of freedom) than the systems which it is designed to test. This is desirable in order to minimize the possible energy transfer to the systems under test so as to reduce the disturbance of the systems under test to a minimum.

Examples of thermometers

a. Mercury in a glass tube. The height of the mercury in the tube is taken as the thermometric parameter ϑ . This is the familiar "mercury-in-glass thermometer."

b. Gas in a bulb, its volume being maintained constant. The mean pressure of the gas is taken as the thermometric parameter ϑ . This is called a "constant-volume gas thermometer."

c. Gas in a bulb, its pressure being maintained constant. The volume of the gas is taken as the thermometric parameter ϑ . This is called a "constant-pressure gas thermometer."

d. An electrical conductor maintained at constant pressure and carrying a current. The electrical resistance of the conductor is the thermometric parameter ϑ . This is called a "resistance thermometer."

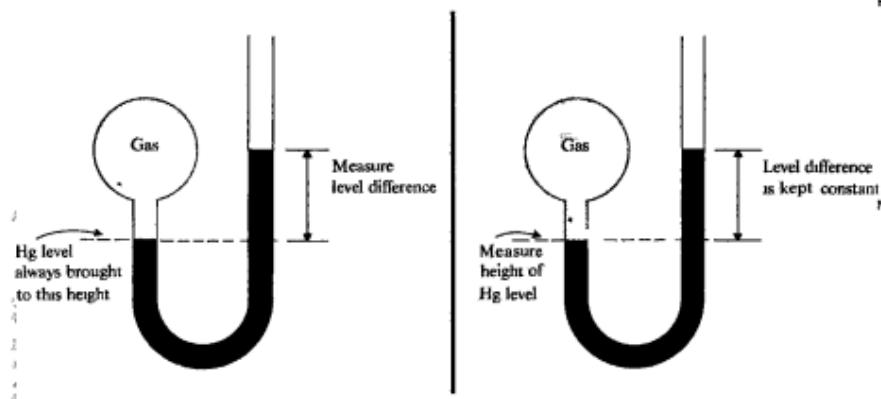


Fig. 3·5·1 Constant-volume and constant-pressure gas thermometers.

A thermometer M is used in the following way. It is successively placed in thermal contact with the systems under test, call them A and B , and is allowed to come to equilibrium with each.

1. If the thermometric parameter ϑ of the thermometer M (e.g., the height of the Hg column of the mercury-in-glass thermometer) has the same value in both cases, one knows that after M has come to equilibrium with A , it remains in equilibrium after being placed in thermal contact with B . Hence the zeroth law allows one to conclude that A and B will remain in equilibrium if brought into contact with each other.

2. If the thermometric parameter of M does *not* have the same value in both cases, then one knows that A and B will *not* remain in equilibrium if brought into thermal contact with each other. For suppose they did remain in equilibrium; then, after M attains thermal equilibrium with A , it would by the zeroth law have to remain in equilibrium when brought into thermal contact with B . Thus the parameter ϑ could not change when M is brought into thermal contact with B .

Consider *any* thermometer M with *any one* parameter ϑ chosen as its thermometric parameter. The value assumed by ϑ when the thermometer M has come to thermal equilibrium with some system A will, by definition, be called the "temperature" of the system A with respect to the particular thermometric parameter ϑ of the particular thermometer M .

According to this definition the temperature can be a length, a pressure, or any other quantity. Note that, even if two different thermometers have parameters of the same dimensions, it is in general not true that they will yield the same value of temperature for the same body. Furthermore, if a body C has a temperature halfway between the temperatures of bodies A and B when measured by one thermometer, this statement is not necessarily true with respect to the temperatures measured by some other thermometer.

Nevertheless, the somewhat arbitrary temperature concept which we have defined has, according to our discussion, the following fundamental and useful property:

Two systems will remain in equilibrium when placed in thermal contact with each other if and only if they have the same temperature (referred to the same thermometer). } (3 · 5 · 2)

Note that if $\psi(\vartheta)$ is *any* single-valued function of ϑ , then it can be used as a thermometric parameter just as well as ϑ itself. This function $\psi(\vartheta)$ also satisfies the property (3 · 5 · 2) and can be equally well designated as the temperature of the system with respect to the particular thermometer chosen for the measurement.

The temperature concept which we have defined is important and useful, but is rather arbitrary in the sense that the temperature assigned to a system depends in an essential way on the peculiar properties of the particular system M used as the thermometer.

On the other hand, we can exploit the properties of the parameter β and use the particular parameter β_M of the thermometer M as its thermometric parameter. Then we know that when the thermometer is in thermal equilibrium with a system A , $\beta_M = \beta_A$. The thermometer measures then, by virtue of (3 · 3 · 9), a fundamental property of the system A , namely, the variation of its density of states with energy. Furthermore, if one uses *any other* thermometer M' , it too will read a value $\beta_{M'} = \beta_A$ when brought into thermal contact with system A . Thus we see that

If the parameter β is used as a thermometric parameter, then *any* thermometer yields the *same* temperature reading when used to measure the temperature of a particular system. Furthermore, this temperature measures a fundamental property of the density of states of the system under test.

The parameter β is, therefore, a particularly useful and fundamental temperature parameter. The corresponding dimensionless quantity $T = (k\beta)^{-1}$ is accordingly called the "absolute temperature." We shall postpone until later a discussion of practical procedures for finding numerical values of β or T by appropriate measurements.

Some properties of the absolute temperature By (3·3·9) the absolute temperature is given by

$$\frac{1}{kT} = \beta = \frac{\partial \ln \Omega}{\partial E} \quad (3\cdot5\cdot3)$$

We saw in Sec. 2·5 that $\Omega(E)$ is ordinarily a very rapidly increasing function of the energy E . Hence (3·5·3) shows that ordinarily

$$\beta > 0 \quad \text{or} \quad T > 0 \quad (3\cdot5\cdot4)$$

***Remark** This is true for all ordinary systems where one takes into account the kinetic energy of the particles. Such systems have no upper bound on their possible energy (a lower bound, of course, always exists—namely, the quantum mechanical ground-state energy of the system); and as we have seen in Sec. 2·5, $\Omega(E)$ increases then roughly like E^f , where f is the number of degrees of freedom of the system. Exceptional situations may arise, however, where one does not want to take into account the translational degrees of freedom of a system (i.e., coordinates and momenta), but focuses attention *only* on its spin degrees of freedom. In that case the system has an upper bound to its possible energy (e.g., all spins lined up antiparallel to the field) as well as a lower bound (e.g., all spins lined up parallel to the field). Correspondingly, the *total* number of states (irrespective of energy) available to the system is finite. In this case the number of possible spin states $\Omega_{\text{spin}}(E)$ at first increases, as usual, with increasing energy; but then it reaches a maximum and decreases again. Thus it is possible to get absolute *spin* temperatures which are negative as well as positive.

If one disregards such exceptional cases where systems have an upper bound to their possible total energy, T is always positive and some further general statements can readily be made. In (2·5·10) it was shown that the functional dependence of $\Omega(E)$ is roughly given by

$$\Omega(E) \propto E^f$$

where f is the number of degrees of freedom of the system and the energy E is measured with respect to its ground state. Thus

$$\ln \Omega \approx f \ln E + \text{constant}$$

Thus, when $E = \bar{E} \approx \bar{E}$, one gets

$$\beta = \frac{\partial \ln \Omega(E)}{\partial E} \approx \frac{f}{\bar{E}} \quad (3 \cdot 5 \cdot 5)$$

and

$$kT \approx \frac{\bar{E}}{f} \quad (3 \cdot 5 \cdot 6)$$

Thus the quantity kT is a rough measure of the mean energy, above the ground state, per degree of freedom of a system.

The condition of equilibrium (3 · 3 · 8) between two systems in thermal contact asserts that their respective absolute temperatures must be equal. By virtue of (3 · 5 · 6) we see that this condition is roughly equivalent to the fairly obvious statement that the total energy of the interacting systems is shared between them in such a way that the mean energy per degree of freedom is the same for both systems.

In addition, the absolute temperature T has the general property of indicating the direction of heat flow between two systems in thermal contact. This is most readily seen in the case where an infinitesimal amount of heat Q is transferred between two systems A and A' , originally at slightly different initial temperatures β_i and $\beta_{i'}$. Using the notation of Sec. 3 · 3, the condition that the probability (3 · 3 · 7) must increase in the process (or equivalently, the condition (3 · 4 · 7) for the entropy) can be written as

$$\frac{\partial \ln \Omega(\bar{E}_i)}{\partial E} (\bar{E}_f - \bar{E}_i) + \frac{\partial \ln \Omega'(\bar{E}_{i'})}{\partial E'} (\bar{E}'_f - \bar{E}'_i) \geq 0$$

Using the definition of β as well as the relations (3 · 4 · 8), this becomes

$$(\beta_i - \beta_{i'})Q \geq 0$$

Thus, if $Q > 0$, then

$$\beta_i \geq \beta_{i'}$$

$$T_i \leq T_{i'}$$

if T_i and $T_{i'}$ are positive. Hence positive heat is always absorbed by the system with higher β and given off by the system with lower β . Or, in the ordinary case where the absolute temperatures are positive, heat is absorbed by the system at the lower absolute temperature T and given off by the system at the higher absolute temperature T . Since the words "colder" and "warmer" were defined in Sec. 3 · 4 in terms of the direction of heat flow, one can say (in the case of ordinary positive absolute temperatures) that the warmer system has a higher absolute temperature than the colder one. (We shall see later that the same conclusions are true in situations where a *finite* amount of heat is transferred.)

3 · 6 Heat reservoirs

The thermal interaction between two systems is particularly simple if one of them is very much larger than the other one (i.e., if it has many more degrees

of freedom). To be precise, suppose that A' denotes the large system and A any relatively small system with which it may interact. The system A' is then said to act as a "heat reservoir," or "heat bath," with respect to the smaller system if it is so large that its temperature parameter remains essentially unchanged irrespective of any amount of heat Q' which it may absorb from the smaller system. In symbols, this condition says that A' is such that

$$\left| \frac{\partial \beta'}{\partial E'} Q' \right| \ll \beta' \quad (3 \cdot 6 \cdot 1)$$

Here $\partial \beta'/\partial E'$ is of the order of β'/\bar{E}' , where \bar{E}' is the mean energy of A' measured from its ground state,* while the heat Q' absorbed by A' is at most of the order of the mean energy \bar{E} of the small system A above its ground state. Hence one expects (3·6·1) to be valid if

$$\frac{\bar{E}}{\bar{E}'} \ll 1$$

i.e., if A' is sufficiently large compared to A .

Note that the concept of a heat reservoir is a relative one. A glass of tea acts approximately as a heat reservoir with respect to a slice of lemon immersed in it. On the other hand, it is certainly not a heat bath with respect to the whole room; indeed, the relationship there is the opposite.

If the macroscopic system A' has $\Omega'(E')$ accessible states and absorbs heat $Q' = \Delta \bar{E}'$, one can express the resulting change in $\ln \Omega'$ by a Taylor's expansion. Thus

$$\begin{aligned} \ln \Omega'(E' + Q') - \ln \Omega'(E') &= \left(\frac{\partial \ln \Omega'}{\partial E'} \right) Q' + \frac{1}{2} \left(\frac{\partial^2 \ln \Omega'}{\partial E'^2} \right) Q'^2 + \dots \\ &= \beta' Q' + \frac{1}{2} \frac{\partial \beta'}{\partial E'} Q'^2 + \dots \end{aligned} \quad (3 \cdot 6 \cdot 2)$$

where we have used the definition (3·3·9). But if A' acts as a heat reservoir so that (3·6·1) is satisfied, then β' does not change appreciably and higher-order terms on the right of (3·6·2) are negligible. Thus (3·6·2) reduces simply to

$$\ln \Omega'(E' + Q') - \ln \Omega'(E') = \beta' Q' = \frac{Q'}{kT'} \quad (3 \cdot 6 \cdot 3)$$

The left side expresses, by the definition (3·3·12), the entropy change of the heat reservoir. Thus one arrives at the simple result that, if a heat reservoir at temperature T' absorbs heat Q' , its resulting entropy change is given by

►
$$\Delta S' = \frac{Q'}{T'} \quad (\text{for a heat reservoir}) \quad (3 \cdot 6 \cdot 4)$$

A similar relation holds for any system which is at absolute temperature $T = (k\beta)^{-1}$ and which absorbs an *infinitesimal* amount of heat dQ from some

* Indeed, assuming the approximate dependence $\Omega' \propto E'^f$ of (2·5·10), it follows for $E' = \bar{E}'$ that $\beta' \approx (\partial \ln \Omega'/\partial \bar{E}') \approx f'/\bar{E}'$; hence $|\partial \beta'/\partial E'| \approx f'/\bar{E}'^2 \approx \beta'/\bar{E}'$.

other system at a slightly different temperature. Since $dQ \ll E$, where E is the energy of the system under consideration, it follows that

$$\ln \Omega(E + dQ) - \ln \Omega(E) = \frac{\partial \ln \Omega}{\partial E} dQ = \beta dQ$$

or, since $S = k \ln \Omega$, that

$$dS = \frac{dQ}{T} \quad (3 \cdot 6 \cdot 5)$$

where dS is the increase in entropy of the system.

3 · 7 Sharpness of the probability distribution

In Sec. 3 · 3 we argued that the probability $P(E)$ that A has an energy E exhibits a very sharp maximum. Let us now investigate more quantitatively just how sharp this maximum really is.

Our method of approach is identical to that used in Sec. 1 · 5. To investigate the behavior of $P(E)$ near its maximum $E = \tilde{E}$, we consider the more slowly varying function $\ln P(E)$ of (3 · 3 · 7) and expand it in a power series of the energy difference

$$\eta \equiv E - \tilde{E} \quad (3 \cdot 7 \cdot 1)$$

Expanding $\ln \Omega(E)$ in Taylor's series about \tilde{E} , one gets

$$\ln \Omega(E) = \ln \Omega(\tilde{E}) + \left(\frac{\partial \ln \Omega}{\partial E} \right) \eta + \frac{1}{2} \left(\frac{\partial^2 \ln \Omega}{\partial E^2} \right) \eta^2 + \dots \quad (3 \cdot 7 \cdot 2)$$

Here the derivatives are evaluated at $E = \tilde{E}$. Let us use the abbreviations

$$\beta \equiv \left(\frac{\partial \ln \Omega}{\partial E} \right) \quad (3 \cdot 7 \cdot 3)$$

$$\text{and} \quad \lambda \equiv - \left(\frac{\partial^2 \ln \Omega}{\partial E^2} \right) = - \left(\frac{\partial \beta}{\partial E} \right) \quad (3 \cdot 7 \cdot 4)$$

The minus sign has been introduced for convenience since we shall see that the second derivative is intrinsically negative.

Hence (3 · 7 · 2) can be written

$$\ln \Omega(E) = \ln \Omega(\tilde{E}) + \beta \eta - \frac{1}{2} \lambda \eta^2 + \dots \quad (3 \cdot 7 \cdot 5)$$

One can write down a corresponding expression for $\ln \Omega'(E')$ near $E' = \tilde{E}'$. By conservation of energy $E' = E^{(0)} - E$, so that

$$E' - \tilde{E}' = -(E - \tilde{E}) = -\eta \quad (3 \cdot 7 \cdot 6)$$

Analogously to (3·7·5) one thus obtains

$$\ln \Omega'(E') = \ln \Omega'(\tilde{E}') + \beta'(-\eta) - \frac{1}{2}\lambda'(-\eta)^2 + \dots \quad (3 \cdot 7 \cdot 7)$$

where β' and λ' are the parameters (3·7·3) and (3·7·4) correspondingly defined for system A' and evaluated at the energy $E' = \tilde{E}'$. Adding (3·7·5) and (3·7·7) gives

$$\ln [\Omega(E)\Omega'(E')] = \ln [\Omega(\tilde{E})\Omega'(\tilde{E}')] + (\beta - \beta')\eta - \frac{1}{2}(\lambda + \lambda')\eta^2 \quad (3 \cdot 7 \cdot 8)$$

At the maximum of $\Omega(E)\Omega'(E')$ it follows by (3·3·8) that $\beta = \beta'$, so that the term linear in η vanishes as it should. Hence (3·7·7) yields for (3·3·8) the result

$$\ln P(E) = \ln P(\tilde{E}) - \frac{1}{2}\lambda_0\eta^2$$

or

► $P(E) = P(\tilde{E}) e^{-\frac{1}{2}\lambda_0(E-\tilde{E})^2} \quad (3 \cdot 7 \cdot 9)$

where

$$\lambda_0 \equiv \lambda + \lambda' \quad (3 \cdot 7 \cdot 10)$$

Note that λ_0 cannot be negative, since then the probability $P(E)$ would not exhibit a maximum value, i.e., the combined system $A^{(0)}$ would not attain a well-defined final equilibrium situation as, physically, we know it must. Furthermore, neither λ nor λ' can be negative. Indeed, one could choose for A' a system for which $\lambda' \ll \lambda$; in that case $\lambda \approx \lambda_0$ and, since we already argued that this last quantity cannot be negative, it follows that $\lambda \geq 0$. Similar reasoning shows that $\lambda' \geq 0$.*

The same conclusion follows also from the argument that ordinarily $\Omega \propto E^f$. Indeed, using the definition (3·7·4), one obtains from (3·5·5)

$$\lambda = -\left(-\frac{f}{\tilde{E}^2}\right) = \frac{f}{\tilde{E}^2} > 0 \quad (3 \cdot 7 \cdot 11)$$

The preceding discussion leads to several interesting remarks. In Sec. 2·3 we concluded from the general behavior of the densities of states of the interacting systems that the probability $P(E)$ has a unique maximum at some energy \tilde{E} . We have now shown more specifically that, for E not too far from \tilde{E} , the probability $P(E)$ is described by the Gaussian distribution (3·7·9). It then follows by (1·6·8) that the mean energy \bar{E} is given by

$$\bar{E} = \tilde{E} \quad (3 \cdot 7 \cdot 12)$$

Thus the mean energy of A is indeed equal to the energy \tilde{E} corresponding to the situation of maximum probability. Furthermore, (3·7·9) shows that $P(E)$ becomes negligibly small compared to its maximum value when

* The equals sign corresponds to exceptional circumstances. An example might be a system consisting of a mixture of ice and water in equilibrium. The addition of energy to this system results in melting some of the ice, but does not change its temperature parameter. Thus $\lambda = -\partial\beta/\partial E = 0$.

$\frac{1}{2}\lambda_0(E - \bar{E})^2 \gg 1$, i.e., when $|E - \bar{E}| \gg \lambda_0^{-\frac{1}{2}}$. In other words, it is very improbable that the energy of A lies far outside the range $\bar{E} + \Delta^*E$ where *

$$\Delta^*E = \lambda_0^{-\frac{1}{2}} \quad (3 \cdot 7 \cdot 13)$$

Suppose now that A is the system with the larger value of the parameter λ . Then

$$\lambda_0 \approx \lambda \approx \frac{f}{\bar{E}^2} = \frac{f}{\bar{E}^2}$$

and

$$\Delta^*E \approx \frac{\bar{E}}{\sqrt{f}}$$

where \bar{E} is the mean energy of A above its ground state. The fractional width of the maximum in $P(E)$ is then given by

$$\blacktriangleright \quad \frac{\Delta^*E}{\bar{E}} \approx \frac{1}{\sqrt{f}} \quad (3 \cdot 7 \cdot 14)$$

If A contains a mole of particles, $f \approx N_a \approx 10^{24}$ and $(\Delta^*E/\bar{E}) \approx 10^{-12}$.

Hence the probability distribution has indeed an exceedingly sharp maximum when one is dealing with macroscopic systems containing very many particles. In our example, the probability $P(E)$ becomes already negligibly small if the energy E differs from its mean value by more than 1 part in 10^{12} ! This is an example of a general characteristic of macroscopic systems. Since the number of particles is so very large, fluctuations in any macroscopic parameter y (e.g., energy or pressure) are ordinarily utterly negligible. This means that one nearly always observes the mean value \bar{y} of the parameter and tends, therefore, to remain unaware of the statistical aspects of the macroscopic world. It is only when one is making very precise measurements or when one is dealing with very small systems that the existence of fluctuations becomes apparent.

The condition $\lambda \geq 0$ implies by (3 · 7 · 4) that

$$\begin{aligned} \lambda &= -\frac{\partial^2 \ln \Omega}{\partial E^2} = -\frac{\partial \beta}{\partial E} \geq 0 \\ \text{or} \quad \frac{\partial \beta}{\partial E} &\leq 0 \end{aligned} \quad (3 \cdot 7 \cdot 15)$$

Using $\beta = (kT)^{-1}$, the equivalent condition for T becomes

$$\frac{\partial \beta}{\partial T} \frac{\partial T}{\partial E} = -\frac{1}{kT^2} \frac{\partial T}{\partial E} \leq 0$$

Thus

$$\frac{\partial T}{\partial E} \geq 0 \quad (3 \cdot 7 \cdot 16)$$

i.e., the absolute temperature of any system increases with its energy.

The relation (3 · 7 · 15) allows one to establish the general connection between absolute temperature and the direction of heat flow. In the situation of Sec. 3 · 4, suppose that initially $\beta_i \neq \beta'_i$. If A absorbs positive heat Q ,

* The result (1 · 6 · 9) applied to the Gaussian distribution (3 · 7 · 9) shows that λ_0^{-1} is indeed the dispersion of the energy.

(3·7·15) implies that its value of β must decrease. At the same time, A' must give off heat so that its value of β' must increase. Since β is a continuous function of E for each system, the β values of the systems change in this way continuously until they reach the common final value β_f ; this must therefore be such that $\beta_f < \beta_i$ and $\beta_f > \beta'_i$. Thus $\beta_i > \beta'_i$ and the positive heat Q gets absorbed by the system with the higher value of β . Correspondingly, for ordinary positive absolute temperatures, positive heat gets absorbed by the system at the lower absolute temperature T .*

***Remark on the total number of accessible states** It is of some interest to calculate the total number of states $\Omega^{(0)}_{\text{tot}}$ accessible to the entire system $A^{(0)}$. Since the probability distribution is so sharply peaked, practically all states lie in a range within a width $\Delta^*E = \lambda_0^{-1}$ of \bar{E} (see Fig. 3·3·3). Since the density of states near $E = \bar{E}$ is equal to $\Omega^{(0)}(\bar{E})/\delta E$, the total number of states is approximately given by

$$\Omega^{(0)}_{\text{tot}} \approx \frac{\Omega^{(0)}(\bar{E})}{\delta E} \Delta^*E = K\Omega(\bar{E}) \quad (3\cdot7\cdot17)$$

where

$$K = \frac{\Delta^*E}{\delta E} \quad (3\cdot7\cdot18)$$

By (3·7·17) it follows that

$$\ln \Omega^{(0)}_{\text{tot}} = \ln \Omega^{(0)}(\bar{E}) + \ln K \approx \ln \Omega(\bar{E}) \quad (3\cdot7\cdot19)$$

The last result is true to excellent approximation because $\ln K$ is utterly negligible. This is another striking consequence of the fact that we are dealing with such large numbers. The reason is that, no matter what reasonable energy subdivision interval we may choose, the number K in (3·7·18) should certainly not be much bigger than f , where f may be of the order of Avogadro's number N_a . Thus $\ln K$ is of order $\ln f$ or less. On the other hand, (2·5·9) shows that $\ln \Omega$ is some number of the order of f . But when f is large, say $f \approx N_a \approx 10^{24}$, $f >> \ln f$ (10^{24} compared to only 55). Thus $\ln K$ is completely negligible compared to $\ln \Omega(\bar{E})$. The relation (3·7·19) asserts then that the probability distribution is so sharply peaked around its maximum that, for purposes of calculating logarithms, the total number of states is equal to the maximum number of states. From this it also follows that, if in analogy to (3·3·13), one defines the total entropy of $A^{(0)}$ in terms of the total number of its accessible states so that

$$S^{(0)} = k \ln \Omega^{(0)}_{\text{tot}}$$

Then

$$S^{(0)} = k \ln \Omega^{(0)}(\bar{E}) = k \ln [\Omega(\bar{E})\Omega'(\bar{E}')] = k \ln \Omega(\bar{E}) + k \ln \Omega'(\bar{E}') \\ \text{or} \quad S^{(0)} = S(\bar{E}) + S'(\bar{E}') \quad (3\cdot7\cdot20)$$

Thus the entropy so defined has the above simple additive property.

* The statement relating to T is restricted to the ordinary case of positive temperatures, since otherwise T is not a continuous function of E . Indeed, for spin systems with an upper bound of possible energies, Ω has a maximum, and hence β passes continuously through the value $\beta = 0$. But correspondingly, $T = (k\beta)^{-1}$ jumps from ∞ to $-\infty$.

GENERAL INTERACTION BETWEEN MACROSCOPIC SYSTEMS

3 · 8 Dependence of the density of states on the external parameters

Now that we have examined in detail the thermal interaction between systems, let us turn to the general case where mechanical interaction can also take place, i.e., where the external parameters of the systems are also free to change. We begin, therefore, by investigating how the density of states depends on the external parameters.

For the sake of simplicity, consider the situation where only a single external parameter x of the system is free to vary; the generalization to the case where there are several such parameters will be immediate. The number of states accessible to this system in the energy range between E and $E + \delta E$ will also depend on the particular value assumed by this external parameter; we shall denote it by $\Omega(E,x)$. We are interested in examining how Ω depends on x .

When x is changed by an amount dx , the energy $E_r(x)$ of each microstate r changes by an amount $(\partial E_r/\partial x) dx$. The energies of different states are, in general, changed by different amounts. Let us denote by $\Omega_Y(E,x)$ the number of those states which have an energy in the range between E and $E + \delta E$ when the external parameter has the value x , and which are such that their derivative $\partial E_r/\partial x$ has a value in range between Y and $Y + \delta Y$. The total number of states is then given by

$$\Omega(E,x) = \sum_Y \Omega_Y(E,x) \quad (3 \cdot 8 \cdot 1)$$

where the summation is over all possible values of Y .

Consider a particular energy E . When the external parameter is changed, some states which originally had an energy less than E will acquire an energy greater than E , and vice versa. What then is the total number of states $\sigma(E)$ whose energy is changed from a value less than E to a value greater than E when the parameter changes from x to $x + dx$? Those states for which $\partial E_r/\partial x$ has the particular value Y change their energy by the infinitesimal amount $Y dx$. Hence all these states located within an energy $Y dx$ below E will change their energy from a value smaller to one greater than E (see Fig. 3 · 8 · 1). The number $\sigma_Y(E)$ of such states is thus given by the number per unit

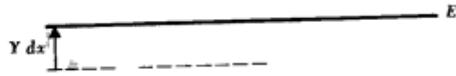


Fig. 3 · 8 · 1 The shaded area indicates the energy range occupied by states with a value of $\partial E_r/\partial x = Y$, whose energy changes from a value smaller than E to one greater than E when the external parameter is changed from x to $x + dx$.

energy multiplied by the energy range $Y dx$, i.e., by

$$\sigma_Y(E) = \frac{\Omega_Y(E,x)}{\delta E} Y dx \quad (3 \cdot 8 \cdot 2)$$

Different states have their energy changed by different amounts $Y dx$ (positive or negative). Hence the total number of states $\sigma(E)$ whose energy is changed from a value less to a value greater than E is given by summing (3·8·2) over all possible values of $\partial E_r / \partial x = Y$; thus

$$\sigma(E) = \sum_Y \frac{\Omega_Y(E,x)}{\delta E} Y dx = \frac{\Omega(E,x)}{\delta E} \bar{Y} dx \quad (3 \cdot 8 \cdot 3)$$

where we have used the definition

$$\bar{Y} = \frac{1}{\Omega(E,x)} \sum_Y \Omega_Y(E,x) Y \quad (3 \cdot 8 \cdot 4)$$

as the mean value of Y over all accessible states, each state being considered equally likely. The mean value thus defined is, of course, a function of E and x , i.e., $\bar{Y} = \bar{Y}(E,x)$. Since $Y = \partial E_r / \partial x$, one has

$$\bar{Y} = \frac{\partial \bar{E}_r}{\partial x} \equiv -\bar{X} \quad (3 \cdot 8 \cdot 5)$$

where \bar{X} is, by the definition (2·9·6), the mean generalized force conjugate to the external parameter x .

Let us now consider the total number of states $\Omega(E,x)$ between E and $E + \delta E$ (see Fig. 3·8·2). When the parameter changes from x to $x + dx$, the number of states in this energy range changes by an amount $[\partial \Omega(E,x)/\partial x] dx$ which must be due to [the net number of states which enter this range by having their energy changed from a value less than E to one greater than E] minus [the net number of states which leave this range by having their energy changed from a value less than $E + \delta E$ to one greater than $E + \delta E$]. In symbols this can be written

$$\frac{\partial \Omega(E,x)}{\partial x} dx = \sigma(E) - \sigma(E + \delta E) = -\frac{\partial \sigma}{\partial E} \delta E$$

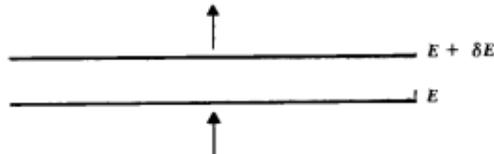
Using (3·8·3) this becomes

$$\frac{\partial \Omega}{\partial x} = -\frac{\partial}{\partial E} (\Omega \bar{Y}) \quad (3 \cdot 8 \cdot 6)$$

or

$$\frac{\partial \Omega}{\partial x} = -\frac{\partial \Omega}{\partial E} \bar{Y} - \Omega \frac{\partial \bar{Y}}{\partial E}$$

Fig. 3·8·2 The number of states in the shaded energy range changes when the external parameter is varied because states enter and leave this energy range.



By dividing both sides by Ω , this can be written

$$\frac{\partial \ln \Omega}{\partial x} = - \frac{\partial \ln \Omega}{\partial E} \bar{Y} - \frac{\partial \bar{Y}}{\partial E} \quad (3 \cdot 8 \cdot 7)$$

But for a large system the first term on the right is of the order of $(f/E) \bar{Y}$, since one has approximately $\Omega \propto E^f$. On the other hand, the second term on the right is only of the order of $\partial \bar{Y} / \partial E \approx \bar{Y}/E$ and is therefore smaller than the first term by a factor f . For macroscopic systems where f is of the order of Avogadro's number, the second term on the right is thus utterly negligible. Hence (3 · 8 · 7) becomes to excellent approximation

$$\frac{\partial \ln \Omega}{\partial x} = - \frac{\partial \ln \Omega}{\partial E} \bar{Y} = \beta \bar{X} \quad (3 \cdot 8 \cdot 8)$$

where we have used (3 · 8 · 5) and the definition (3 · 3 · 9) of the temperature parameter β .

When there are several external parameters x_1, \dots, x_n (so that $\Omega = \Omega(E, x_1, \dots, x_n)$) the above relation is clearly valid for each one of them. One obtains then for each external parameter x_α and its corresponding mean generalized force \bar{X}_α the general relation

$$\frac{\partial \ln \Omega}{\partial x_\alpha} = \beta \bar{X}_\alpha \quad (3 \cdot 8 \cdot 9)$$

3 · 9 Equilibrium between interacting systems

Consider two systems A and A' which can interact both by exchanging heat and by doing work on each other. A specific example might be the situation illustrated in Fig. 2 · 8 · 1 where the piston is free to move and is capable of conducting heat. System A has energy E and is characterized by some external parameters x_1, \dots, x_n which are free to vary. Similarly, A' has energy E' and is characterized by some external parameters x'_1, \dots, x'_n . The combined system $A^{(0)} \equiv A + A'$ is isolated. Hence

$$E + E' = E^{(0)} = \text{constant} \quad (3 \cdot 9 \cdot 1)$$

The energy E' of A' is then determined if the energy E of A is known. Furthermore, if A and A' can interact mechanically, this implies that the parameters x' are some functions of the parameters x .

Example In Fig. 2 · 8 · 1, the gas A is described by one external parameter x , its volume V ; similarly, the gas A' is described by its volume V' . But as the piston moves, the total volume remains unchanged, i.e.,

$$V + V' = V^{(0)} = \text{constant} \quad (3 \cdot 9 \cdot 2)$$

The total number of states accessible to $A^{(0)}$ is thus a function of E and of the parameters $x_\alpha (\alpha = 1, \dots, n)$. Once again $\Omega^{(0)}(E; x_1, \dots, x_n)$ will

have a very sharp maximum for some values $E = \bar{E}$ and $x_\alpha = \bar{x}_\alpha$. The equilibrium situation corresponds then to the one of maximum probability where practically all systems $A^{(0)}$ have values of E and x_α very close to \bar{E} and \bar{x}_α . The mean values of these quantities in equilibrium will thus be equal to $\bar{E} = \bar{E}$ and $\bar{x}_\alpha = \bar{x}_\alpha$.

Infinitesimal quasi-static process Consider a quasi-static process in which the system A , by virtue of its interaction with system A' , is brought from an equilibrium state described by \bar{E} and $\bar{x}_\alpha (\alpha = 1, \dots, n)$ to an infinitesimally different equilibrium state described by $\bar{E} + d\bar{E}$ and $\bar{x}_\alpha + d\bar{x}_\alpha$. What is the resultant change in the number of states Ω accessible to A ?

Since $\Omega = \Omega(E; x_1, \dots, x_n)$, one can write for the resultant change in $\ln \Omega$ the purely mathematical result

$$d \ln \Omega = \frac{\partial \ln \Omega}{\partial E} d\bar{E} + \sum_{\alpha=1}^n \frac{\partial \ln \Omega}{\partial x_\alpha} d\bar{x}_\alpha \quad (3 \cdot 9 \cdot 3)$$

By using the relation (3·8·9), this can be written

$$d \ln \Omega = \beta(d\bar{E} + \sum_{\alpha} \bar{X}_\alpha d\bar{x}_\alpha) \quad (3 \cdot 9 \cdot 4)$$

By (2·9·5) the last term in the parentheses is just the macroscopic work dW done by A in this infinitesimal process. Hence (3·9·4) can be written

$$d \ln \Omega = \beta(d\bar{E} + dW) \equiv \beta dQ \quad (3 \cdot 9 \cdot 5)$$

where we have used the definition (2·8·3) for the infinitesimal heat absorbed by A . Equation (3·9·5) is a fundamental relation valid for any quasi-static infinitesimal process. By (3·3·10) and (3·3·12) it can also be written in the form

► $dQ = T dS = d\bar{E} + dW \quad (3 \cdot 9 \cdot 6)$

or equivalently

► $dS = \frac{dQ}{T} \quad (3 \cdot 9 \cdot 7)$

Thus the relation (3·6·5) remains valid even if the external parameters of the system are changed quasi-statically. Note that in the special case when the system is thermally insulated (i.e., when the process is "adiabatic") the absorbed heat $dQ = 0$ and (3·9·7) asserts that

$$dS = 0$$

This shows that S , or $\ln \Omega$, does not change even if the external parameters are varied *quasi-statically* by a *finite* amount. Hence one has the important result that

If the external parameters of a *thermally isolated* system are changed *quasi-statically* by any amount, $\Delta S = 0$. (3 · 9 · 8)

Thus the performance of quasi-static work changes the energy of a thermally isolated system, but does not affect the number of states accessible to it. In accordance with the discussion of Sec. 3 · 2, such a process is thus reversible.

It is worth emphasizing that even if a system is thermally isolated so that it absorbs *no* heat, its entropy will *increase* if processes take place which are *not* quasi-static. For example, each total system $A^{(0)}$ is thermally isolated in the three examples discussed at the beginning of Sec. 3 · 1, yet the number of states accessible to it, and hence its entropy, increases.

Equilibrium conditions Consider the equilibrium between the systems A and A' in the simple case where the external parameters are the volumes V and V' of the two systems. The number of states available to the combined system $A^{(0)}$ is, as in (3 · 3 · 5), given by the simple product

$$\Omega^{(0)}(E, V) = \Omega(E, V)\Omega'(E', V') \quad (3 \cdot 9 \cdot 9)$$

where E' and V' are related to E and V by (3 · 9 · 1) and (3 · 9 · 2).

Taking logarithms of (3 · 9 · 9), one gets

$$\ln \Omega^{(0)} = \ln \Omega + \ln \Omega' \quad (3 \cdot 9 \cdot 10)$$

or $S^{(0)} = S + S' \quad (3 \cdot 9 \cdot 11)$

The maximum value of $\Omega^{(0)}$ or $S^{(0)}$ is then determined by the condition that

$$d \ln \Omega^{(0)} = d(\ln \Omega + \ln \Omega') = 0 \quad (3 \cdot 9 \cdot 12)$$

for arbitrary changes of dE and dV .

But $d \ln \Omega = \frac{\partial \ln \Omega}{\partial E} dE + \frac{\partial \ln \Omega}{\partial V} dV = \beta dE + \beta \bar{p} dV \quad (3 \cdot 9 \cdot 13)$

where we have used (3 · 8 · 8) and the generalized force $\bar{X} = -(\partial \bar{E}_r / \partial \bar{V})$ is, by (2 · 10 · 2), just the mean pressure \bar{p} exerted by A . Similarly, one has for A'

$$d \ln \Omega' = \beta' dE' + \beta' \bar{p}' dV' = -\beta' dE - \beta' \bar{p}' dV \quad (3 \cdot 9 \cdot 14)$$

since the conservation conditions (3 · 9 · 1) and (3 · 9 · 2) imply that $dE' = -dE$ and $dV' = -dV$, respectively. Hence the condition of maximum entropy (3 · 9 · 12) becomes

$$(\beta - \beta') dE + (\beta \bar{p} - \beta' \bar{p}') dV = 0 \quad (3 \cdot 9 \cdot 15)$$

Since this must be satisfied for arbitrary values of dE and dV , it follows that the coefficients of both of these differentials must separately vanish.

Hence

$$\left. \begin{aligned} \beta - \beta' &= 0 \\ \beta \bar{p} - \beta' \bar{p}' &= 0 \end{aligned} \right\} \quad (3 \cdot 9 \cdot 16)$$

and

or

and

As might be expected, these conditions simply assert that the temperatures of the systems must be equal to guarantee their thermal equilibrium and that their mean pressures must be equal to guarantee their mechanical equilibrium.

As a particularly simple example consider the mechanical interaction of a system A with a purely mechanical device A' whose energy E' is a function only of some external parameter x . The situation might be the one illustrated in Fig. 3·9·1 where A' is a spring whose elongation is measured by the distance x . The total number $\Omega^{(0)}$ of states accessible to the system $A + A'$ is then simply proportional to the number of states $\Omega(E, x)$ accessible to A .* But if $E^{(0)}$ is the constant *total* energy, the energy of A is

$$E = E^{(0)} - E'(x) \quad (3 \cdot 9 \cdot 17)$$

and is thus a function of x . If x is free to adjust itself, it will tend to change so that the system approaches an equilibrium situation where Ω is a maximum, i.e., where

$$\frac{\partial}{\partial x} \ln \Omega(E, x) = 0$$

This means that

$$\frac{\partial \ln \Omega}{\partial E} \frac{\partial E}{\partial x} + \frac{\partial \ln \Omega}{\partial x} = 0$$

By (3·9·17) and (3·8·9) this becomes

$$\beta \left(-\frac{\partial E'}{\partial x} \right) + \beta \bar{X} = 0$$

or

$$\bar{X} = \frac{\partial E'}{\partial x}$$

This condition asserts simply that in equilibrium the mean force \bar{X} exerted by the gas A must be equal to the force $\partial E'/\partial x$ exerted by the spring.

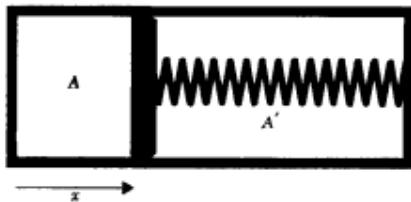


Fig. 3·9·1 A gas A interacting through a movable piston with a spring A' .

3 · 10 Properties of the entropy

Entropy and exact differentials In Sec. 2·9 we discussed the fact that the infinitesimal heat dQ is *not* an exact differential. The relation (3·9·7) establishes, however, the following remarkable result: Although the heat dQ absorbed in a quasi-static infinitesimal process is *not* an exact differential, the quantity

* The number of states accessible to A' remains unchanged—this is what one means by a purely mechanical device which is completely described by its external parameters; i.e., A' itself is thermally insulated, and its external parameters are changed sufficiently slowly (compared to its significant relaxation time) that its entropy remains constant.

$dQ/T = dS$ obtained by dividing dQ by the absolute temperature T of the system is an exact differential. This is so because the entropy S is a function characteristic of each macrostate of the system, and dS is simply the difference between two such functions for adjacent macrostates.

Remark If multiplication by a factor converts an inexact differential into an exact differential, then this factor is in mathematical terminology said to be an "integrating factor" for the inexact differential. Thus one can say that the absolute temperature T is characterized by the property that T^{-1} is an integrating factor for dQ .

It follows that, given any two macrostates i and f of a system, the entropy difference between them can be written as

$$S_f - S_i = \int_i^f dS = \int_i^f \frac{dQ}{T} \quad (3 \cdot 10 \cdot 1)$$

where we have used the result (3 · 9 · 7) in the last integral. We have labeled this integral by the subscript "eq" (standing for "equilibrium") to emphasize explicitly the fact that it is to be evaluated for any process by which the system is brought *quasi-statically* through a sequence of near-equilibrium situations from its initial to its final macrostates. The temperature in the integrand is thus well defined at all stages of the process. Since the left side of (3 · 10 · 1) refers only to the initial and final macrostates, it follows that the integral on the right side must be independent of whatever particular quasi-static process leading from i to f is chosen to evaluate the integral. In symbols,

$$\int_i^f \frac{dQ}{T} \quad \text{is independent of the process} \quad (3 \cdot 10 \cdot 2)$$

Example Consider the two quasi-static processes indicated by the solid and dashed lines of the \bar{p} versus V diagram in Fig. 2 · 11 · 2. The integral $\int_i^f dQ$ giving the total heat absorbed in going from i to f will be different when evaluated for the two processes. But the integral $\int_i^f (dQ/T)$ will yield the same result when evaluated for these processes.

Let us be clear as to how such an integral is to be evaluated. At any stage of the process the system is characterized by a certain value of V and corresponding value \bar{p} given by the graph. This information is adequate to determine a definite temperature T for this macrostate of the system. In going to an adjacent value of V , an amount of heat dQ is absorbed. Hence one knows both T and dQ , and can sum all the successive quantities dQ/T as one proceeds to increase the volume from V_i to V_f .

Implications of the statistical definition of entropy In (3·3·12) we defined the entropy in terms of the number Ω of accessible states in the range between E and $E + \delta E$ by the relation

$$S = k \ln \Omega \quad (3 \cdot 10 \cdot 3)$$

It is important to note that if the macrostate of a system is specified, i.e., if one knows the external parameters and the energy E of the system, then the number Ω of states accessible to the system is completely determined if the system is discussed in terms of quantum mechanics. Hence the entropy S has, by (3·10·3), a *unique* value calculable from a knowledge of the microscopic constitution of the system.*

Remark This last statement would *not* be true if the system were described in terms of classical mechanics. If the system has f degrees of freedom, phase space is subdivided (as in Sec. 2·1) into cells of arbitrarily chosen volume h_0^f . The total number of cells, or states, available to the system is then obtained by dividing the accessible volume of phase space (contained between the energies E and $E + \delta E$) by the volume per cell. Thus

$$\Omega = \frac{1}{h_0^f} \int \cdots \int dq_1 \cdots dq_f dp_1 \cdots dp_f$$

or $S = k \ln \left(\int \cdots \int dq_1 \cdots dq_f dp_1 \cdots dp_f \right) - kf \ln h_0 \quad (3 \cdot 10 \cdot 4)$

These relations show that Ω depends in an essential way on the size of the cells into which phase space is subdivided. Correspondingly, S contains an additive constant which depends on this cell size. In a classical description the value of the entropy is thus not unique, but only defined to within an arbitrary additive constant. What happens effectively in quantum mechanics is that there exists a natural unit of cell size according to which h_0 is to be put equal to Planck's constant; the additive constant becomes then uniquely determined.

Limiting behavior of the entropy As one goes to lower energy, every system described by quantum mechanics approaches the lowest possible energy E_0 of its ground state. Corresponding to this energy there exists usually only one possible state of the system; or there may be a relatively small number of such states, all of the same energy E_0 (the ground state is then said to be "degenerate"). When one considers energies somewhat greater than E_0 , the number of states $\Omega(E)$ increases, of course, very rapidly; by (2·5·10), it then behaves roughly like $\Omega \propto (E - E_0)^f$ if f is the number of degrees of freedom of the system. The general dependence of $\ln \Omega$ on the energy E of the system is thus of the form sketched in Fig. 3·10·1.

* The value of S is certainly unique for a given choice of the energy subdivision interval δE . In addition, we showed at the end of Sec. 3·3 that the value of the entropy is also utterly insensitive to the exact choice of the magnitude of δE .

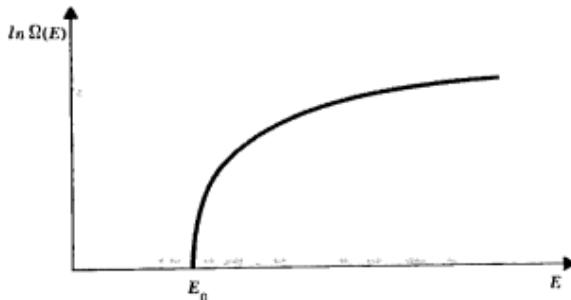


Fig. 3·10·1 Behavior of $\ln \Omega(E)$ for energies $E > E_0$. Note that β , the slope of the curve, becomes very large for $E \rightarrow E_0$ and that $d\beta/dE < 0$.

Whenever the energy of the system is significantly greater than its ground state energy E_0 , its entropy S is the order of $k \ln \Omega$, i.e., by (2·5·9), of the order of kf . As its energy approaches E_0 , the number of states $\Omega(E)$ within the given interval δE falls off rapidly; it ultimately becomes a number of the order of f itself or smaller since the ground state itself consists of only one state, or at most a relatively small number of states. But then $S = k \ln \Omega$ approaches a number of the order of $k \ln f$ or less, and this is utterly negligible compared to the magnitude kf of the entropy at higher energies. Thus one can assert to excellent approximation that the entropy becomes vanishingly small as the system approaches its ground-state energy. In symbols,

$$\text{as } E \rightarrow E_0, \quad S \rightarrow 0 \quad (3 \cdot 10 \cdot 5)$$

Remark We note again the quantum-mechanical basis for the validity of the above discussion. In the framework of classical mechanics there would not exist a situation of lowest energy with an associated definite small number of states.

This limiting behavior of S can also be expressed in terms of the temperature of the system. We know by (3·7·15) that $\partial\beta/\partial E < 0$, or equivalently that $\partial T/\partial E > 0$. Hence it follows that as E decreases toward E_0 , β increases and becomes very large, while $T = (k\beta)^{-1}$ decreases and becomes very small. In the limiting case as $T \rightarrow 0$, E must increasingly approach its ground-state value E_0 . By virtue of (3·10·5), the entropy must then become negligibly small. Thus,

$$\text{as } T \rightarrow 0, \quad S \rightarrow 0 \quad (3 \cdot 10 \cdot 6)$$

In applying (3·10·6) to situations of experimental interest one must, as usual, be sure to ascertain that equilibrium arguments are applicable to the system under consideration. At very low temperatures one must be particularly careful in this respect since the rate of attaining equilibrium may then become quite slow. Another question which may arise in practice concerns the extent to which the limiting situation $T \rightarrow 0$ has indeed been reached. In

other words, how small a temperature is sufficiently small to apply (3·10·6)?

It is only possible to answer this question by knowing something about the particular system under consideration. A case of very frequent occurrence is that where the nuclei of the atoms in the system have nuclear spins. If one brings such a system to some sufficiently low temperature T_0 , the entropy (or number of states) associated with the degrees of freedom *not* involving nuclear spins may have become quite negligible. Yet the number of states Ω_s corresponding to the possible nuclear spin orientations may be very large, indeed just as large as at very much higher temperatures. The reason is that nuclear magnetic moments are very small; interactions affecting the spin orientation of a nucleus are thus also very small. Hence each nuclear spin is ordinarily oriented completely at random, even at temperatures as low as T_0 .

Suppose, for example, that the system consists of atoms having nuclear spin $\frac{1}{2}$ (e.g., the system might be a silver spoon). Each spin can have two possible orientations. If spin-independent interactions are very small, these two orientations do not differ significantly in energy and each spin is equally likely to point "up" or "down." If there are N nuclei in the system, there are then $\Omega_s = 2^N$ possible spin states accessible to the system, even at a temperature as low as T_0 . Only at temperatures very much less than T_0 will the interactions involving nuclear spins become of significance. For instance, the situation where all nuclear spins are aligned parallel to each other ("nuclear ferromagnetism") may have slightly lower energy than other spin orientations. The system will then (if it attains equilibrium) settle down at the very lowest temperatures ($T \ll T_0$) into this *one* very lowest state where all nuclei are aligned.

- - -

In the case under discussion, where one knows that the nuclear spins are still randomly oriented at a temperature as low as T_0 and do not depart from this random configuration until one goes to temperatures very much less than T_0 , one can still make a useful statement. Thus one can assert that as T decreases towards T_0 , the entropy approaches a value S_0 which is simply given by the number of possible nuclear spin orientations, i.e., $S_0 = k \ln \Omega_s$. In symbols this yields the result that

$$\text{as } T \rightarrow 0_+, \quad S \rightarrow S_0 \quad (3\cdot10\cdot7)$$

Here $T \rightarrow 0_+$ denotes a limiting temperature which is very small, yet large enough so that spins remain randomly oriented. Furthermore, S_0 is a definite constant which depends only on the kinds of atomic nuclei of which the system is composed, but which is completely independent of any details concerned with the energy levels of the system. One can say, in short, that S_0 is independent of all parameters of the system in the widest sense, i.e., independent of the spatial arrangement of its atoms or of the interactions between them. The statement (3·10·7) is a useful one because it can be applied at temperatures which are not prohibitively low.

SUMMARY OF FUNDAMENTAL RESULTS

3 · 11 Thermodynamic laws and basic statistical relations

The entire discussion of this chapter has been based on the fundamental statistical postulates of Sec. 2 · 3. We have elaborated some details and given various illustrative examples to gain familiarity with the main properties of macroscopic systems. The main ideas have, however, been quite simple; most of them are contained in Secs. 3 · 1, 3 · 3, and 3 · 9. The discussion of this chapter encompasses, nevertheless, all the fundamental results of classical thermodynamics and all the essential results of statistical mechanics. Let us now summarize these results in the form of very general statements and group them into two categories. The first of these will consist of purely *macroscopic* statements which make *no* reference to the *microscopic* properties of the systems, i.e., to the molecules of which they consist. These statements we shall call “thermodynamic laws.” The other class of statements will refer to the microscopic properties of the systems and will be designated as “statistical relations.”

Thermodynamic laws The first statement is the fairly trivial one discussed in Sec. 3 · 5.

► **Zeroth law:** If two systems are in thermal equilibrium with a third system, they must be in thermal equilibrium with each other.

Next there is a statement expressing the conservation of energy and discussed in Sec. 2 · 8.

► **First law:** An equilibrium macrostate of a system can be characterized by a quantity \bar{E} (called “internal energy”) which has the property that for an isolated system, $\bar{E} = \text{constant}$ (3 · 11 · 1)

If the system *is* allowed to interact and thus goes from one macrostate to another, the resulting change in \bar{E} can be written in the form

$$\Delta\bar{E} = -W + Q \quad (3 \cdot 11 \cdot 2)$$

where W is the macroscopic work done *by* the system as a result of the system’s change in external parameters. The quantity Q , *defined* by (3 · 11 · 2), is called the “heat absorbed by the system.”

Next we were led to introduce the entropy S which had the simple properties discussed in Secs. 3 · 1 and 3 · 9. Thus we obtain the following results:

► **Second law:** An equilibrium macrostate of a system can be characterized by a quantity S (called “entropy”), which has the properties that
a. In any process in which a thermally *isolated* system goes from one macrostate to another, the entropy tends to increase, i.e.,

$$\Delta S \geq 0 \quad (3 \cdot 11 \cdot 3)$$

b. If the system is not isolated and undergoes a quasi-static infinitesimal process in which it absorbs heat dQ , then

$$dS = \frac{dQ}{T} \quad (3 \cdot 11 \cdot 4)$$

where T is a quantity characteristic of the macrostate of the system. (T is called the "absolute temperature" of the system.)

Finally, there is a last statement based on (3 · 10 · 7).

► *Third law:* The entropy S of a system has the limiting property that

$$\text{as } T \rightarrow 0_+, \quad S \rightarrow S_0 \quad (3 \cdot 11 \cdot 5)$$

where S_0 is a constant independent of all parameters of the particular system.

Note again that the above four laws are completely *macroscopic* in content. We have introduced three quantities (\bar{E} , S , and T) which are asserted to be defined for each macrostate of the system, and we have made some statements about properties relating these quantities. But *nowhere* in these four laws have we made any explicit reference to the microscopic nature of the system (e.g., to the properties of the molecules constituting the system or to the forces of interaction between them).

Statistical relations There is first the general connection (3 · 3 · 12)

$$S = k \ln \Omega \quad (3 \cdot 11 \cdot 6)$$

This allows one to relate the quantities entering in the thermodynamic laws to one's microscopic knowledge of the system. Indeed, if one knows the nature of the particles constituting the system and the interactions between them, then one can in principle use the laws of mechanics to compute the possible quantum states of the system and thus to find Ω .

Furthermore, one can use the fundamental statistical postulate of Sec. 2 · 3 to make statements about the probability P of finding an isolated system in a situation characterized by certain parameters y_1, \dots, y_n . If the corresponding number of accessible states is Ω , then in equilibrium

$$P \propto \Omega \propto e^{S/k} \quad (3 \cdot 11 \cdot 7)$$

A large number of conclusions follow from the purely *macroscopic* statements which we have called the laws of thermodynamics. Indeed, the whole discipline of classical thermodynamics assumes these laws as basic postulates and then proceeds to deduce their consequences in a macroscopic discussion which never refers to the microscopic description of matter in terms of atoms or molecules. The approach is sufficiently fruitful to have given rise to a large body of important results. We shall discuss it further, particularly in Chapter 5. This approach was also historically the oldest, since it could arise in a context where the atomic constitution of matter was not yet known or suffi-

ciently understood.* If one does make use of microscopic information and uses statistical mechanics to calculate Ω , one's powers of prediction are, of course, tremendously increased. Not only can one then calculate thermodynamic quantities from first principles by using (3·11·6), but one can also calculate probabilities and thus the fluctuations of physical quantities about their mean values. Statistical mechanics is thus the more inclusive discipline which encompasses all of classical thermodynamics; to emphasize this fact it is sometimes called "statistical thermodynamics."

3 · 12 Statistical calculation of thermodynamic quantities

It is worth pointing out explicitly how a knowledge of the number of states $\Omega = \Omega(E; x_1, \dots, x_n)$ of a system allows one to calculate important macroscopic quantities characterizing the system in equilibrium. The quantity Ω is a function of the energy of the system under consideration and of its external parameters. The relations of particular interest are (3·3·9) and (3·8·9), namely,

$$\beta = \frac{\partial \ln \Omega}{\partial E} \quad \text{and} \quad \bar{X}_a = \frac{1}{\beta} \frac{\partial \ln \Omega}{\partial x_a} \quad (3 \cdot 12 \cdot 1)$$

These allow one to compute the absolute temperature and the mean generalized forces of the system from a knowledge of Ω . For example, in the particular case when $x_a = V$ is the volume V of the system, the corresponding mean generalized force \bar{X}_a is, by (2·10·2), the mean pressure \bar{p} given by

$$\bar{p} = \frac{1}{\beta} \frac{\partial \ln \Omega}{\partial V} \quad (3 \cdot 12 \cdot 2)$$

The equations (3·12·1) permit one to find relations connecting the generalized forces, the external parameters, and the absolute temperature T . Such relations are called "equations of state" and are important since they relate parameters that are readily measured by experiment. For example, one can find how the mean pressure \bar{p} depends on the temperature T and volume V of the system; the relation $\bar{p} = \bar{p}(T, V)$ would be the corresponding "equation of state."

Remark Note that the relations (3·12·1) are implied by the statement (3·11·4). The latter yields for an infinitesimal quasi-static process the entropy change

$$dS = \frac{dQ}{T} = \frac{1}{T} \left(d\bar{E} + \sum_{a=1}^n \bar{X}_a dx_a \right) \quad (3 \cdot 12 \cdot 3)$$

* See the references at the end of this chapter for books treating the subject entirely from a macroscopic point of view.

But since the entropy S is a function of the energy and the external parameters, one can also write the purely mathematical result

$$dS = \left(\frac{\partial S}{\partial \bar{E}} \right) d\bar{E} + \sum_{a=1}^n \left(\frac{\partial S}{\partial \bar{x}_a} \right) d\bar{x}_a \quad (3 \cdot 12 \cdot 4)$$

Since (3 · 12 · 3) and (3 · 12 · 4) must be identically equal for arbitrary values of $d\bar{E}$ and $d\bar{x}_a$, the corresponding coefficients multiplying the differentials must be equal. Hence one obtains

$$\frac{1}{T} = \frac{\partial S}{\partial \bar{E}} \quad \text{and} \quad \frac{\bar{x}_a}{T} = \left(\frac{\partial S}{\partial \bar{x}_a} \right) \quad (3 \cdot 12 \cdot 5)$$

where the derivatives are to be evaluated corresponding to the energy $\bar{E} = \bar{E}$ and parameters $\bar{x}_a = \bar{x}_a$ of the equilibrium state under consideration. The relations (3 · 12 · 5) are identical to (3 · 12 · 1), since $S = k \ln \Omega$ and $T = (k\beta)^{-1}$.

Let us illustrate the preceding comments by applying them to a very simple but important system, an ideal gas. In (2 · 5 · 14) we showed that for an ideal gas of N molecules in a volume V the quantity Ω is of the form

$$\Omega \propto V^N \chi(E) \quad (3 \cdot 12 \cdot 6)$$

where $\chi(E)$ is independent of V and depends only on the energy E of the gas. Hence

$$\ln \Omega = N \ln V + \ln \chi(E) + \text{constant} \quad (3 \cdot 12 \cdot 7)$$

Thus (3 · 12 · 2) yields immediately for the mean pressure of the gas the simple relation

$$\bar{p} = \frac{N}{V} \frac{1}{k} = \frac{N}{V} kT \quad (3 \cdot 12 \cdot 8)$$

or

$$\bar{p} = nkT \quad (3 \cdot 12 \cdot 9)$$

where $n \equiv N/V$ is the number of molecules per unit volume. This is the equation of state for an ideal gas. Alternatively one can write $N = \nu N_a$, where ν is the number of moles of gas present and N_a is Avogadro's number. Then (3 · 12 · 8) becomes

$$\bar{p}V = \nu RT \quad (3 \cdot 12 \cdot 10)$$

where $R \equiv N_a k$ is called the "gas constant." Note that neither the equation of state nor the constant R depends on the kind of molecules constituting the ideal gas.

By (3 · 12 · 1) and (3 · 12 · 7) one obtains further

$$\beta = \frac{\partial \ln \chi(E)}{\partial E}$$

evaluated for the mean energy $E = \bar{E}$ of the gas. Here the right side is only a function of E , but *not* of V . Thus it follows that, for an ideal gas $\beta = \beta(\bar{E})$

or
$$\bar{E} = \bar{E}(T) \quad (3 \cdot 12 \cdot 11)$$

Hence one reaches the important conclusion that the mean energy of an ideal gas depends only on its temperature and is independent of its volume. This result is physically plausible. An increase in volume of the container increases the mean distance between the molecules and thus changes, in general, their mean potential energy of mutual interaction. But in the case of an *ideal* gas this interaction energy is negligibly small, while the kinetic and internal energies of the molecules do not depend on the distances between them. Hence the total energy of the gas remains unchanged.

SUGGESTIONS FOR SUPPLEMENTARY READING

The following books give a statistical discussion somewhat similar to the one of this text:

- C. Kittel: "Elementary Statistical Physics," secs. 1-10, John Wiley & Sons, Inc., New York, 1958.
- R. Becker: "Theorie der Wärme," secs. 32-35, 45, Springer-Verlag, Berlin, 1955. (A good book, but in German.)
- L. Landau and E. M. Lifshitz: "Statistical Physics," secs. 1-13, Addison-Wesley, Reading, Mass., 1963.

The following books are good introductions to classical thermodynamics from a completely macroscopic point of view:

- M. W. Zemansky: "Heat and Thermodynamics," 4th ed, McGraw-Hill Book Company, New York, 1957.
- E. Fermi: "Thermodynamics," Dover Publications, New York, 1957.
- H. B. Callen: "Thermodynamics," John Wiley & Sons, Inc., New York, 1960. (More sophisticated in approach than the preceding books.)

PROBLEMS

- 3.1** A box is separated by a partition which divides its volume in the ratio 3:1. The larger portion of the box contains 1000 molecules of Ne gas; the smaller, 100 molecules of He gas. A small hole is punctured in the partition, and one waits until equilibrium is attained.

(a) Find the mean number of molecules of each type on either side of the partition.

(b) What is the probability of finding 1000 molecules of Ne gas in the larger portion and 100 molecules of He gas in the smaller (i.e., the same distribution as in the initial system)?

- 3.2** Consider a system of N localized weakly interacting particles, each of spin $\frac{1}{2}$ and magnetic moment μ , located in an external magnetic field H . This system was already discussed in Problem 2.4.
- Using the expression for $\ln \Omega(E)$ calculated in Problem 2.4b and the definition $\beta = \partial \ln \Omega / \partial E$, find the relation between the absolute temperature T and the total energy E of this system.
 - Under what circumstances is T negative?
 - The total magnetic moment M of this system is related to its energy E . Use the result of part (a) to find M as a function of H and the absolute temperature T .
- 3.3** Consider two spin systems A and A' placed in an external field H . System A consists of N weakly interacting localized particles of spin $\frac{1}{2}$ and magnetic moment μ . Similarly, system A' consists of N' weakly interacting localized particles of spin $\frac{1}{2}$ and magnetic moment μ' . The two systems are initially isolated with respective total energies $bN\mu H$ and $b'N'\mu'H$. They are then placed in thermal contact with each other. Suppose that $|b| \ll 1$ and $|b'| \ll 1$ so that the simple expressions of Problem 2.4c can be used for the densities of states of the two systems.
- In the most probable situation corresponding to the final thermal equilibrium, how is the energy \bar{E} of system A related to the energy \bar{E}' of system A' ?
 - What is the value of the energy \bar{E} of system A ?
 - What is the heat Q absorbed by system A in going from the initial situation to the final situation when it is in equilibrium with A' ?
 - What is the probability $P(E) dE$ that A has its final energy in the range between E and $E + dE$?
 - What is the dispersion $(\Delta^*E)^2 \equiv \overline{(E - \bar{E})^2}$ of the energy E of system A in the final equilibrium situation?
 - What is the value of the relative energy spread $|\Delta^*E/\bar{E}|$ in the case when $N' \gg N$?
- 3.4** Suppose that a system A is placed into thermal contact with a heat reservoir A' which is at an absolute temperature T' and that A absorbs an amount of heat Q in this process. Show that the entropy increase ΔS of A in this process satisfies the inequality $\Delta S \geq Q/T'$, where the equals sign is only valid if the initial temperature of A differs infinitesimally from the temperature T' of A' .
- 3.5** A system consists of N_1 molecules of type 1 and N_2 molecules of type 2 confined within a box of volume V . The molecules are supposed to interact very weakly so that they constitute an ideal gas mixture.
- How does the total number of states $\Omega(E)$ in the range between E and $E + \delta E$ depend on the volume V of this system? You may treat the problem classically.
 - Use this result to find the equation of state of this system, i.e., to find its mean pressure \bar{p} as a function of V and T .
- 3.6** A glass bulb contains air at room temperature and at a pressure of 1 atmosphere. It is placed in a chamber filled with helium gas at 1 atmosphere and at room temperature. A few months later, the experimenter happens to read in a journal article that the particular glass of which the bulb is made is quite permeable to helium, although not to any other gases. Assuming that equilibrium has been attained by this time, what gas pressure will the experimenter measure inside the bulb when he goes back to check?

Macroscopic parameters and their measurement

SECTION 3·11 CONTAINS all the results necessary for an extensive discussion of systems in equilibrium. We shall begin this discussion by exploring a few of the purely *macroscopic* consequences of the theory. The present chapter will consider briefly some of the parameters which are commonly used in the description of macroscopic systems. Many of these parameters, such as heat, absolute temperature, and entropy, have already been introduced. They have been defined in terms of the *microscopic* mechanical concepts applicable to the particles of a system, and their properties and interrelations have already been established on the basis of the microscopic theory. But we have yet to examine how these quantities are to be determined operationally by suitable macroscopic measurements on a system. An examination of this kind of question is, of course, essential to any physical theory, since one must show how theoretical constructs and predictions can be compared with well-defined experimental measurements. In this chapter we shall discuss how the theory suggests *what* quantities are to be studied experimentally and *how* they are to be measured. In Chapter 5 we then shall show how the theory is capable of predicting various important relationships between such measurable macroscopic quantities.

4 · 1 Work and internal energy

The macroscopic work done by a system is very easily determined, since one can readily measure the external parameters of the system and the associated mean generalized forces. For example, if the volume of a system is changed quasi-statically from V_i to V_f , and throughout this process the mean pressure of the system has the measurable value $\bar{p}(V)$, the macroscopic work done by the system is given by calculating the integral (2·10·4)

$$W = \int_{V_i}^{V_f} \bar{p}(V) dV \quad (4 \cdot 1 \cdot 1)$$

The determination of the internal energy \bar{E} of a system is, by (3·11·2), reducible to a measurement of macroscopic work. If one considers a system which is *thermally insulated* so that it cannot absorb any heat, then $Q = 0$ and one has simply

$$\Delta\bar{E} = -W \quad \text{or} \quad \bar{E}_b - \bar{E}_a = -W_{ab} = - \int_a^b dW \quad (4\cdot1\cdot2)$$

This relation defines internal energy *differences* in terms of the macroscopic work W_{ab} done by the system in going from macrostate a to macrostate b . Only such energy differences are of physical significance; i.e., the mean energy is defined only to within an arbitrary additive constant (just as potential energy in mechanics is defined only to within an arbitrary constant). Thus one can choose one particular macrostate a of a system as a standard state from which the mean energy is measured. For example, one can adopt the convention of putting $\bar{E}_a = 0$. To determine the internal energy \bar{E}_b of any other macrostate b of the system, it is only necessary to insulate the system thermally and to go from state a to state b (or conversely, from state b to state a) by the performance of a suitable amount of macroscopic work. Equation (4·1·2) shows that the work thus done is independent of the particular process used in going from a to b . (This was the essential content of the first law of thermodynamics discussed in Sec. 2·11.) Hence the work thus done by the system is guaranteed to yield a unique measurable number. The internal energy of macrostate b is thus uniquely determined by (4·1·2) as

$$\bar{E}_b = -W_{ab} = W_{ba} \quad (4\cdot1\cdot3)$$

The above procedure can be applied to all macrostates b of the system and allows one, therefore, to characterize each such state by a definite operationally measurable value of the internal energy parameter \bar{E}_b . Note that the units of energy are the same as those of work, i.e., ergs or joules.

Example 1 Consider a system consisting of a vessel which contains a liquid and a paddle wheel which is free to rotate (see Fig. 2·7·5). If the system is kept at fixed pressure, its macrostate is completely specified by its internal energy \bar{E} . Equivalently, it can be specified by its temperature ϑ measured with respect to any *arbitrary* thermometer, since \bar{E} and ϑ are functionally related. A falling weight can do macroscopic work on the system by rotating the paddle wheel.

Consider some standard macrostate a , where $\vartheta = \vartheta_a$ and $\bar{E} = \bar{E}_a$. By doing some measurable amount of work W on the system one can attain a different macrostate characterized by a different temperature ϑ and a larger internal energy $\bar{E} = E_a + W$. Similarly, one can determine the internal energy of a macrostate of internal energy lower than \bar{E}_a by starting from this macrostate, characterized by some temperature ϑ , and measuring the amount of work W which must be done on the system to bring its temperature to ϑ_a .

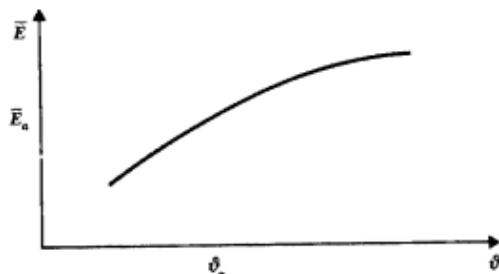


Fig. 4·1·1 Schematic curve showing the dependence of the measured mean energy \bar{E} on the arbitrary temperature parameter ϑ characterizing the system of Fig. 2·7·5.

The energy of the initial macrostate is then $\bar{E}_a - W$. In this way one can construct a curve of \bar{E} versus ϑ (see Fig. 4·1·1). The energy \bar{E}_a of the standard state can, of course, be set equal to zero.

Example 2 Consider a system consisting of an electric resistor (e.g., a coil of platinum wire). If this system is kept at fixed pressure, its macrostate can again be specified completely by an arbitrary temperature parameter ϑ . Here one can determine values of the internal energy of this system corresponding to various values of ϑ by connecting a battery to the resistor and doing electrical work on it. Except for the fact that electrical measurements are usually more convenient and accurate than mechanical ones, the analysis of this example is identical to that of the preceding one.

Example 3 Consider the system of Fig. 4·1·2 consisting of a cylinder containing a gas. The macrostate of this system can be specified by two parameters, e.g., its volume V and internal energy \bar{E} . The mean gas pressure \bar{p} is then determined. (Alternatively, one can specify the macrostate by specifying V and \bar{p} as independent variables; the mean energy \bar{E} is then determined.) Consider a standard macrostate a of volume V_a and mean pressure \bar{p}_a , where $\bar{E} = \bar{E}_a$. How would one determine the mean energy \bar{E}_b of any other macrostate b of volume V_b and mean pressure \bar{p}_b ?

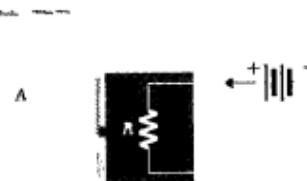


Fig. 4·1·2 A system consisting of a cylinder containing a gas. The volume V of the gas is determined by the position of the movable piston. The resistor R can be brought into thermal contact with this system.

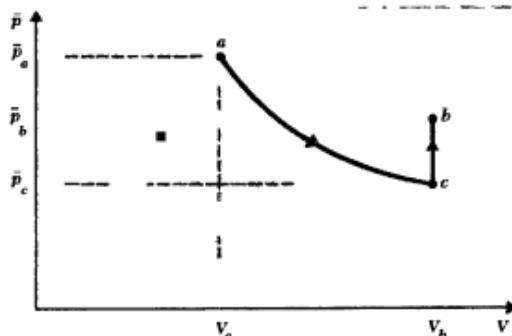


Fig. 4·1·3 Diagram illustrating the volumes and mean pressures describing different macrostates of the gas of Fig. 4·1·2.

Each macrostate can be represented by a point on a \bar{p} V diagram (see Fig. 4·1·3). One could proceed as follows.

1. Let the gas expand against the piston until its volume changes from its initial value V_a to its desired final value V_b . As a result, the mean pressure of the gas decreases to some value \bar{p}_c . Denote by W_{ac} the work done by the gas on the piston in this process.

2. To change the mean pressure to the desired final value \bar{p}_b while keeping the volume constant, bring the system into thermal contact with another system whose internal energy is already known: e.g., bring the system into thermal contact with the electrical resistor of the preceding example. Do electric work W_R on this resistor of an amount just sufficient to bring the gas pressure to \bar{p}_b . In this process the internal energy of the resistor changes by a measurable amount $\Delta\bar{\epsilon}$, and an amount of energy $W - \Delta\bar{\epsilon}$ (in the form of heat) is transferred to the system of interest.

The total internal energy of the system in state b is then given by*

$$\bar{E}_b = \bar{E}_a - W_{ac} + (W - \Delta\bar{\epsilon})$$

4·2 Heat

The heat Q_{ab} absorbed by a system in going from a macrostate a to another macrostate b is, by (2·8·2), defined as

$$Q_{ab} = (\bar{E}_b - \bar{E}_a) + W_{ab} \quad (4·2·1)$$

where W_{ab} is the macroscopic work done by the system in this process. Since we have already discussed how to measure the work W_{ab} and the internal energy

* Since the performance of work on the resistor results in an increase in the pressure of the gas, the above procedure could not be used if $\bar{p}_b < \bar{p}_c$. But then one could proceed in the reverse direction by measuring the work required to go from state b to state a .

\bar{E} , the relation (4·2·1) yields a well-defined number for the heat absorbed. Note that the units of heat are the same as those of work, i.e., ergs or joules.

In practice, two slightly different methods are commonly used for doing "calorimetry," i.e., for measuring the heat absorbed by a system.

Direct measurement in terms of work Suppose it is desired to measure the heat Q_{ab} absorbed by a system A with fixed external parameters (e.g., the gas of Fig. 4·1·2 with piston clamped in position). Then one can bring A into thermal contact with some other system on which work can be done, e.g., an electrical resistor (or equivalently, the paddle-wheel system of Fig. 2·7·5). By doing a measurable amount of electrical work \mathfrak{W} on the resistor, one can bring A from state a to state b . Since the combined system consisting of A and the resistor does not exchange heat with any outside system, Eq. (4·2·1) applied to this combined system yields

$$\mathfrak{W} = \Delta\bar{E} + \Delta\bar{\epsilon}$$

where $\Delta\bar{E}$ is the change in mean energy of A and $\Delta\bar{\epsilon}$ is the change in mean energy of the resistor. Since A itself does no work, the same Eq. (4·2·1) applied to A implies that

$$Q_{ab} = \Delta\bar{E}$$

Hence

$$Q_{ab} = \mathfrak{W} - \Delta\bar{\epsilon} \quad (4\cdot2\cdot2)$$

If the resistor system is sufficiently small compared to A , then $\Delta\bar{\epsilon} \ll \Delta\bar{E}$ and the term $\Delta\bar{\epsilon}$ in (4·2·2) is negligibly small. Otherwise, the internal energy $\bar{\epsilon}$ of the resistor can be considered known as a function of its macrostate (e.g. of its temperature ϑ) from prior measurements involving the performance of electrical work on the isolated resistor system. Equation (4·2·2) thus determines the heat absorbed by A in going from macrostate a to b .

Comparison method (sometimes called the "method of mixtures") While keeping all external parameters fixed, bring the system A into thermal contact with a reference system B whose internal energy is already known as a function of its parameters. No work is done in this process. The conservation of energy for the isolated combined system implies then that in the process of going [from the initial situation, where the systems are in equilibrium and isolated from each other] to [the final situation, where the systems are in equilibrium and in thermal contact], the changes in internal energies satisfy the condition

$$\Delta\bar{E}_A + \Delta\bar{E}_B = 0$$

In terms of the heats absorbed by the two systems, this can be written

$$Q_A + Q_B = 0 \quad (4\cdot2\cdot3)$$

Since $Q_B = \Delta\bar{E}_B$ is known for the reference system B in terms of the changes of its parameters in the process, one has thus measured

$$Q_A = -Q_B \quad (4\cdot2\cdot4)$$

A familiar example of this method is that where water is used as the reference

system B . The system A is immersed in the water. The resulting temperature change of the water can be measured and determines the internal energy change of the water. The heat absorbed by A is then known.

4 · 3 Absolute temperature

We discussed in Sec. 3 · 5 the measurement of temperature with respect to an arbitrary thermometric parameter of some arbitrary thermometer. We now want to consider the operational determination of the *absolute temperature* T of a system. Compared to an arbitrary temperature parameter, the absolute temperature has the following two important properties:

1. As discussed in Sec. 3 · 5, the absolute temperature provides one with a temperature parameter which is completely independent of the nature of the particular thermometer used to perform the temperature measurement.
2. The absolute temperature T is a parameter of fundamental significance which enters all the theoretical equations. Hence all theoretical predictions will involve this particular temperature parameter.

Any theoretical relation involving the absolute temperature T can be used as the basis of an experimental determination of T . We can distinguish between two classes of relations on which such a determination can be based.

a. Theoretical relations involving *microscopic* aspects of the theory. For example, one can apply statistical mechanics to a particular system to calculate from microscopic considerations the equation of state of this system. The equation of state (discussed in Sec. 3 · 12) is a relation between macroscopic parameters of the system and the absolute temperature T . Hence it can be used as a basis for measuring T .

b. Theoretical relations based on the purely *macroscopic* statements of the theory. For example, the second law states that $dS = dQ/T$ for an infinitesimal quasi-static process. This relation involves the absolute temperature T and can thus be used for measuring T . (An example of this procedure will be discussed in Sec. 11 · 3.)

The simplest and most important illustration of the first method is that based on the equation of state for an ideal gas. In (3 · 12 · 8) we found that this equation of state can be written in the form

$$\bar{p}V = NkT \quad (4 \cdot 3 \cdot 1)$$

or equivalently,
where

$$\bar{p}V = \nu RT \quad (4 \cdot 3 \cdot 2)$$

$$R \equiv N_a k \quad (4 \cdot 3 \cdot 3)$$

Here ν is the number of moles of gas, and N_a is Avogadro's number. In practice, the ideal gas conditions of negligible interaction between molecules can be achieved by working with gases in the limit of very high dilution so that the mean intermolecular separation is large.

- In the limit of sufficiently high dilution the mean intermolecular separation also becomes large compared to the mean De Broglie wavelength corresponding to the mean momentum of a gas molecule. Quantum mechanical effects thus become unimportant in this limit and the equation of state (4·3·1), derived on the basis of classical statistical mechanics, must also be valid. (The strictly quantum-mechanical derivation of (4·3·1) will be given in Chapter 9.)

The equation of state (4·3·1) makes some definite predictions. For example, it asserts that, if the temperature is kept fixed, one has the relation

$$\bar{p}V = \text{constant}$$

This result is the familiar and historically important "Boyle's law." Another consequence of (4·3·1) is that this equation of state is the same *irrespective* of the particular gas considered; e.g., it applies equally well to helium, hydrogen, or methane, as long as these gases are sufficiently dilute that they can be considered ideal.

To use the equation of state (4·3·1) as a means for determining the absolute temperature T , one can proceed as follows. Keep the volume V of the given amount of gas fixed. Then one has a constant-volume gas thermometer of the type described in Sec. 3·5. Its thermometric parameter is the pressure \bar{p} . By virtue of (4·3·1) one knows that \bar{p} is directly proportional to the absolute temperature T of the gas. Indeed, once one chooses the value of the arbitrary constant k still at one's disposal (or equivalently, the value of the constant R), Eq. (4·3·1) determines a definite value of T . We now describe how this constant k is conventionally chosen, or equivalently, how the absolute temperature *scale* is chosen.

When the constant-volume gas thermometer is brought into thermal contact with some system A with which it is allowed to come to equilibrium, its pressure will attain some definite value \bar{p}_A . When it is brought into thermal contact with some other system B (this may be a system of the same kind but in some other macrostate) with which it is allowed to come to equilibrium, its pressure will attain some other definite value \bar{p}_B . By (4·3·1), the pressure ratio is then given by

$$\frac{\bar{p}_A}{\bar{p}_B} = \frac{T_A}{T_B} \quad (4·3·4)$$

where T_A and T_B are the absolute temperatures of the systems A and B . Thus any ideal gas thermometer can be used to measure absolute temperature ratios. In particular, if system B is chosen as some standard system in some standard macrostate, then the gas thermometer can be used to measure the ratio of the absolute temperature T of any system to the temperature T_B of this standard system.

The relation (4·3·4) is another consequence of the equation of state (4·3·1) which can be checked experimentally, for it implies that this pressure ratio ought to be the same no matter what gas is used in the gas thermometer, provided that the gas is sufficiently dilute. In symbols,

$$\lim_{v \rightarrow 0} \frac{\bar{p}_A}{\bar{p}_B} \rightarrow \text{constant independent of nature of gas} \quad (4\cdot3\cdot5)$$

where v denotes the number of moles of gas used in the bulb of the thermometer.

This comment also provides an experimental criterion for deciding when a gas is sufficiently dilute to be considered ideal. The pressure ratio \bar{p}_A/\bar{p}_B can be measured with a given amount of gas in the bulb. This measurement can then be repeated with successively smaller amounts of gas in the bulb; the ratio \bar{p}_A/\bar{p}_B must then, by (4·3·5), reach a limiting constant value. When this is the case, the gas is known to be sufficiently dilute to exhibit ideal behavior.

By international convention one chooses as the standard system pure water, and as the standard macrostate of this system the situation where the solid, liquid, and gas phases of this system (i.e., ice, water, and water vapor) are in equilibrium with each other. (This macrostate is called the "triple

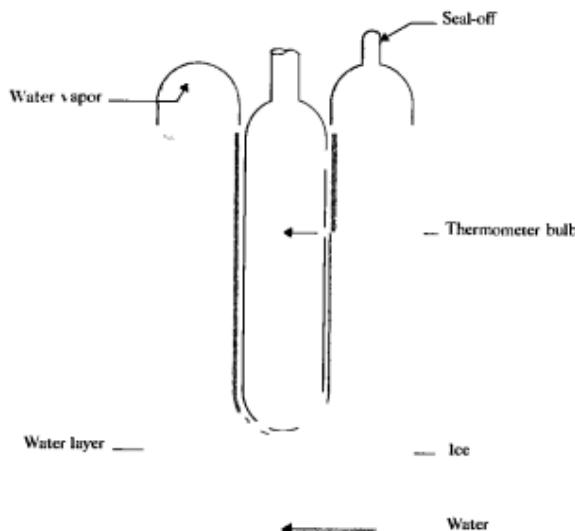


Fig. 4·3·1 Diagram illustrating a triple-point cell designed to calibrate a thermometer at the triple point of water. A freezing mixture is first introduced into the central well to produce some of the ice. After the freezing mixture is removed, the thermometer bulb is placed in the well and the system is allowed to come to thermal equilibrium.

point" of water.) The reason for this choice is that there is only one definite value of pressure and temperature at which all these three phases can coexist in equilibrium and, as can readily be verified experimentally, the temperature of this system is unaffected by any changes in the relative amounts of solid, liquid, and gas present under these circumstances. The triple point provides, therefore, a very reproducible standard of temperature. By international convention one chooses to assign to the absolute temperature T_t of water at its triple point the value

$$T_t \equiv 273.16 \quad \text{exactly} \quad (4 \cdot 3 \cdot 6)$$

This particular choice was motivated by the desire to keep the modern temperature scale, adopted by international convention in 1954, as nearly identical as possible with a historically older temperature scale.

The choice (4 · 3 · 6) fixes a scale factor for T which we can indicate by assigning to T the unit* of "degree K." This stands for "degree Kelvin" (so named after the famous British physicist of the last century), and is commonly abbreviated as "°K." Whenever we use the unit "degree" without further qualifications, we shall always mean "degree K." We mentioned in Sec. 3 · 3 that kT has the dimensions of energy. It thus follows that the constant k has the units of ergs/degree.

Once the above conventions have been adopted, the absolute temperature T_A of any system A is completely determinable by the gas thermometer. If the thermometer exhibits pressure \bar{p}_A when in thermal contact with this system, and pressure \bar{p}_t when in thermal contact with water at the triple point, then

$$T_A = 273.16 \frac{\bar{p}_A}{\bar{p}_t} \quad (4 \cdot 3 \cdot 7)$$

Here the pressure ratio is to be evaluated in the ideal gas limit, i.e., in the limit when the gas used in the thermometer is made sufficiently dilute. Thus the absolute temperature of any system can be directly determined by measuring the pressure of a constant volume gas thermometer. This is in practice a relatively simple method of measuring the absolute temperature, provided that the temperature is not so low or so high that the use of gas thermometers becomes impracticable.

Once the temperature scale has been fixed by (4 · 3 · 6), one can return to the equation of state (4 · 3 · 2) and determine the constant R . Taking ν moles of any gas at the triple-point temperature $T_t = 273.16^\circ\text{K}$, one need only measure its volume V (in cm^3) and its corresponding mean pressure \bar{p} (in dynes/ cm^2); this information permits computation of R by (4 · 3 · 2). Careful measurements of this type yield for the gas constant the value

$$R = (8.3143 \pm 0.0012) \text{ joules mole}^{-1} \text{ deg}^{-1} \quad (4 \cdot 3 \cdot 8)$$

* This unit is a unit in the same sense as the degree of angular measure; it does not involve length, mass, or time.

(1 joule = 10^7 ergs). Knowing Avogadro's number ("unified scale," atomic weight of C¹² = 12 exactly)

$$N_a = (6.02252 \pm 0.00028) \times 10^{23} \text{ molecules mole}^{-1} \quad (4 \cdot 3 \cdot 9)$$

one can use (4·3·3) to find the value of k . This important constant is called "Boltzmann's constant" in honor of the Austrian physicist who contributed so significantly to the development of kinetic theory and statistic mechanics. Its value is found to be*

$$k = (1.38054 \pm 0.00018) \times 10^{-16} \text{ ergs degree}^{-1} \quad (4 \cdot 3 \cdot 10)$$

On the absolute-temperature scale defined by the choice of (4·3·6), an energy of 1 ev (electron volt) corresponds to an energy kT , where $T \approx 11\,600^\circ\text{K}$. Also, room temperature is approximately 300°K on this scale and corresponds to an energy $kT \approx \frac{1}{40}$ ev.

Another temperature scale sometimes used is the Celsius (or centigrade) temperature θ defined in terms of the absolute temperature T by the relation

$$\theta = T - 273.15 \quad \text{degrees Celsius} \quad (4 \cdot 3 \cdot 11)$$

(abbreviated as " ${}^\circ\text{C}$ "). On this scale water at atmospheric pressure freezes at *approximately* 0°C and boils at *approximately* 100°C .

Historical remark We wish to mention briefly the historical context which motivates the particular choice of numerical value in (4·3·6). The basic reason was the adoption, before the full significance of the absolute temperature concept had become clear, of the Celsius (or centigrade) temperature scale based on *two* fixed standard temperatures. In this scheme the Celsius temperature θ was chosen as a linear function of the thermometric parameter. For example, in the case of the constant-volume gas thermometer where the thermometric parameter is the pressure \bar{p} , the Celsius temperature was taken to be

$$\theta = a\bar{p} + b \quad (4 \cdot 3 \cdot 12)$$

where a and b are constants to be determined in terms of the two fixed standard temperatures. The latter were again based on water as the standard system, but the two standard macrostates of this substance were chosen as follows:

1. The state where ice is in equilibrium with air-saturated water at atmospheric pressure. This is the so-called "ice point" of water. By definition one assigned to the temperature of this state the value $\theta = 0$.
2. The state where water is in equilibrium with water vapor at a pressure of 1 atmosphere. This is the so-called "steam point" of water. By definition one assigned to the temperature of this state the value $\theta = 100$.

* The values of these physical constants are those of the least-squares adjustment of E. R. Cohen and J. W. M. DuMond and approved for adoption by the National Research Council in April, 1963. See the Table of Numerical Constants at the end of the book.

(We note parenthetically that these points are more difficult to reproduce experimentally than the triple point of water, which involves only pure water (without the presence of air) and needs no specification of the applied pressure.)

If the pressure readings of the gas thermometer at the ice and steam points are denoted by \bar{p}_i and \bar{p}_s , respectively, then Eq. (4·3·12) yields the following two relations when applied to these two standard states:

$$\begin{aligned} 0 &= a\bar{p}_i + b \\ 100 &= a\bar{p}_s + b \end{aligned}$$

These two equations can be solved for a and b in terms of \bar{p}_i and \bar{p}_s . The relation (4·3·12) then becomes

$$\theta = 100 \frac{\bar{p} - \bar{p}_i}{\bar{p}_s - \bar{p}_i} \quad (4·3·13)$$

Alternatively, this can be used to express \bar{p} in terms of θ . One finds

$$\bar{p} = \bar{p}_i \left(1 + \frac{\theta}{\theta_0} \right) \quad (4·3·14)$$

where $\theta_0 = 100 \left(\frac{\bar{p}_s}{\bar{p}_i} - 1 \right)^{-1}$ (4·3·15)

depends only on a pressure ratio and is thus independent of the nature of the gas used. Thus θ_0 is a universal constant for all gases and can be *measured* by using a gas thermometer at the ice and steam points. One finds thus

$$\theta_0 = 273.15 \quad (4·3·16)$$

By (4·3·14), measurements of two systems at the respective temperatures θ_A and θ_B yield for the corresponding pressure ratio

$$\frac{\bar{p}_A}{\bar{p}_B} = \frac{\theta_0 + \theta_A}{\theta_0 + \theta_B} \quad (4·3·17)$$

This is of the same form as (4·3·4) if one defines the absolute temperature T by the relation

$$T = \theta_0 + \theta \quad (4·3·18)$$

If the triple point of water is measured on this temperature scale, one finds $\theta \approx 0.01^\circ\text{C}$, or by (4·3·18), $T \approx 273.16^\circ\text{K}$ approximately.

It is clear that this old-fashioned procedure for establishing a temperature scale is cumbersome, logically not very satisfying, and not of the highest possible accuracy. The modern convention using a single fixed point is far more satisfactory in all these respects. But by choosing T_i to be exactly 273.16°K rather than some other number, one gains the convenience that all the older temperature measurements based on the former temperature scale will (within the limits of accuracy with which the triple point of water was measured on that scale) agree numerically with the values based on the modern convention.

4 · 4 Heat capacity and specific heat

Consider a macroscopic system whose macrostate can be specified by its absolute temperature T and some other macroscopic parameter (or set of macroscopic parameters) y . For example, y might be the volume or the mean pressure of the system. Suppose that, starting with the system at temperature T , an infinitesimal amount of heat dQ is added to the system while all its parameters y are kept fixed. The resulting change dT in the temperature of the system depends on the nature of the system as well as on the parameters T and y specifying the macrostate of the system. We define the ratio

$$\left(\frac{dQ}{dT} \right)_y \equiv C_y \quad (4 \cdot 4 \cdot 1)$$

in the limit as $dQ \rightarrow 0$ (or $dT \rightarrow 0$) as the “heat capacity” of the system. Here we have used the subscript y to denote explicitly the parameters kept constant in the process of adding heat. The quantity C_y depends, of course, on the nature of the system and on the particular macrostate under consideration, i.e., in general

$$C_y = C_y(T, y) \quad (4 \cdot 2 \cdot 2)$$

The amount of heat dQ which needs to be added to produce a given temperature change dT of a homogeneous system will be proportional to the quantity of matter contained therein. Hence it is convenient to define a quantity, the “specific heat,” which depends only on the nature of the substance under consideration, not on the amount present. This can be achieved by dividing the heat capacity C_y of ν moles (or m grams) of the substance by the corresponding number of moles (or of grams). The “specific heat per mole” or “heat capacity per mole” is thus defined as

$$c_y \equiv \frac{1}{\nu} C_y = \frac{1}{\nu} \left(\frac{dQ}{dT} \right)_y \quad (4 \cdot 4 \cdot 3)$$

Equivalently, the “specific heat per gram” is defined as

$$c_y' \equiv \frac{1}{m} C_y = \frac{1}{m} \left(\frac{dQ}{dT} \right)_y \quad (4 \cdot 4 \cdot 4)$$

The cgs units of the molar specific heat are, by (4 · 4 · 3), ergs degree⁻¹ mole⁻¹.

It should be noted from the operational definition (4 · 4 · 1) of the heat capacity C_y that this quantity does depend on which particular parameters y of the system are kept constant in the process of adding the heat. Suppose we consider a substance, e.g., a gas or liquid, whose macrostate can be specified by two parameters, say the temperature T and the volume V (see Fig. 4 · 4 · 1). When the system is in a given macrostate, we can ask for the following two quantities: (1) c_V , the molar specific heat at constant volume of the system in this state, and (2) c_p , the molar specific heat at constant pressure of the system in this state.



Fig. 4·4·1 Diagram illustrating specific heat measurements of a gas kept at constant volume or at constant pressure.

1. To determine c_v , we clamp the piston in position so that the volume of the system is kept fixed. In this case the system cannot do any work, and the heat dQ added to the system goes entirely to increase the internal energy of the system

$$dQ = d\bar{E} \quad (4 \cdot 4 \cdot 5)$$

2. To determine c_p , the piston is left completely free to move, the weight of the piston being equal to the constant force per unit area (or mean pressure \bar{p}) on the system. In this case the piston will move when heat dQ is added to the system; as a result, the system does also mechanical work. Thus the heat dQ is used *both* to increase the internal energy of the system *and* to do mechanical work on the piston; i.e.

$$dQ = d\bar{E} + \bar{p} dV \quad (4 \cdot 4 \cdot 6)$$

For a *given* amount of heat absorbed, the internal energy \bar{E} will therefore increase by a smaller amount (and hence the temperature T will also increase by a smaller amount) in the second case compared to the first. By (4·4·1) one expects, therefore, that the heat capacity is greater in the second case; i.e., one expects that

$$c_p > c_v \quad (4 \cdot 4 \cdot 7)$$

Remark Note that the specific heat at constant volume may itself still be a function of the volume V ; i.e., $c_v = c_v(T, V)$ in general. For example, the heat required to raise the temperature of a gas from 300 to 301 degrees is in general not the same if its volume is kept constant at 50 cm^3 in the process of adding heat, as it is if its volume is kept constant at 1000 cm^3 in the process of adding heat.

Since the second law allows us to write $dQ = T dS$, the heat capacity (4·4·1) can be written in terms of the entropy as

$$C_v = T \left(\frac{\partial S}{\partial T} \right)_v \quad (4 \cdot 4 \cdot 8)$$

If S in this expression is the entropy per mole of substance, then C_v is the molar specific heat.

If one contemplates a situation where all the *external* parameters of a

system are kept fixed, then the system does no macroscopic work, $dW = 0$, and the first law reduces simply to the statement $dQ = d\bar{E}$. For example, if the volume V is the only external parameter, one can write

$$C_V = T \left(\frac{\partial S}{\partial T} \right)_V = \left(\frac{\partial \bar{E}}{\partial T} \right)_V \quad (4 \cdot 4 \cdot 9)$$

By virtue of (3.7.16) it follows that this quantity is always positive.

Measurements of the specific heat involve measurements of heat of the type discussed in Sec. 4.2. In measuring heat by the comparison method (or method of mixtures), it used to be popular to select water as the reference substance. Hence a knowledge of its specific heat was of particular importance. Such a measurement can be made directly in terms of work and was first performed by Joule during the years 1843–1849. The specific heat of water at a pressure of 1 atmosphere and a temperature of 15°C (288.2°K) is found to be 4.18 joules deg^{-1} gram^{-1} .

Before the nature of heat as a form of energy was understood, it was customary to define a unit of heat, the "calorie," as the heat required to raise the temperature of water at 1 atmosphere from 14.5 to 15.5°C. Joule's measurement of the heat capacity of water in terms of work allowed expression of the calorie in terms of absolute energy units. The calorie unit is gradually becoming obsolete and is now *defined* in terms of energy units by the relation

$$1 \text{ calorie} \equiv 4.1840 \text{ joules} \quad (4 \cdot 4 \cdot 10)$$

Example Let us consider heat measurements by the method of mixtures in terms of the specific heats of the substances involved. Consider that two substances A and B , of respective masses m_A and m_B , are brought into thermal contact under conditions where the pressure is kept constant. (For example, a copper block is immersed in water, the pressure being atmospheric.) Suppose that at this pressure the specific heats per gram of the respective substances are $c_A'(T)$ and $c_B'(T)$. Assume that before the substances are brought into thermal contact their respective equilibrium temperatures are T_A and T_B , respectively. Denote their final common temperature, after equilibrium is reached, by T_f . No work gets done in this process, so that the conservation of energy is expressed by (4.2.3) as

$$Q_A + Q_B = 0 \quad (4 \cdot 4 \cdot 11)$$

But by (4.4.4) the heat absorbed by a substance when its temperature is increased by an amount dT is given by $dQ = mc'dT$. Hence the heat absorbed by A in going from temperature T_A to temperature T_f is given by

$$Q_A = \int_{T_A}^{T_f} m_A c_A'(T') dT'$$

or
$$Q_A = m_A c_A'(T_f - T_A)$$

If the temperature dependence of c_A' is negligible. Similar expressions hold

for B . Hence the fundamental condition (4·4·11) can be written

$$m_A \int_{T_A}^{T_f} c_A' dT' + m_B \int_{T_B}^{T_f} c_B' dT' = 0 \quad (4\cdot4\cdot12)$$

This relation allows, for example, computation of the final temperature T_f if the other quantities are known. The situation is particularly simple if the specific heats c_A' and c_B' are temperature independent. In this case (4·4·12) becomes simply

$$m_A c_A' (T_f - T_A) + m_B c_B' (T_f - T_B) = 0 \quad (4\cdot4\cdot13)$$

If desired, this can be solved explicitly for the final temperature T_f to give

$$T_f = \frac{m_A c_A' T_A + m_B c_B' T_B}{m_A c_A' + m_B c_B'}$$

4 · 5 Entropy

The entropy can readily be determined by using the second-law statement (3·11·4) that $dS = dQ/T$ for an infinitesimal quasi-static process. Given any macrostate b of the system, one can find the entropy difference between this state and some standard state a by considering *any quasi-static* process which carries the system from state a to state b and calculating for this process the integral

$$S_b - S_a = \int_a^b \frac{dQ}{T} \quad (4\cdot5\cdot1)$$

The evaluation of this integral yields, as discussed in Sec. 3·10, the same unique value of $S_b - S_a$ *irrespective* of what quasi-static process may be chosen to bring the system from state a to state b .

Let us emphasize again that the process chosen for calculating the integral in (4·5·1) *must* be quasi-static. This means that we must somehow bring the system from state a to state b by continuously changing its parameters so slowly (compared to significant relaxation times) that it is at all times very close to an equilibrium situation. In general, this will require the use of other auxiliary systems on which the system can do work and from which it can absorb heat. For example, if we need to change the volume of the system, we could do it by successively moving a piston by small amounts, proceeding sufficiently slowly to allow the system to reach equilibrium at all stages. Or, if we need to change the temperature of the system, we could do it by bringing the system successively into contact with a large number of heat reservoirs of slightly different temperatures, again proceeding sufficiently slowly to allow

the system to reach equilibrium at all stages. Clearly, the temperature T is a well-defined quantity as we go through this succession of equilibrium situations, and the heat dQ absorbed in going from one macrostate to an adjacent one is also a measurable quantity. The evaluation of the entropy difference by (4·5·1) thus presents no conceptual difficulties. Note that the units of entropy are, by (4·5·1), ergs/degree or joules/degree.

Suppose that the macrostate of a body is only specified by its temperature, since all its other parameters y (e.g., its volume V or its mean pressure \bar{p}) are kept constant. If one knows the heat capacity $C_y(t)$ of the body under these conditions, then its entropy difference (for the given values of the parameters y) is given by

$$S(T_b) - S(T_a) = \int_a^b \frac{dQ}{T} = \int_{T_a}^{T_b} \frac{C_y(T') dT'}{T'} \quad (4\cdot5\cdot2)$$

In the particular case that $C_y(T)$ is independent of T , this becomes simply

$$S(T_b) - S(T_a) = C_y \ln \frac{T_b}{T_a} \quad (4\cdot5\cdot3)$$

Example Consider the example discussed at the end of Sec. 4·4 where two systems A and B , with constant specific heats c_A' and c_B' and originally at respective temperatures T_A and T_B , are brought into thermal contact with each other. After the systems come to equilibrium, they reach a common final temperature T_f . What is the entropy change of the entire system in this process? The process which occurred here was certainly *not* a quasi-static one (unless $T_B = T_A$); nonequilibrium conditions prevailed between the initial and final situations. To calculate the entropy change of system A , we can imagine that it is brought from its initial temperature T_A to its final temperature T_f by a succession of infinitesimal heat additions whereby the system at any intermediate equilibrium temperature T absorbs from a heat reservoir at an infinitesimally higher temperature ($T + dT$) a small amount of heat $dQ = m_A c_A' dT$. Thus the entropy change of A is given by

$$\Delta S_A = S_A(T_f) - S_A(T_A) = \int_{T_A}^{T_f} \frac{m_A c_A' dT}{T} = m_A c_A' \ln \frac{T_f}{T_A}$$

A similar expression holds for system B . Hence the entropy change of the total system is given by

$$\Delta S_A + \Delta S_B = m_A c_A' \ln \frac{T_f}{T_A} + m_B c_B' \ln \frac{T_f}{T_B} \quad (4\cdot5\cdot4)$$

Since this represents the total entropy change of the isolated system ($A + B$), we know by the second law (3·11·3) that this can never be negative. To verify this explicitly, we make use of the simple inequality (proved in Appendix A·8)

$$\ln x \leq x - 1 \quad (= \text{sign for } x = 1) \quad (4\cdot5\cdot5)$$

Hence $-\ln x \geq -x + 1$

or, putting $y = 1/x$,

$$\ln y \geq 1 - \frac{1}{y} \quad (= \text{sign for } y = 1) \quad (4 \cdot 5 \cdot 6)$$

Equation (4 · 5 · 4) implies, therefore, the inequality

$$\begin{aligned}\Delta S_A + \Delta S_B &\geq m_A c_A' \left(1 - \frac{T_A}{T_f}\right) + m_B c_B' \left(1 - \frac{T_B}{T_f}\right) \\ &= T_f^{-1}[m_A c_A'(T_f - T_A) + m_B c_B'(T_f - T_B)] \\ &= 0 \quad \text{by (4 · 4 · 13)}\end{aligned}$$

Thus $\Delta S_A + \Delta S_B \geq 0$ (4 · 5 · 7)

The equals sign holds only if $T_A/T_f = 1$ and $T_B/T_f = 1$, i.e., if $T_B = T_A$; then equilibrium is indeed preserved and no irreversible process occurs after the systems are brought into thermal contact with each other.

The relation (4 · 5 · 2) is actually quite interesting because it gives an explicit connection between two different types of information about the system under consideration. On the one hand, (4 · 5 · 2) involves the heat capacity $C(T)$ obtainable from macroscopic measurements of absorbed heat. On the other hand, it involves the entropy which is related to a microscopic knowledge of the quantum states of the system and which can be calculated either from first principles or from experimental information obtainable from spectroscopic data.

Example As a simple illustration, suppose that one is dealing with a simple system of N magnetic atoms, each with spin $\frac{1}{2}$. If this system is known to be ferromagnetic at sufficiently low temperatures, all spins must be completely aligned as $T \rightarrow 0$ so that the number of accessible states $\Omega \rightarrow 1$, or $S = k \ln \Omega \rightarrow 0$ (in accordance with the third law). But at sufficiently high temperatures all spins must be completely randomly oriented so that $\Omega = 2^N$ and $S = kN \ln 2$. Hence it follows that this system must have a heat capacity $C(T)$ which satisfies, by (4 · 5 · 2), the equation

$$\int_0^\infty \frac{C(T') dT'}{T'} = kN \ln 2$$

This relation must be valid irrespective of the details of the interactions which bring about ferromagnetic behavior and irrespective of the temperature dependence of $C(T)$.

These comments should make it apparent that the measuring of heat capacities is not just a dull activity undertaken to fill up handbooks with data useful to engineers concerned with the properties of materials. Accurate measurements of heat capacities may be of considerable interest because they

can provide some important information about the nature of the energy levels of physical systems.

4 · 6 *Consequences of the absolute definition of entropy*

In many applications it is true that only entropy *differences*, i.e., values of the entropy measured with respect to some chosen standard state, are of importance. In this respect the entropy is similar to the internal energy \bar{E} of a system, and the entropy of any state with respect to the standard state can be determined by the integral (4 · 5 · 1). But we do know, as discussed in Sec. 3 · 10, that the entropy S is a completely calculable number and is *not* merely defined to within an arbitrary additive constant. This reflects itself in the third law statement that the entropy approaches, as $T \rightarrow 0$, a definite value S_0 (usually $S_0 = 0$) independent of all parameters of the system. To obtain an absolute value of the entropy one can either use statistical mechanics to calculate the absolute value of the entropy in the standard state, or one can measure entropy differences from a standard state chosen at $T \rightarrow 0$ where $S = S_0$ is known to have a definite value independent of all parameters of the system.

The fact that the entropy has a definite value (without any arbitrary additive constant) makes possible physically significant statements which could not be made otherwise. The following two examples will serve to illustrate this point.

Example 1 Consider the case of a solid which can exist in two different crystal structures. A classical example of this kind is tin, which exists in two rather different forms: one of these is "white" tin, which is a metal; the other is "gray" tin, which is a semiconductor. Gray tin is the stable form at temperatures below $T_0 \equiv 292^\circ\text{K}$, while white tin is the stable form above this temperature. At the temperature T_0 the two forms are in equilibrium with each other. They can then coexist indefinitely in arbitrary proportions, and a positive amount of heat Q_0 must be absorbed to transform one mole of gray tin into the white tin modification.

Although white tin is the unstable form at temperatures below T_0 , the speed with which the transformation proceeds to the gray form is very slow compared to times of experimental interest. It is thus very easy to work with white tin, the ordinary metal, down to very low temperatures. (One is, in practice, scarcely aware that there is a tendency for the metal to transform to the gray form.) A sample of white tin readily achieves internal equilibrium, although it exhibits a negligible tendency to transform to the gray form. One can thus easily measure the molar specific heat $C^{(w)}(T)$ of white tin in the temperature range $T < T_0$. There is, of course, no difficulty in working with a sample of gray tin in this temperature range; one can thus also make measurements of the molar specific heat $C^{(g)}(T)$ of gray tin at temperatures $T < T_0$.*

* All quantities in this discussion refer to measurements at the same constant pressure.

Since the transformation from the white to the gray form proceeds at a negligible rate, the situation would not be changed significantly if one imagined imposing a constraint which would prevent the transformation altogether. In that case white tin can be considered in a genuine equilibrium situation statistically distributed over all states consistent with the crystal structure of white tin; similarly, gray tin can be considered in an equilibrium situation statistically distributed over all states consistent with the crystal structure of gray tin. Arguments of equilibrium statistical mechanics can then be applied to each of these systems. In particular, let us consider the limit as the temperature $T \rightarrow 0$. By this limit we shall mean a reasonably low temperature (say 0.1°K), but not one so extremely low (say, less than 10^{-6}°K) that the random spin orientation of the tin nuclei would be affected.* In accordance with the discussion of Sec. 3·10, a sample consisting of a mole of white tin approaches a ground-state configuration consistent with the white-tin crystal structure. Correspondingly its entropy $S^{(w)}$ tends to zero except for the contribution $S_0 = k \ln \Omega_S$ associated with the Ω_S states of possible nuclear spin orientations. Similarly, a sample consisting of a mole of gray tin approaches a ground state configuration consistent with the gray tin crystal structure. Correspondingly, its entropy $S^{(g)}$ also tends to zero except for the contribution due to the possible nuclear spin orientations. Since one is dealing with the *same* number of the *same* kind of nuclei, there are again Ω_S possible spin states, and this nuclear spin contribution to the entropy is again $S_0 = k \ln \Omega_S$. Thus, as $T \rightarrow 0$,

$$\begin{aligned} S^{(w)}(T) &\rightarrow S_0 \quad \text{and} \quad S^{(g)}(T) \rightarrow S_0 \\ \text{i.e.,} \quad S^{(w)}(0) &= S^{(g)}(0) \end{aligned} \quad (4 \cdot 6 \cdot 1)$$

The relation (4·6·1) expresses just the content of the third law that the entropy approaches, as $T \rightarrow 0$, a value *independent* of all parameters of the system (in this case independent of crystal structure). We shall now show how this statement can be combined with a knowledge of the specific heats to calculate the heat of transformation Q_0 from gray to white tin at the transition temperature T_0 . For suppose that it is desired to calculate the entropy $S^{(w)}(T_0)$ of a mole of white tin at $T = T_0$. One can use two different quasi-static processes in going from $T = 0$ to the same final macrostate.

1. Bring a mole of white tin quasi-statically from $T = 0$ to $T = T_0$. This yields a final entropy

$$S^{(w)}(T_0) = S^{(w)}(0) + \int_0^{T_0} \frac{C^{(w)}(T')}{T'} dT' \quad (4 \cdot 6 \cdot 2)$$

2. Take a mole of gray tin at $T = 0$ and bring it first quasi-statically to temperature T_0 . Then transform it quasi-statically (at this equilibrium transition temperature) to white tin; its entropy change in this transformation is

* The magnetic field H produced by a nuclear moment μ at a neighboring nucleus at an interatomic distance r is of the order of μ/r^3 , i.e., of the order of 5 gauss if μ is a nuclear magneton (5×10^{-24} ergs/gauss) and $r = 10^{-8}$ cm. Departures from random nuclear spin orientation can therefore only be expected at a temperature T low enough so that $kT \lesssim \mu H$, the interaction energy between nuclei.

simply Q_0/T_0 . Hence one can write*

$$S^{(o)}(T_0) = S^{(o)}(0) + \int_0^{T_0} \frac{C^{(o)}(T')}{T'} dT' + \frac{Q_0}{T_0} \quad (4 \cdot 6 \cdot 3)$$

Making use of the result (4 · 6 · 1), one thus obtains the relation

$$\frac{Q_0}{T_0} = \int_0^{T_0} \frac{C^{(w)}(T')}{T'} dT' - \int_0^{T_0} \frac{C^{(o)}(T')}{T'} dT' \quad (4 \cdot 6 \cdot 4)$$

By using the experimental specific heat measurements and performing the integrations numerically, one finds that the first integral has the value 51.4 joules/deg, while the second integral has the value 44.1 joules/deg. Using $T_0 = 292^\circ\text{K}$, one then calculates $Q_0 = (292)(7.3) = 2130$ joules, which compares favorably with the value of 2240 joules obtained by direct measurement of the heat of transformation. Note that this calculation of Q_0 would have been impossible if the third law had not allowed the comparison (4 · 6 · 1) between the entropies at $T = 0$.

Example 2 As a second example illustrating the significance of the third law for calculations of the entropy, consider a system *A* consisting of a mole of solid lead (Pb) and a mole of solid sulfur (S) separated by a partition. Consider also another system *B* consisting of a mole of the solid compound lead sulfide (PbS). Although the systems *A* and *B* are very different, they consist of the same atoms. Hence the third law asserts that the entropies of both systems approach the same value as $T \rightarrow 0$, a value corresponding merely to the number of possible orientations of the nuclear spins. In symbols,

$$S^{(\text{Pb+S})}(0) = S^{(\text{PbS})}(0) \quad (4 \cdot 6 \cdot 5)$$

if $S^{(\text{Pb+S})}(T)$ denotes the entropy of system *A* and $S^{(\text{PbS})}$ denotes the entropy of system *B*.

Suppose that the systems *A* and *B* are both at atmospheric pressure. Suppose further that one knows as a function of temperature the heat capacity per mole (at constant atmospheric pressure) $C^{(\text{Pb})}$ of solid lead, $C^{(\text{S})}$ of solid sulfur, and $C^{(\text{PbS})}$ of lead sulfide (PbS). Then one can write for the entropy of the system *A* consisting of Pb and S separately

$$S^{(\text{Pb+S})}(T) = S^{(\text{Pb+S})}(0) + \int_0^T \frac{C^{(\text{Pb})}(T')}{T'} dT' + \int_0^T \frac{C^{(\text{S})}(T')}{T'} dT' \quad (4 \cdot 6 \cdot 6)$$

For the entropy of the system *B* consisting of PbS one can write at the same temperature

$$S^{(\text{PbS})}(T) = S^{(\text{PbS})}(0) + \int_0^T \frac{C^{(\text{PbS})}(T')}{T'} dT' \quad (4 \cdot 6 \cdot 7)$$

* The integrals in (4 · 6 · 2) and (4 · 6 · 3) must converge properly, since all other quantities are finite. Hence the specific heats must approach zero as $T \rightarrow 0$. This is a general property of the specific heats of all substances and one that is well verified experimentally.

By virtue of (4·6·5) the last two relations provide then a quite unique value for the entropy difference [$S^{(Pb+S)}(T) - S^{(PbS)}(T)$], even though the calculations involve only a knowledge of the specific heats and no information whatever about how Pb and S might react to form PbS. This implies also a very definite prediction as to the value of the integral $\int dQ/T$ which one would obtain by carrying out a quasi-static process in which PbS at a temperature T would be transformed through successive equilibrium states into separated Pb and S at the same temperature. Such a quasi-static process could be performed as follows: Heat the PbS slowly until it evaporates, then heat it further until all the PbS molecules are completely dissociated into Pb and S atoms; then separate the gases slowly with the aid of a semipermeable membrane; then lower the temperature of the whole system back to the temperature T while keeping the membrane in position.

Remark on the separation of gases by semipermeable membranes

It is possible to conceive of a membrane which is completely permeable to molecules of one type and completely impermeable to all other molecules. (One can also realize such membranes in practice; e.g., hot palladium metal is permeable to hydrogen (H_2) gas but to no other gases.) With the aid of such membranes one can unmix gases as illustrated in Fig. 4·6·1. For example, to unmix the A and B molecules in a quasi-static way, one needs only to move the two membranes slowly until they meet somewhere in the container. Then all the A molecules will be in the left part and all the B molecules in the right part of the container.

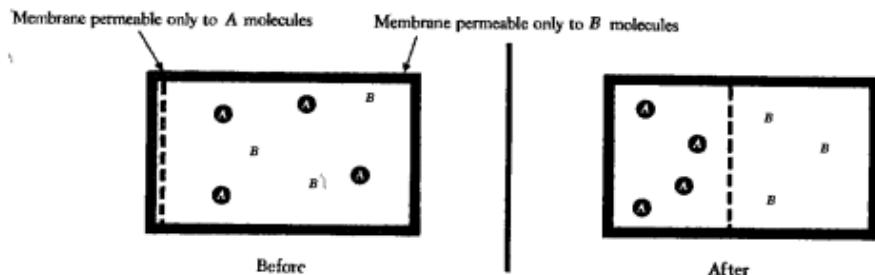


Fig. 4·6·1 Separation of two gases A and B by means of semipermeable membranes.

4 · 7 Extensive and intensive parameters

The macroscopic parameters specifying the macrostate of a homogeneous system can be classified into two types. Let y denote such a parameter. Consider that the system is divided into two parts, say by introducing a partition, and denote by y_1 and y_2 the values of this parameter for the two subsystems.

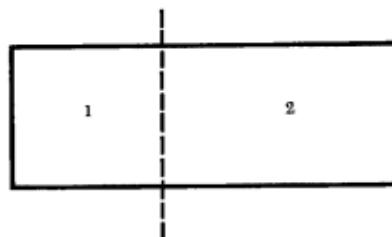


Fig. 4·7·1 Division of a homogeneous system into two parts by a partition.

Then two cases can arise:

1. One has $y_1 + y_2 = y$, in which case the parameter y is said to be *extensive*.
2. One has $y_1 = y_2 = y$, in which case the parameter y is said to be *intensive*.

In simple terms one can say that an extensive parameter gets doubled if the size of the system is doubled, while an intensive parameter remains unchanged.

Thus the volume V of a system is an extensive parameter, as is the total mass M of a system. On the other hand, the density ρ of a system, $\rho = M/V$, is an intensive parameter. Indeed, it is clear that the ratio of any two extensive parameters is an intensive parameter.

The mean pressure of a system is an intensive parameter, since both parts of a system, after subdivision, will have the same pressure as before. Similarly, the temperature T of a system is an intensive parameter.

The internal energy E of a system is an extensive quantity. Indeed, no work is required to subdivide the system into two parts (*if* one neglects the work involved in creating the two new surfaces; this is negligible for large systems for which the ratio of the number of molecules near the boundary to the number of molecules in the bulk of the system is very small). Thus the total energy of the system is the same after subdivision as it was before, i.e., $\bar{E}_1 + \bar{E}_2 = \bar{E}$.

The heat capacity C , being the ratio of an energy increase divided by a fixed small temperature increment, is similarly an extensive quantity. The specific heat per mole (or per gram) is, by its definition C/ν (where ν is the number of moles in the system), obviously an intensive quantity.

The entropy S is also an extensive quantity. This follows from the relation $\Delta S = \int dQ/T$, since the heat absorbed $dQ = C dT$ is an extensive quantity. It also follows from the statistical definition, e.g., from (3·7·20).

When dealing with extensive quantities such as the entropy S , it is often convenient to talk in terms of the quantity per mole S/ν which is an intensive parameter independent of the size of the system. It is sometimes convenient to denote the quantity per mole by a small letter, e.g., to denote the entropy per mole by s . Thus $S = \nu s$.

SUGGESTIONS FOR SUPPLEMENTARY READING

Macroscopic discussion of internal energy, heat, and temperature

M. W. Zemansky: "Heat and Thermodynamics," 4th ed., chaps. 1 and 4, McGraw-Hill Book Company, New York, 1959.

H. B. Callen: "Thermodynamics," chap. 1, John Wiley & Sons, Inc., New York, 1960.

Consequences of the third law

E. Fermi: "Thermodynamics," chap. 8, Dover Publications, New York, 1956.

J. Wilks: "The Third Law of Thermodynamics," Oxford University Press, Oxford, 1961. (A more advanced book.)

PROBLEMS

- 4.1** (a) One kilogram of water at 0°C is brought into contact with a large heat reservoir at 100°C . When the water has reached 100°C , what has been the change in entropy of the water? of the heat reservoir? of the entire system consisting of both water and heat reservoir?

(b) If the water had been heated from 0°C to 100°C by first bringing it in contact with a reservoir at 50°C and then with a reservoir at 100°C , what would have been the change in entropy of the entire system?

(c) Show how the water might be heated from 0°C to 100°C with no change in the entropy of the entire system.

- 4.2** A 750-g copper calorimeter can containing 200 g of water is in equilibrium at a temperature of 20°C . An experimenter now places 30 g of ice at 0°C in the calorimeter and encloses the latter with a heat-insulating shield.

(a) When all the ice has melted and equilibrium has been reached, what will be the temperature of the water? (The specific heat of copper is 0.418 joules g^{-1} deg^{-1} . Ice has a specific gravity of 0.917 and its heat of fusion is 333 joules g^{-1} ; i.e., it requires 333 joules of heat to convert 1 g of ice to water at 0°C .)

(b) Compute the total entropy change resulting from the process of part (a).

(c) After all the ice has melted and equilibrium has been reached, how much work, in joules, must be supplied to the system (e.g., by means of a stirring rod) to restore all the water to 20°C ?

- 4.3** The heat absorbed by a mole of ideal gas in a quasi-static process in which its temperature T changes by dT and its volume V by dV is given by

$$dQ = c dT + \bar{p} dV$$

where c is its constant molar specific heat at constant volume and \bar{p} is its mean pressure, $\bar{p} = RT/V$. Find an expression for the change of entropy of this gas in a quasi-static process which takes it from initial values of temperature T_i and volume V_i to the final values T_f and V_f . Does your answer depend on the process involved in going from the initial to the final state?

- 4.4** A solid contains N magnetic atoms having spin $\frac{1}{2}$. At sufficiently high temperatures, each spin is completely randomly oriented, i.e., equally likely to be in either of its two possible states. But at sufficiently low temperatures the interactions between the magnetic atoms causes them to exhibit ferromagnetism,

with the result that all their spins become oriented along the same direction as $T \rightarrow 0$. A very crude approximation suggests that the spin-dependent contribution $C(T)$ to the heat capacity of this solid has an approximate temperature dependence given by

$$\begin{aligned} C(T) &= C_1 \left(2 \frac{T}{T_1} - 1 \right) && \text{if } \frac{1}{2}T_1 < T < T_1 \\ &= 0 && \text{otherwise} \end{aligned}$$

The abrupt increase in specific heat as T is reduced below T_1 is due to the onset of ferromagnetic behavior.

Use entropy considerations to find an explicit expression for the maximum value C_1 of the heat capacity.

- 4.5 A solid contains N magnetic iron atoms having spin S . At sufficiently high temperatures, each spin is completely randomly oriented, i.e., equally likely to be in any of its $2S + 1$ possible states. But at sufficiently low temperatures the interactions between the magnetic atoms causes them to exhibit ferromagnetism, with the result that all their spins become oriented along the same direction as $T \rightarrow 0$. The magnetic atoms contribute to the solid a heat capacity $C(T)$ which has, very crudely, an approximate temperature dependence given by

$$\begin{aligned} C(T) &= C_1 \left(2 \frac{T}{T_1} - 1 \right) && \text{if } \frac{1}{2}T_1 < T < T_1 \\ &= 0 && \text{otherwise} \end{aligned}$$

The abrupt increase in specific heat as T is reduced below T_1 is due to the onset of ferromagnetic behavior.

If one dilutes the magnetic atoms by replacing 30 percent of the iron atoms by nonmagnetic zinc atoms, then the remaining 70 percent of the iron atoms still become ferromagnetic at sufficiently low temperatures. The magnetic atoms now contribute to the solid a heat capacity $C(T)$ with the different temperature dependence given, very crudely, by

$$\begin{aligned} C(T) &= C_2 \frac{T}{T_2} && \text{if } 0 < T < T_2 \\ &= 0 && \text{otherwise} \end{aligned}$$

Since the interactions between magnetic ions have been reduced, the ferromagnetic behavior now sets in at a temperature T_2 lower than the previous temperature T_1 , and the heat capacity falls off more slowly below the temperature T_2 .

Use entropy considerations to compare the magnitude of the specific heat maximum C_2 in the dilute case with the magnitude of the specific heat maximum C_1 in the undiluted case. Find an explicit expression for C_2/C_1 .

Simple applications of macroscopic thermodynamics

5

IN THIS chapter we shall explore the purely macroscopic consequences of our theory in order to derive various important relationships between macroscopic quantities. The whole chapter will be based solely on the general statements, called "thermodynamic laws," which were derived in Chapter 3 and summarized in Sec. 3·11. Despite their apparent innocuousness, these statements allow one to draw an impressive number of remarkable conclusions which are completely *independent* of any specific models assumed to describe the microscopic constituents of a system.

Since the discussion of this chapter will be completely macroscopic, quantities such as energy E and pressure p will always refer to their respective mean values. For simplicity, we shall therefore omit the averaging bar symbols above the letters designating these quantities.

Most of the systems considered in this chapter will be characterized by a single external parameter, the volume V . The macrostate of such a system can then be specified completely by two macroscopic variables: its external parameter V and its internal energy E .^{*} The other macroscopic parameters, like temperature T or pressure p , are then determined. But the quantities V and E do not always represent the most convenient choice of independent variables. Any two other macroscopic parameters, e.g., E and p , or T and V , might equally well be chosen as independent variables. In either case, E and V would then be determined.

Most of the mathematical manipulations encountered in making thermodynamic calculations involve changing variables and taking partial derivatives. To avoid ambiguity, it is customary to indicate explicitly by subscripts which of the independent variables are kept constant in evaluating a given partial derivative. For example, if T and V are chosen as independent variables, $(\partial E / \partial T)_V$ denotes a partial derivative where the other independent variable V is kept constant. On the other hand, if T and p are chosen as independent

* These are, of course, the same variables that specify the number of states $\Omega(E, V)$ accessible to the system.

variables, $(\partial E/\partial T)_p$ denotes a partial derivative where the other independent variable p is kept constant. These two partial derivatives are, in general, not equal. If one simply wrote the partial derivative $(\partial E/\partial T)$ without a subscript, it would not be clear which is the other independent variable kept constant in the differentiation.*

The first law (3·11·2) applied to any infinitesimal process yields the relation

$$dQ = dE + dW$$

where dE is the change of internal energy of the system under consideration. If the process is quasi-static, the second law (3·11·2) allows one to express the heat dQ absorbed by the system in this process in terms of the change of entropy of the system, i.e., $dQ = T dS$; furthermore, the work done by the system when its volume is changed by an amount dV in the process is simply given by $dW = p dV$. Hence one obtains the fundamental thermodynamic relation

► $T dS = dE + p dV$

Most of this chapter will be based on this one equation. Indeed, it is usually simplest to make this fundamental relation the starting point for discussing any problem.

PROPERTIES OF IDEAL GASES

5·1 *Equation of state and internal energy*

Macroscopically, an ideal gas is described by the equation of state relating its pressure p , volume V , and absolute temperature T . For ν moles of gas, this equation of state is given by

$$pV = \nu RT \quad (5\cdot1\cdot1)$$

We derived this relation in (3·12·10) by *microscopic* arguments of statistical mechanics applied to an ideal gas in the classical limit. But from the point of view of the present macroscopic discussion, Eq. (5·1·1) merely characterizes the kind of system we are talking about; thus (5·1·1) might equally well be considered as a purely phenomenological relation summarizing experimental measurements on the system.

An ideal gas has a second important property already proved in (3·12·11) on the basis of microscopic statistical mechanics: its internal energy does not depend on its volume, but only on its temperature. Thus

$$E = E(T) \quad \text{independent of } V \quad (5\cdot1\cdot2)$$

This property is actually a direct consequence of the equation of state (5·1·1). Thus, even if (5·1·1) were to be considered a purely empirical equation of state describing a particular gas, the thermodynamic laws would allow

* Further discussion of partial derivatives can be found in Appendix A·9.

us to conclude immediately that this gas must satisfy the property (5·1·2). Let us show this explicitly.

Quite generally, the internal energy E of ν moles of any gas can be considered a function of T and V ,

$$E = E(T, V) \quad (5 \cdot 1 \cdot 3)$$

Thus we can write the purely mathematical statement

$$dE = \left(\frac{\partial E}{\partial T} \right)_V dT + \left(\frac{\partial E}{\partial V} \right)_T dV \quad (5 \cdot 1 \cdot 4)$$

But the fundamental thermodynamic relation for a quasi-static continuous change of parameters can be written as

$$\blacktriangleright \quad T dS = dQ = dE + p dV \quad (5 \cdot 1 \cdot 5)$$

Using (5·1·1) to express p in terms of V and T , (5·1·5) becomes

$$dS = \frac{1}{T} dE + \frac{\nu R}{V} dV$$

or by (5·1·4)

$$dS = \frac{1}{T} \left(\frac{\partial E}{\partial T} \right)_V dT + \left[\frac{1}{T} \left(\frac{\partial E}{\partial V} \right)_T dT + \frac{\nu R}{V} \right] dV \quad (5 \cdot 1 \cdot 6)$$

The mere fact that dS on the left side of (5·1·6) is the exact differential of a well-defined function allows us to draw an important conclusion. We can consider S dependent on T and V . Thus $S = S(T, V)$, and we can write the mathematical statement

$$dS = \left(\frac{\partial S}{\partial T} \right)_V dT + \left(\frac{\partial S}{\partial V} \right)_T dV \quad (5 \cdot 1 \cdot 7)$$

Since this expression must be true for all values of dT and dV , comparison with (5·1·4) shows immediately that

$$\begin{aligned} \left(\frac{\partial S}{\partial T} \right)_V &= \frac{1}{T} \left(\frac{\partial E}{\partial T} \right)_V \\ \left(\frac{\partial S}{\partial V} \right)_T &= \frac{1}{T} \left(\frac{\partial E}{\partial V} \right)_T + \frac{\nu R}{V} \end{aligned} \quad (5 \cdot 1 \cdot 8)$$

But the equality of the second derivatives, irrespective of order of differentiation,

$$\frac{\partial^2 S}{\partial V \partial T} = \frac{\partial^2 S}{\partial T \partial V} \quad (5 \cdot 1 \cdot 9)$$

implies a definite connection between the expressions on the right side of (5·1·8). Thus

$$\left(\frac{\partial}{\partial V} \right)_T \left(\frac{\partial S}{\partial T} \right)_V = \left(\frac{\partial}{\partial T} \right)_V \left(\frac{\partial S}{\partial V} \right)_T$$

by (5·1·9), and

$$\frac{1}{T} \left(\frac{\partial^2 E}{\partial V \partial T} \right) = \left[-\frac{1}{T^2} \left(\frac{\partial E}{\partial V} \right)_T + \frac{1}{T} \left(\frac{\partial^2 E}{\partial T \partial V} \right) \right] + 0$$

using (5·1·8). Since the second derivatives of E are again (analogously to (5·1·9)) equal irrespective of the order of differentiation, this last relation shows immediately that

$$\left(\frac{\partial E}{\partial V}\right)_T = 0 \quad (5\cdot1\cdot10)$$

This establishes that E is independent of V and completes the proof that (5·1·2) follows from (5·1·1).

Historical remark on the “free expansion” experiment The fact that the internal energy E of a gas does not depend on its volume (if the gas is sufficiently dilute that it can be considered ideal) was verified in a classical experiment by Joule. He made use of the “free expansion” of an ideal gas as illustrated in Fig. 5·1·1.

A container consisting of two compartments separated by a valve is immersed in water. Initially, the valve is closed and one compartment is filled with the gas under investigation, while the other compartment is evacuated. Suppose that the valve is now opened so that the gas is free to expand and fill both compartments. In this process no work gets done by the system consisting of the gas and container. (The container walls are rigid and nothing moves.) Hence one can say, by the first law, that the heat Q absorbed by this system equals its increase in internal energy,

$$Q = \Delta E \quad (5\cdot1\cdot11)$$

Assume that the internal energy change of the (thin-walled) container is negligibly small. Then ΔE measures simply the energy change of the gas.

Joule found that the temperature of the water did not change in this experiment. (Because of the large heat capacity of the water, any anticipated temperature change is, however, quite small; Joule's actual sensitivity of temperature measurement was, in retrospect, rather inadequate.) Thus the water absorbed no heat from the gas; consequently, the heat Q absorbed by the gas

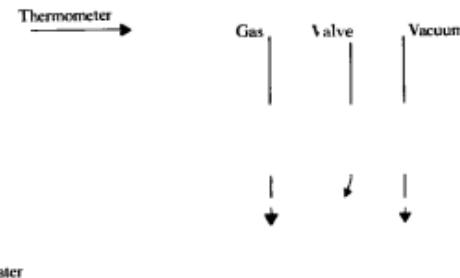


Fig. 5·1·1 Experimental setup for studying the free expansion of a gas.

also vanished. All that happens in the experiment is that the temperature of the gas remains unchanged while its volume changes from its initial value V_i to its final value V_f . Since $Q = 0$, Joule's experiment leads by virtue of (5·1·11) to the conclusion

$$E(T, V_f) - E(T, V_i) = 0$$

which verifies that $E(T, V)$ is independent of the volume V .

5 · 2 Specific heats

The heat absorbed in an infinitesimal process is given by the first law as

$$dQ = dE + p \, dV \quad (5 \cdot 2 \cdot 1)$$

Let us first obtain an expression for the molar specific heat c_v at constant volume. Then $dV = 0$ and (5·2·1) reduces simply to

$$dQ = dE$$

Hence one obtains

$$c_v \equiv \frac{1}{v} \left(\frac{dQ}{dT} \right)_V = \frac{1}{v} \left(\frac{\partial E}{\partial T} \right)_V \quad (5 \cdot 2 \cdot 2)$$

The specific heat c_v may itself, of course, be a function of T . But by virtue of (5·1·2), it is independent of V for an ideal gas.

Since E is independent of V , so that E is only a function of T , the general relation (5·1·4) reduces to

$$dE = \left(\frac{\partial E}{\partial T} \right)_V \, dT \quad (5 \cdot 2 \cdot 3)$$

i.e., the change of energy depends *only* on the temperature change of the gas, even if the volume of the gas also changes. Using (5·2·2), one can then write quite generally

$$dE = \nu c_v \, dT \quad (5 \cdot 2 \cdot 4)$$

for an *ideal* gas.

Let us now obtain an expression for the molar specific heat c_p at constant pressure. Here the pressure is constant, but the volume V changes, in general, as heat is added to the gas. Since the general expression (5·2·4) for dE is still valid, one can substitute it into (5·2·2) to get

$$dQ = \nu c_v \, dT + p \, dV \quad (5 \cdot 2 \cdot 5)$$

We now want to make use of the fact that the pressure p is kept constant. By the equation of state (5·1·1), a volume change dV and a temperature change dT are then related by

$$p \, dV = \nu R \, dT \quad (5 \cdot 2 \cdot 6)$$

Substituting this into (5·2·5) yields for the heat absorbed at constant pressure

$$dQ = \nu c_v \, dT + \nu R \, dT \quad (5 \cdot 2 \cdot 7)$$

But, by definition,

$$c_p = \frac{1}{\nu} \left(\frac{dQ}{dT} \right)_p$$

Using (5·2·7), this becomes

$$\blacktriangleright \quad c_p = c_v + R \quad (5·2·8)$$

Thus $c_p > c_v$, in general agreement with (4·4·7), and these molar specific heats of an ideal gas differ precisely by the gas constant R .

The ratio γ of the specific heats is then given by

$$\gamma \equiv \frac{c_p}{c_v} = 1 + \frac{R}{c_v} \quad (5·2·9)$$

The quantity γ can be determined from the velocity of sound in the gas and can also be measured directly by other methods. Table 5·2·1 lists some representative experimental values of c_v for a few gases. It also illustrates the extent of agreement between the values of γ computed by (5·2·9) and the experimentally measured values of this quantity.

Table 5·2·1 Specific heats of some gases (at 15°C and 1 atm)*

Gas	Symbol	c_v (experimental) (joules mole ⁻¹ deg ⁻¹)	γ (experimental)	γ (computed by (5·2·9))
Helium	He	12.5	1.666	1.666
Argon	Ar	12.5	1.666	1.666
Nitrogen	N ₂	20.6	1.405	1.407
Oxygen	O ₂	21.1	1.396	1.397
Carbon dioxide	CO ₂	28.2	1.302	1.298
Ethane	C ₂ H ₆	39.3	1.220	1.214

* Experimental values taken from J. R. Partington and W. G. Shilling, "The Specific Heats of Gases," p. 201, Benn, London, 1924.

Microscopic calculation of specific heats If one is willing to make use of microscopic information, one can, of course, make many more interesting statements. The situation of a *monatomic* ideal gas is particularly simple. In (2·5·19) we found for the number of states of such a gas in some small energy range δE the expression

$$\Omega(E, V) = B V^N E^{3N/2}$$

where N is the number of molecules in the gas and B is some constant independent of E and V . Hence

$$\ln \Omega = \ln B + N \ln V + \frac{3N}{2} \ln E$$

The temperature parameter $\beta = (kT)^{-1}$ is then given by

$$\beta = \frac{\partial \ln \Omega}{\partial E} = \frac{3N}{2} \frac{1}{E}$$

Thus $E = \frac{3N}{2\beta} = \frac{3N}{2} kT \quad (5 \cdot 2 \cdot 10)$

This expresses directly the relation between the internal energy and the absolute temperature of the gas. If N_a denotes Avogadro's number, $N = \nu N_a$ and (5 · 2 · 10) can also be written

$$E = \frac{3}{2}\nu(N_a k)T = \frac{3}{2}\nu RT \quad (5 \cdot 2 \cdot 11)$$

where $R = N_a k$ is the gas constant.

The molar specific heat at constant volume of a monatomic ideal gas is then, by (5 · 2 · 2) and (5 · 2 · 11),

$$c_v = \frac{1}{\nu} \left(\frac{\partial E}{\partial T} \right)_v = \frac{3}{2} R \quad (5 \cdot 2 \cdot 12)$$

By (4 · 3 · 8) this has the numerical value

$$c_v = 12.47 \text{ joules deg}^{-1} \text{ mole}^{-1} \quad (5 \cdot 2 \cdot 13)$$

Furthermore, (5 · 2 · 8) then gives

$$c_p = \frac{3}{2}R + R = \frac{5}{2}R \quad (5 \cdot 2 \cdot 14)$$

and $\gamma \equiv \frac{c_p}{c_v} = \frac{5}{3} = 1.667 \quad (5 \cdot 2 \cdot 15)$

These simple microscopic arguments lead thus to very definite quantitative predictions. The experimental values given in Table 5 · 2 · 1 for the monatomic gases helium and argon show very satisfactory agreement with the theoretical values (5 · 2 · 13) and (5 · 2 · 15).

5 · 3 Adiabatic expansion or compression

Suppose that the temperature of a gas is maintained constant by being kept in thermal contact with a heat reservoir. If the gas is allowed to expand quasi-statically under such "isothermal" (i.e., "same temperature") conditions, the pressure p and volume V satisfy, by virtue of the equation of state (5 · 1 · 1), the relation

$$pV = \text{constant} \quad (5 \cdot 3 \cdot 1)$$

Suppose, however, that the gas is thermally insulated from its surroundings (i.e., that it is maintained under adiabatic conditions). If the gas is allowed to expand under these conditions, it will do work at the expense of its internal energy; as a result its temperature will also change. In a quasi-static adiabatic process of this kind, how is the pressure p of the gas related to its volume V ?

Our starting point is again the first law (5·2·1). Since no heat is absorbed in the adiabatic process here contemplated, $dQ = 0$. Using Eq. (5·2·4) for an ideal gas, (5·2·1) becomes

$$0 = \nu c_V dT + p dV \quad (5\cdot3\cdot2)$$

This relation involves the three variables p , V , and T . By the equation of state (5·1·1), one can express one of these in terms of the other two. Thus (5·1·1) yields

$$p dV + V dp = \nu R dT \quad (5\cdot3\cdot3)$$

Let us solve this for dT and substitute the result into (5·3·2). This gives a relation between dp and dV .

$$0 = \frac{c_V}{R} (p dV + V dp) + p dV = \left(\frac{c_V}{R} + 1 \right) p dV + \frac{c_V}{R} V dp$$

or

$$(c_V + R)p dV + c_V V dp = 0$$

Dividing both sides of this equation by the quantity $c_V p V$ yields the equation

$$\gamma \frac{dV}{V} + \frac{dp}{p} = 0 \quad (5\cdot3\cdot4)$$

where, by (5·2·9),

$$\gamma \equiv \frac{c_V + R}{c_V} = \frac{c_p}{c_V} \quad (5\cdot3\cdot5)$$

Now c_V is temperature independent for most gases. In other cases it may be a slowly varying function of T . Thus it is always an excellent approximation to assume that the specific heat ratio γ is independent of T in a limited temperature range. Then (5·3·4) can be immediately integrated to give

$$\gamma \ln V + \ln p = \text{constant}$$

or

$$\blacktriangleright \quad p V^\gamma = \text{constant} \quad (5\cdot3\cdot6)$$

Since $\gamma > 1$ by virtue of (5·3·5), p will vary more rapidly with V than in the isothermal case (5·3·1) where $pV = \text{constant}$.

From (5·3·6) one can, of course, also obtain corresponding relations between V and T , or between p and T . For example, since $p = \nu RT/V$, (5·3·6) implies that

$$V^{\gamma-1} T = \text{constant} \quad (5\cdot3\cdot7)$$

5 · 4 Entropy

The entropy of an ideal gas can readily be computed from the fundamental thermodynamic relation (5 · 1 · 5) by the procedure of Sec. 4 · 5. By virtue of (5 · 2 · 4) and the equation of state (5 · 1 · 1), the relation (5 · 1 · 5) becomes

$$T dS = \nu c_V(T) dT + \frac{\nu R T}{V} dV$$

$$\text{or } dS = \nu c_V(T) \frac{dT}{T} + \nu R \frac{dV}{V} \quad (5 \cdot 4 \cdot 1)$$

This allows one to find, by integration, the entropy of ν moles of this gas at any arbitrary temperature T and volume V compared to the entropy of the gas in some standard macrostate.

Let us choose as the standard macrostate of this kind of gas one where ν_0 moles of the gas occupy a volume V_0 at the temperature T_0 . We denote the *molar* entropy of the gas in this standard state by s_0 . To calculate the entropy $S(T, V; \nu)$ of ν moles of this gas at temperature T and volume V , we need merely to consider *any* quasi-static process whereby we bring these ν moles of gas from the standard state to the final state of interest. Let us then first divide off by a partition ν moles of gas in the standard state; these will have entropy νs_0 and occupy a volume $V_0(\nu/\nu_0)$. Take these ν moles of gas and slowly increase the temperature to the value T while keeping the volume constant at $V_0(\nu/\nu_0)$. Then change the volume slowly to the value V while keeping the temperature constant at T . For the process just described, integration of (5 · 4 · 1) gives the result

$$S(T, V; \nu) - \nu s_0 = \nu \int_{T_0}^T \frac{c_V(T') dT'}{T'} + \nu R \int_{V_0(\nu/\nu_0)}^V \frac{dV'}{V'} \quad (5 \cdot 4 \cdot 2)$$

The last integration is immediate:

$$\int_{V_0(\nu/\nu_0)}^V \frac{dV'}{V'} = [\ln V']_{V_0(\nu/\nu_0)}^V = \ln V - \ln \left(V_0 \frac{\nu}{\nu_0} \right) = \ln \frac{V}{\nu} - \ln \frac{V_0}{\nu_0}$$

Hence (5 · 4 · 2) becomes

$$S(T, V; \nu) = \nu \left[\int_{T_0}^T \frac{c_V(T')}{T'} dT' + R \ln \frac{V}{\nu} - R \ln \frac{V_0}{\nu_0} + s_0 \right] \quad (5 \cdot 4 \cdot 3)$$

$$\text{or } S(T, V; \nu) = \nu \left[\int \frac{c_V(T')}{T'} dT' + R \ln V - R \ln \nu + \text{constant} \right] \quad (5 \cdot 4 \cdot 4)$$

In this last expression we have lumped all the quantities referring to the standard state into a single constant. The expressions (5 · 4 · 3) or (5 · 4 · 4) give the dependence of the entropy S on T , V , and ν . In the special case when c_V is temperature independent, the integral over temperature becomes, of course, trivial; i.e.,

$$\int \frac{c_V}{T'} dT' = c_V \ln T$$

if c_V is constant.

GENERAL RELATIONS FOR A HOMOGENEOUS SUBSTANCE

5 · 5 *Derivation of general relations*

We consider a homogeneous system whose volume V is the only external parameter of relevance. The starting point of our whole discussion is again the fundamental thermodynamic relation for a quasi-static infinitesimal process

$$dQ = T dS = dE + p dV \quad (5 \cdot 5 \cdot 1)$$

This equation gives rise to a wealth of other relations which we want to exhibit presently.

Independent variables S and V Equation (5 · 5 · 1) can be written

► $dE = T dS - p dV \quad (5 \cdot 5 \cdot 2)$

This shows how E depends on independent variations of the parameters S and V . If these are considered the two independent parameters specifying the system, then

$$E = E(S, V)$$

and one can write the corresponding purely mathematical statement

$$dE = \left(\frac{\partial E}{\partial S} \right)_V dS + \left(\frac{\partial E}{\partial V} \right)_S dV \quad (5 \cdot 5 \cdot 3)$$

Since (5 · 5 · 2) and (5 · 5 · 3) must be equal for all possible values of dS and dV , it follows that the corresponding coefficients of dS and dV must be the same. Hence

$$\begin{aligned} \left(\frac{\partial E}{\partial S} \right)_V &= T \\ \left(\frac{\partial E}{\partial V} \right)_S &= -p \end{aligned} \quad (5 \cdot 5 \cdot 4)$$

The important content of the relation (5 · 5 · 2) is that the combination of parameters on the right side is always equal to the exact differential of a quantity, which in this case is the energy E . Hence the parameters T , S , p , and V which occur on the right side of (5 · 5 · 2) cannot be varied completely arbitrarily; there must exist some connection between them to guarantee that their combination yields the differential dE . To obtain this connection, it is only necessary to note that the second derivatives of E must be independent of the order of differentiation, i.e.,

$$\frac{\partial^2 E}{\partial V \partial S} = \frac{\partial^2 E}{\partial S \partial V}$$

or $\left(\frac{\partial}{\partial V} \right)_S \left(\frac{\partial E}{\partial S} \right)_V = \left(\frac{\partial}{\partial S} \right)_V \left(\frac{\partial E}{\partial V} \right)_S$

Hence one obtains by (5 · 5 · 4) the result

$$\left(\frac{\partial T}{\partial V}\right)_S = - \left(\frac{\partial p}{\partial S}\right)_V \quad (5 \cdot 5 \cdot 5)$$

This useful relation reflects merely the fact that dE is the exact differential of a well-defined quantity E characteristic of the macrostate of the system.

Independent variables S and p Equation (5 · 5 · 2) exhibits the effect of independent variations of S and V . One might equally well exhibit the effect of independent variations of S and p . One can easily pass from the expression $p dV$, where the variation dV appears, to an equivalent expression where dp appears by the simple transformation

$$p dV = d(pV) - V dp$$

Let us then return to (5 · 5 · 2) and transform this into an expression involving dp rather than dV . We get

$$dE = T dS - p dV = T dS - d(pV) + V dp$$

or $d(E + pV) = T dS + V dp$

We can write this as

$$dH = T dS + V dp \quad (5 \cdot 5 \cdot 6)$$

where we have introduced the definition

$$H \equiv E + pV \quad (5 \cdot 5 \cdot 7)$$

The function H is called the “enthalpy.”

Considering S and p as independent variables, one can write

$$H = H(S, p)$$

and $dH = \left(\frac{\partial H}{\partial S}\right)_p dS + \left(\frac{\partial H}{\partial p}\right)_S dp \quad (5 \cdot 5 \cdot 8)$

Comparison between (5 · 5 · 6) and (5 · 5 · 8) yields the relations

$$\begin{aligned} \left(\frac{\partial H}{\partial S}\right)_p &= T \\ \left(\frac{\partial H}{\partial p}\right)_S &= V \end{aligned} \quad (5 \cdot 5 \cdot 9)$$

The important aspect of (5 · 5 · 6) is again the fact that the combination of parameters on the right side is equal to the exact differential of a quantity which we happen to have designated by the letter H . The equality of the cross derivatives of this quantity, i.e., the equality

$$\frac{\partial^2 H}{\partial p \partial S} = \frac{\partial^2 H}{\partial S \partial p}$$

then implies immediately the relation

$$\left(\frac{\partial T}{\partial p}\right)_S = \left(\frac{\partial V}{\partial S}\right)_p \quad (5 \cdot 5 \cdot 10)$$

This equation is analogous to (5·5·5) and represents again a necessary connection between the parameters T, S, p, V . By this time it should be clear what kind of game one plays to get thermodynamic relations of this sort. All we need do is play it to the bitter end by performing all other possible changes of variable in the fundamental equation (5·5·2).

Independent variables T and V We transform (5·5·2) into an expression involving dT rather than dS . Thus we can write

$$\begin{aligned} dE &= T dS - p dV = d(TS) - S dT - p dV \\ \text{or} \quad dF &= -S dT - p dV \end{aligned} \quad (5 \cdot 5 \cdot 11)$$

where we have introduced the definition

$$F \equiv E - TS \quad (5 \cdot 5 \cdot 12)$$

The function F is called the "Helmholtz free energy."

Considering T and V as independent variables,

$$\begin{aligned} F &= F(T, V) \\ \text{and} \quad dF &= \left(\frac{\partial F}{\partial T}\right)_V dT + \left(\frac{\partial F}{\partial V}\right)_T dV \end{aligned} \quad (5 \cdot 5 \cdot 13)$$

Comparison of (5·5·11) with (5·5·13) yields

$$\begin{aligned} \left(\frac{\partial F}{\partial T}\right)_V &= -S \\ \left(\frac{\partial F}{\partial V}\right)_T &= -p \end{aligned} \quad (5 \cdot 5 \cdot 14)$$

Equality of the cross derivatives

$$\frac{\partial^2 F}{\partial V \partial T} = \frac{\partial^2 F}{\partial T \partial V}$$

then implies

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial p}{\partial T}\right)_V \quad (5 \cdot 5 \cdot 15)$$

Independent variables T and p We finally transform (5·5·2) into an expression involving dT and dp rather than dS and dV . Thus we can write

$$\begin{aligned} dE &= T dS - p dV = d(TS) - S dT - d(pV) + V dp \\ \text{or} \quad dG &= -S dT + V dp \end{aligned} \quad (5 \cdot 5 \cdot 16)$$

where we have introduced the definition

$$G \equiv E - TS + pV \quad (5 \cdot 5 \cdot 17)$$

The function G is called the “Gibbs free energy.” In terms of the previous definitions (5·5·7) or (5·5·12), we could also write $G = H - TS$, or $G = F + pV$.

Considering T and p as independent variables,

$$\text{and } G = G(T, p) \quad dG = \left(\frac{\partial G}{\partial T}\right)_p dT + \left(\frac{\partial G}{\partial p}\right)_T dp \quad (5 \cdot 5 \cdot 18)$$

Comparison of (5·5·16) with (5·5·18) yields

$$\begin{aligned} \left(\frac{\partial G}{\partial T}\right)_p &= -S \\ \left(\frac{\partial G}{\partial p}\right)_T &= V \end{aligned} \quad (5 \cdot 5 \cdot 19)$$

Equality of the cross derivatives

$$\frac{\partial^2 G}{\partial p \partial T} = \frac{\partial^2 G}{\partial T \partial p}$$

then implies

$$-\left(\frac{\partial S}{\partial p}\right)_T = \left(\frac{\partial V}{\partial T}\right)_p \quad (5 \cdot 5 \cdot 20)$$

5 · 6 Summary of Maxwell relations and thermodynamic functions

Maxwell relations The entire discussion of the preceding section was based upon the fundamental thermodynamic relation

► $dE = T dS - p dV \quad (5 \cdot 6 \cdot 1)$

From this statement we derived the important relations (5·5·5), (5·5·10), (5·5·15), and (5·5·20), which are repeated below:

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial p}{\partial S}\right)_V \quad (5 \cdot 6 \cdot 2)$$

$$\left(\frac{\partial T}{\partial p}\right)_S = \left(\frac{\partial V}{\partial S}\right)_p \quad (5 \cdot 6 \cdot 3)$$

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial p}{\partial T}\right)_V \quad (5 \cdot 6 \cdot 4)$$

$$\left(\frac{\partial S}{\partial p}\right)_T = -\left(\frac{\partial V}{\partial T}\right)_p \quad (5 \cdot 6 \cdot 5)$$

These are known as “Maxwell’s relations.” They are a direct consequence of the fact that the variables T , S , p , and V are not completely independent, but are related through the fundamental thermodynamic relation (5·6·1). All

the Maxwell relations are basically equivalent;* any one of them can be derived from any other one by a simple change of independent variables.

It is worth recalling explicitly why there exists this connection between variables which is expressed by the Maxwell relations. The basic reason is as follows: It is possible to give a complete macroscopic description of a system in equilibrium if one knows the number of states Ω accessible to the system (or equivalently, its entropy $S = k \ln \Omega$) as a function of its energy E and its one external parameter V . But both the temperature T and mean pressure p of the system can be expressed in terms of $\ln \Omega$ or S ; in Chapter 3 we found the explicit expressions (3·12·1) or, equivalently, (3·12·5)

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_V \quad \text{and} \quad p = T \left(\frac{\partial S}{\partial V} \right)_E \quad (5 \cdot 6 \cdot 6)$$

It is the fact that both T and p are expressible in terms of the same function S which leads to the connection (5·6·1) and hence to the Maxwell relations.

Indeed, one has

$$\begin{aligned} dS &= \left(\frac{\partial S}{\partial E} \right)_V dE + \left(\frac{\partial S}{\partial V} \right)_E dV \\ &= \frac{1}{T} dE + \frac{p}{T} dV \quad \text{by (5·6·6)} \end{aligned}$$

and the latter expression is simply the fundamental relation (5·6·1)

Note that the fundamental relation (5·6·1) involves the variables on the right side in pairs, one pair consisting of T and S , the other of p and V . In these two pairs

$$(T, S) \quad \text{and} \quad (p, V)$$

the first involves quantities like entropy and temperature which describe the density of accessible states of the system, whereas the second involves an external parameter and its corresponding generalized force. The essential content of the Maxwell relations is the existence of a connection between the cross derivatives of these two kinds of quantities. Specifically, each of the Maxwell relations is a statement asserting that [the derivative of a variable of the first pair with respect to a variable of the second pair] is (except for sign) equal to [the corresponding derivative of the other variable of the second pair with respect to the other variable of the first pair].

The above property characterizing the Maxwell relations makes it very easy to read them off directly from the fundamental relation (5·6·1). The proper sign can be obtained in the following way. If the two variables with respect to which one differentiates are the same variables S and V which occur as differentials in (5·6·1), then the minus sign that occurs in (5·6·1) also

* They can, for example, all be summarized by the single statement that the Jacobian determinant $\partial(T, S)/\partial(p, V) = 1$.

occurs in the Maxwell relation. Any one permutation away from these particular variables introduces a change of sign.

Example A minus sign occurs in (5·6·2) because the variables S and V with respect to which one differentiates are the same as those appearing as differentials in (5·6·1). On the other hand, in (5·6·3) the derivatives are with respect to S and p , whereas S and V appear as differentials in (5·6·1). The switch from p to V implies one sign change with respect to the minus sign in (5·6·1); hence there is a plus sign in (5·6·3).

Thermodynamic functions The Maxwell relations constitute the most important result of the last section. It is, however, also worth summarizing for future reference the various thermodynamic functions that were introduced in that section. We list them below, together with the most convenient independent variables used in conjunction with each of these functions (i.e., the variables in terms of which the fundamental relation (5·6·1) is expressed most simply):

$$\begin{array}{ll} E & E = E(S, V) \\ H = E + pV & H = H(S, p) \\ F = E - TS & F = F(T, V) \\ G = E - TS + pV & G = G(T, P) \end{array} \quad (5 \cdot 6 \cdot 7)$$

Next we summarize the thermodynamic relations satisfied by each of these functions

$$dE = T dS - p dV \quad (5 \cdot 6 \cdot 8)$$

$$dH = T dS + V dp \quad (5 \cdot 6 \cdot 9)$$

$$dF = -S dT - p dV \quad (5 \cdot 6 \cdot 10)$$

$$dG = -S dT + V dp \quad (5 \cdot 6 \cdot 11)$$

The relations (5·5·4), (5·5·9), (5·5·14), and (5·5·19), involving derivatives of the functions E , H , F , and G , respectively, can immediately be read off from these equations.

The equations (5·6·9) through (5·6·11) are very simply related to the fundamental equation (5·6·8) or (5·6·1). It is only necessary to note that all of them involve the same variable pairs (T, S) and (p, V) , the variables entering as differentials being the independent variables listed in (5·6·7); and that any change of variable away from those used in (5·6·8) introduces a change of sign.

5 · 7 Specific heats

We consider any homogeneous substance whose volume V is the only relevant external parameter. Let us first investigate the general relation existing between the molar specific heat c_V at constant volume and the molar specific

heat c_p at constant pressure. This relation has practical importance, since calculations by statistical mechanics are usually more easily performed for an assumed fixed volume, while experimental measurements are more readily carried out under conditions of constant (say atmospheric) pressure. Thus, to compare the theoretically calculated quantity c_V with the experimentally measured parameter c_p , a knowledge of the relation between these quantities is necessary.

The heat capacity at constant volume is given by

$$C_V = \left(\frac{dQ}{dT} \right)_V = T \left(\frac{\partial S}{\partial T} \right)_V \quad (5 \cdot 7 \cdot 1)$$

The heat capacity at constant pressure is similarly given by

$$C_p = \left(\frac{dQ}{dT} \right)_p = T \left(\frac{\partial S}{\partial T} \right)_p \quad (5 \cdot 7 \cdot 2)$$

We seek a general relation between these two quantities.

Experimentally, the parameters which can be controlled most conveniently are the temperature T and pressure p . Let us consider these as independent variables. Then $S = S(T, p)$ and one obtains the following general expression for the heat dQ absorbed in an infinitesimal quasi-static process

$$dQ = T dS = T \left[\left(\frac{\partial S}{\partial T} \right)_p dT + \left(\frac{\partial S}{\partial p} \right)_T dp \right] \quad (5 \cdot 7 \cdot 3)$$

One can use (5 · 7 · 2) to write this in the form

$$dQ = T dS = C_p dT + T \left(\frac{\partial S}{\partial p} \right)_T dp \quad (5 \cdot 7 \cdot 4)$$

If the pressure is maintained constant, $dp = 0$, and (5 · 7 · 4) reduces to (5 · 7 · 2). But in calculating C_V by (5 · 7 · 1), T and V are used as the independent variables. To express dQ in (5 · 7 · 4) in terms of dT and dV , it is only necessary to express dp in terms of these differential quantities. This gives

$$dQ = T dS = C_p dT + T \left(\frac{\partial S}{\partial p} \right)_T \left[\left(\frac{\partial p}{\partial T} \right)_V dT + \left(\frac{\partial p}{\partial V} \right)_T dV \right] \quad (5 \cdot 7 \cdot 5)$$

The heat dQ absorbed under conditions when V is constant is then immediately obtained by putting $dV = 0$. Dividing this heat by dT gives C_V . Thus

$$C_V = T \left(\frac{\partial S}{\partial T} \right)_V = C_p + T \left(\frac{\partial S}{\partial p} \right)_T \left(\frac{\partial p}{\partial T} \right)_V \quad (5 \cdot 7 \cdot 6)$$

This is a relation between C_V and C_p , but it involves on the right side quantities which are not readily measured. For example, what is $(\partial S / \partial p)_T$? It is not readily measured, but since it is the derivative of a variable from the (T, S) pair with respect to a variable from the (p, V) pair, we can use one of the Maxwell relations. By (5 · 6 · 5)

$$\left(\frac{\partial S}{\partial p} \right)_T = - \left(\frac{\partial V}{\partial T} \right)_p \quad (5 \cdot 7 \cdot 7)$$

Here the quantity on the right is a readily measured and familiar quantity, since it is simply the change of volume with temperature under conditions of constant pressure. Indeed, one defines the intensive quantity

$$\alpha \equiv \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_p \quad (5 \cdot 7 \cdot 8)$$

as the "volume coefficient of expansion" of the substance. Thus

$$\left(\frac{\partial S}{\partial p} \right)_T = -V\alpha \quad (5 \cdot 7 \cdot 9)$$

The derivative $(\partial p / \partial T)_V$ is also not very readily determined, since it implies a measurement where the volume V is kept constant.* It is usually more convenient to control the temperature T and pressure p . But we can express V in terms of T and p . Thus

$$dV = \left(\frac{\partial V}{\partial T} \right)_p dT + \left(\frac{\partial V}{\partial p} \right)_T dp$$

and, under conditions of constant volume where $dV = 0$, this gives for the desired ratio dp/dT at constant volume the result

$$\left(\frac{\partial p}{\partial T} \right)_V = - \frac{\left(\frac{\partial V}{\partial T} \right)_p}{\left(\frac{\partial V}{\partial p} \right)_T} \quad (5 \cdot 7 \cdot 10)$$

(This is simply the result (A · 9 · 5), which we could have written down without rederiving it.) Here the numerator is again related to α by (5 · 7 · 8). The denominator is another familiar quantity, since it measures the change in volume of the substance with increasing pressure at a constant temperature. (The change of volume will be negative, since the volume decreases with increasing pressure.) One defines the positive intensive quantity

$$\kappa \equiv - \frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_T \quad (5 \cdot 7 \cdot 11)$$

as the "isothermal compressibility" of the substance. Hence (5 · 7 · 10) becomes

$$\left(\frac{\partial p}{\partial T} \right)_V = \frac{\alpha}{\kappa} \quad (5 \cdot 7 \cdot 12)$$

Substitution of (5 · 7 · 9) and (5 · 7 · 12) into (5 · 7 · 6) yields then

$$C_V = C_p + T(-V\alpha) \left(\frac{\alpha}{\kappa} \right)$$

or

$$C_p - C_V = VT \frac{\alpha^2}{\kappa} \quad (5 \cdot 7 \cdot 13)$$

* In the case of solids and liquids a small temperature increase at constant volume tends to produce a very large pressure increase. This imposes rather severe demands on the strength of the vessel containing the substance.

If C_p and C_v in this expression are the heat capacities per mole, then the corresponding volume V is the volume per mole.

Equation (5·7·13) provides the desired relation between C_p and C_v in terms of quantities which can be either readily measured directly or computed from the equation of state. For solids and liquids the right side of (5·7·13) is fairly small, so that C_p and C_v do not differ very much.

Numerical example Consider the case of copper at room temperature (298°K) and atmospheric pressure. The density of the metal is 8.9 g cm⁻³ and its atomic weight is 63.5; hence its molar volume is $V = 63.5/8.9 = 7.1 \text{ cm}^3 \text{ mole}^{-1}$. The other observed values* are $\alpha = 5 \times 10^{-6} \text{ deg}^{-1}$, $\kappa = 4.5 \times 10^{-11} \text{ cm}^2 \text{ dyne}^{-1}$, and $c_p = 24.5 \text{ joules deg}^{-1} \text{ mole}^{-1}$. Then one computes by (5·7·13) that $c_p - c_v = 1.2 \cdot 10^7 \text{ ergs deg}^{-1} \text{ mole}^{-1}$. Thus $c_v = 23.3 \text{ joules deg}^{-1} \text{ mole}^{-1}$ and $\gamma = c_p/c_v = 1.05$.

Simple application: ideal gas Let us apply (5·7·13) to the special case of the ideal gas discussed in Sec. 5·2. Then the equation of state (5·1·1) is

$$pV = \nu RT \quad (5\cdot7\cdot14)$$

We calculate first the expansion coefficient α defined by (5·7·8). For constant p

$$\begin{aligned} \text{Hence } & p dV = \nu R dT \\ & \left(\frac{\partial V}{\partial T} \right)_p = \frac{\nu R}{p} \\ \text{and } & \alpha = \frac{1}{V} \left(\frac{\nu R}{p} \right) = \frac{\nu R}{\nu R T} = \frac{1}{T} \end{aligned} \quad (5\cdot7\cdot15)$$

We calculate next the compressibility κ defined in (5·7·11). For constant T , (5·7·14) yields

$$\begin{aligned} \text{Hence } & p dV + V dp = 0 \\ & \left(\frac{\partial V}{\partial p} \right)_T = -\frac{V}{p} \\ \text{and } & \kappa = -\frac{1}{V} \left(-\frac{V}{p} \right) = \frac{1}{p} \end{aligned} \quad (5\cdot7\cdot16)$$

Thus (5·7·13) becomes

$$C_p - C_v = VT \frac{(1/T)^2}{1/p} = \frac{Vp}{T} = \nu R$$

or, per mole,

$$c_p - c_v = R \quad (5\cdot7\cdot17)$$

which agrees with our previous result (5·2·8).

Limiting properties near absolute zero The third law of thermodynamics (3·11·5) asserts that, as the temperature $T \rightarrow 0$, the entropy S of a system

* Data taken from "American Institute of Physics Handbook," 2d ed., McGraw-Hill Book Company, New York, 1963.

approaches smoothly some limiting constant value S_0 independent of all parameters of the system. In symbols*

$$\text{as } T \rightarrow 0, \quad S \rightarrow S_0 \quad (5 \cdot 7 \cdot 18)$$

In the limit as $T \rightarrow 0$, the derivatives $\partial S/\partial T$ appearing in (5·7·1) and (5·7·2) remain thus finite, and one can conclude from these relations that

$$\text{as } T \rightarrow 0, \quad C_V \rightarrow 0 \quad \text{and} \quad C_p \rightarrow 0 \quad (5 \cdot 7 \cdot 19)$$

The argument leading to (5·7·19) becomes particularly clear if one writes (5·7·1) or (5·7·2) in integrated form. For example, if the volume is kept constant,

$$S(T) - S(0) = \int_0^T \frac{C_V(T')}{T'} dT'$$

But since the entropy difference on the left must be finite, it must be true that $C_V(T) \rightarrow 0$ as $T \rightarrow 0$ in order to guarantee proper convergence of the integral on the right. The limiting behavior (5·7·19) of the heat capacities is not surprising. It merely reflects the fact that as $T \rightarrow 0$, the system tends to settle down in its ground state. The mean energy of the system then becomes essentially equal to its ground-state energy, and no further reduction of temperature can result in a further change of mean energy to a smaller value.

Since the limiting value approached by the entropy as $T \rightarrow 0$ is independent of all parameters of the system, it is also independent of volume or pressure variations in this limit. Hence $(\partial S/\partial p)_T \rightarrow 0$, and the Maxwell relation (5·7·7) applied to the definition (5·7·8) allows one to make a statement about the limiting behavior of the coefficient of expansion α ; i.e.,

$$\text{as } T \rightarrow 0, \quad \alpha \rightarrow 0 \quad (5 \cdot 7 \cdot 20)$$

On the other hand, the compressibility κ is a purely mechanical property and a system (e.g., a solid) in its ground state has a well-defined compressibility. Thus κ remains finite as $T \rightarrow 0$.

Since the product $T\alpha^2$ on the right side of (5·7·13) approaches zero very rapidly as $T \rightarrow 0$, it follows that the difference $C_p - C_V$ becomes increasingly negligible compared to C_V itself as one goes to very low temperatures; i.e.,

$$\text{as } T \rightarrow 0, \quad \frac{C_p - C_V}{C_V} \rightarrow 0 \quad (5 \cdot 7 \cdot 21)$$

This statement in no way contradicts the relation (5·7·17), according to which $C_p - C_V$ is a constant for an ideal gas. The reason is that when $T \rightarrow 0$ and the system approaches its ground state, quantum mechanical effects become very important. Hence the classical equation of state $pV = \nu RT$ is no longer valid, even if the interactions between the particles in a gas are so small that the gas can be treated as ideal.

* As usual, this low-temperature limit may be understood to be sufficiently high that the nuclear spin orientations are still completely random.

5 · 8 Entropy and internal energy

Consider the temperature T and volume V of a substance as the independent variables. Then one can write its entropy as

$$S = S(T, V)$$

so that

$$dS = \left(\frac{\partial S}{\partial T}\right)_V dT + \left(\frac{\partial S}{\partial V}\right)_T dV \quad (5 \cdot 8 \cdot 1)$$

But the first of the derivatives is simply related to the heat capacity at constant volume, while the second can be reexpressed in terms of a Maxwell relation. Specifically, by (5 · 7 · 1) and (5 · 6 · 4) one has

$$\left(\frac{\partial S}{\partial T}\right)_V = \frac{1}{T} C_v \quad (5 \cdot 8 \cdot 2)$$

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial p}{\partial T}\right)_V \quad (5 \cdot 8 \cdot 3)$$

Hence (5 · 8 · 1) becomes

$$dS = \frac{C_v}{T} dT + \left(\frac{\partial p}{\partial T}\right)_V dV \quad (5 \cdot 8 \cdot 4)$$

Note that the right side of (5 · 8 · 3) can be evaluated if one knows the equation of state. The quantity C_v is in general a function of both T and V . Its dependence on V can, however, also be calculated from the equation of state. Indeed, from its definition,

$$C_v = T \left(\frac{\partial S}{\partial T}\right)_V$$

Differentiation at a fixed temperature T then yields

$$\begin{aligned} \left(\frac{\partial C_v}{\partial V}\right)_T &= \left(\frac{\partial}{\partial V}\right)_T \left[T \left(\frac{\partial S}{\partial T}\right)_V \right] = T \frac{\partial^2 S}{\partial V \partial T} \\ &= T \frac{\partial^2 S}{\partial T \partial V} = T \left(\frac{\partial}{\partial T}\right)_V \left(\frac{\partial S}{\partial V}\right)_T \\ &= T \left(\frac{\partial}{\partial T}\right)_V \left(\frac{\partial p}{\partial T}\right)_V \quad \text{by (5 · 8 · 3)} \end{aligned}$$

Thus

$$\left(\frac{\partial C_v}{\partial V}\right)_T = T \left(\frac{\partial^2 p}{\partial T^2}\right)_V \quad (5 \cdot 8 \cdot 5)$$

and the right side can be evaluated from a knowledge of the equation of state.

We have already pointed out, most recently in Sec. 5 · 6, that all thermodynamic properties of a system can be calculated from a knowledge of its entropy. Let us then ask what experimental knowledge is necessary in order to calculate

the entropy S and thus also all other thermodynamic functions. It is readily seen that what is required is only a knowledge of

1. The heat capacity as a function of T for some *one* fixed value $V = V_1$ of the volume

2. The equation of state

For example, the equation of state can be used in (5·8·5) to calculate $(\partial C_V / \partial V)_T$ as a function of T and V . This information is then sufficient to relate the heat capacity $C_V(T, V)$ at any volume V to the known heat capacity $C_V(T, V_1)$ at the volume V_1 and the same temperature T ; i.e., one has simply

$$C_V(T, V) = C_V(T, V_1) + \int_{V_1}^V \left(\frac{\partial C_V(T, V')}{\partial V'} \right)_T dV' \quad (5\cdot8\cdot6)$$

Knowing $C_V(T, V)$ and using the knowledge of $(\partial p / \partial T)_V$ as a function of T and V which is provided by the equation of state, one can immediately use (5·8·4) to find $S(T, V)$ at any temperature T and volume V compared to its value in some standard state of temperature T_0 and volume V_0 . One needs only to integrate (5·8·4) by writing

$$S(T, V) - S(T_0, V_0) = [S(T, V) - S(T_0, V)] + [S(T_0, V) - S(T_0, V_0)] \quad (5\cdot8\cdot7)$$

where the first term on the right represents the entropy change at the constant volume V and the second term the entropy change at the constant temperature T_0 . Thus one gets

$$S(T, V) - S(T_0, V_0) = \int_{T_0}^T \frac{C_V(T', V)}{T'} dT' + \int_{V_0}^V \left(\frac{\partial p(T_0, V')}{\partial T} \right)_V dV' \quad (5\cdot8\cdot8)$$

Remark One could, of course, have equally well integrated (5·8·4) in opposite order by writing, instead of (5·8·7),

$$S(T, V) - S(T_0, V_0) = [S(T, V) - S(T, V_0)] + [S(T, V_0) - S(T_0, V_0)] \\ = \int_{V_0}^V \left(\frac{\partial p(T, V')}{\partial T} \right)_V dV' + \int_{T_0}^T \frac{C_V(T', V_0)}{T'} dT' \quad (5\cdot8\cdot9)$$

This expression involves C_V at the volume V_0 instead of the volume V , and $(\partial p / \partial T)_V$ at the temperature T instead of the temperature T_0 . Nevertheless (5·8·9) must yield the result (5·8·8). The reason is, of course, the fundamental one that the entropy is a quantity characteristic of a particular macrostate, so that the entropy difference calculated is independent of the process used to go from the macrostate T_0, V_0 to that corresponding to T, V .

Let us now turn to the internal energy E of the substance and consider it as a function of T and V . The fundamental thermodynamic relation asserts that

$$dE = T dS - p dV$$

By expressing dS in terms of T and V as we already did in (5·8·4), this can be written as

$$dE = C_V dT + \left[T \left(\frac{\partial p}{\partial T} \right)_V - p \right] dV \quad (5\cdot8\cdot10)$$

Comparing this with the purely mathematical result

$$dE = \left(\frac{\partial E}{\partial T} \right)_V dT + \left(\frac{\partial E}{\partial V} \right)_T dV$$

we obtain the relations

► $\left(\frac{\partial E}{\partial T} \right)_V = C_V \quad (5\cdot8\cdot11)$

► $\left(\frac{\partial E}{\partial V} \right)_T = T \left(\frac{\partial p}{\partial T} \right)_V - p \quad (5\cdot8\cdot12)$

Equation (5·8·12) shows that the dependence of the internal energy on the volume can again be calculated from the equation of state. A knowledge of this equation of state and of the heat capacity permits one thus to integrate (5·8·10) to find $E(T, V)$ at any temperature T and volume V compared to the energy $E(T_0, V_0)$ of some standard macrostate.

Example: The van der Waals gas Consider a gas whose equation of state is

$$\left(p + \frac{a}{v^2} \right) (v - b) = RT \quad (5\cdot8\cdot13)$$

where $v = V/n$ is the molar volume. This is an empirical equation known as the van der Waals equation. (With suitable approximations it can also be derived from statistical mechanics; see Chapter 10.) It represents the behavior of real gases more accurately than the ideal gas law by introducing two additional positive constants a and b characteristic of the particular gas under consideration. (Indeed, it is approximately valid even at temperatures and molar volumes so low that the gas has become a liquid.)

From a qualitative microscopic point of view, long-range attractive forces between molecules tend to keep them closer together than would be the case for noninteracting molecules. These forces thus have the same effect as a slight compression of the gas; the term a/v^2 represents this additional positive pressure. On the other hand, there are also short-range repulsive forces between the molecules which keep them apart sufficiently to prevent them from occupying the same place at the same time. The term b represents the volume occupied by the molecules themselves and which must thus be subtracted from the volume available to any one molecule in the container.

For $a = b = 0$, or in the limit where the gas becomes very dilute (so that $v \rightarrow \infty$), Eq. (5·8·13) reduces to the ideal gas equation

$$pv = RT$$

as it must.

We first calculate by Eq. (5·8·12) the volume dependence of the molar energy ϵ . We need to find $(\partial p/\partial T)_v$. Solving (5·8·13) for p , one gets

$$p = \frac{RT}{v - b} - \frac{a}{v^2} \quad (5·8·14)$$

$$\text{Hence } \left(\frac{\partial p}{\partial T}\right)_v = \frac{R}{v - b} \quad (5·8·15)$$

Thus (5·8·12) yields

$$\left(\frac{\partial \epsilon}{\partial v}\right)_T = T \left(\frac{\partial p}{\partial T}\right)_v - p = \frac{RT}{v - b} - p$$

or, by (5·8·14),

$$\left(\frac{\partial \epsilon}{\partial v}\right)_T = \frac{a}{v^2} \quad (5·8·16)$$

For an ideal gas, $a = 0$ so that $(\partial \epsilon/\partial v)_T = 0$ in agreement with our earlier result (5·1·10).

Also we have by (5·8·5) and (5·8·15)

$$\left(\frac{\partial c_V}{\partial v}\right)_T = T \left(\frac{\partial^2 p}{\partial T^2}\right) = T \left(\frac{\partial}{\partial T}\right)_v \left(\frac{R}{v - b}\right) = 0$$

Hence c_V is independent of the molar volume and thus only a function of T , i.e.,

$$c_V = c_V(T) \quad (5·8·17)$$

(The same result is, of course, true *a fortiori* for an ideal gas.) Equation (5·8·10) can then be written

$$d\epsilon = c_V(T) dT + \frac{a}{v^2} dv \quad (5·8·18)$$

If some standard macrostate of the gas is chosen to have temperature T_0 and molar volume v_0 , then integration of (5·8·18) gives

$$\epsilon(T, v) - \epsilon(T_0, v_0) = \int_{T_0}^T c_V(T') dT' - a \left(\frac{1}{v} - \frac{1}{v_0} \right)$$

$$\text{or } \epsilon(T, v) = \int_{T_0}^T c_V(T') dT' - \frac{a}{v} + \text{constant} \quad (5·8·19)$$

If c_V is independent of temperature, this becomes simply

$$\epsilon(T, v) = c_V T - \frac{a}{v} + \text{constant} \quad (5·8·20)$$

Note that here ϵ does depend on the molar volume v . As v increases, ϵ also increases. This makes physical sense because the intermolecular separation increases as v increases, and thus the attractive (i.e., negative) potential energy of interaction between the molecules is decreased in magnitude.

Finally, let us compute the entropy per mole of gas. By using (5·8·15), the relation (5·8·4) becomes

$$ds = \frac{c_V(T)}{T} dT + \frac{R}{v - b} dv \quad (5·8·21)$$

Integrating, one gets

$$s(T, v) - s(T_0, v_0) = \int_{T_0}^T \frac{c_V(T') dT'}{T'} + R \ln \left(\frac{v - b}{v_0 - b} \right) \quad (5 \cdot 8 \cdot 22)$$

If c_V is independent of temperature, this can be written

$$s(T, v) = c_V \ln T + R \ln (v - b) + \text{constant} \quad (5 \cdot 8 \cdot 23)$$

FREE EXPANSION AND THROTTLING PROCESSES

5 · 9 Free expansion of a gas

This experiment is one we have mentioned before. Consider a rigid container which is thermally insulated. It is divided into two compartments separated by a valve which is initially closed (see Fig. 5 · 9 · 1). One compartment of volume V_1 contains the gas under investigation, the other compartment is empty. The initial temperature of the system is T_1 . The valve is now opened and the gas is free to expand so as to fill the entire container of volume V_2 . What is the temperature T_2 of the gas after the final equilibrium state has been reached?

Since the system consisting of gas and container is adiabatically insulated, no heat flows into the system; i.e.,

$$Q = 0$$

Furthermore, the system does no work in the process; i.e.,

$$W = 0$$

Thus it follows by the first law that the total energy of the system is conserved; i.e.,

$$\Delta E = 0 \quad (5 \cdot 9 \cdot 1)$$

Assume that the container itself has negligible heat capacity so that the internal energy of the container itself does not change. (This is a condition difficult to realize in practice; we shall come back to this point later.) Then the energy

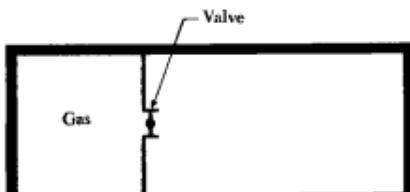


Fig. 5 · 9 · 1 Free expansion of a gas.

change is simply that of the gas, and the conservation of energy (5 · 9 · 1) can be written

$$E(T_2, V_2) = E(T_1, V_1) \quad (5 \cdot 9 \cdot 2)$$

To predict the outcome of the experiment it is only necessary to know the internal energy of the gas $E(T, V)$ as a function of T and V ; for if the initial parameters T_1 and V_1 and the final volume V_2 are known, Eq. (5 · 9 · 2) provides one with an equation for the unknown final temperature T_2 .

Remark The actual free expansion is, of course, a complicated irreversible process involving turbulence and gross nonuniformities of pressure and temperature (to the extent that these quantities can be defined at all for such a marked nonequilibrium situation). Equilibrium conditions prevail only in the initial and final situations. Nevertheless, to predict the outcome of the process, the only knowledge required is that of the energy function E characteristic of equilibrium macrostates of the system.

For an *ideal gas*, E is independent of the volume V ; i.e., $E = E(T)$. Then (5 · 9 · 2) becomes simply $E(T_2) = E(T_1)$, so that one must have $T_2 = T_1$. There is then *no* temperature change in the free expansion of an ideal gas.

More generally, the energy $E(T, V)$ is a function of both T and V . It can be represented in a two-dimensional graph by plotting E versus T for various values of the parameter V , as shown schematically in Fig. 5 · 9 · 2. From such a diagram the result of the experiment can be immediately predicted. Given T_1 and V_1 , one can read off the value $E = E_1$. By (5 · 9 · 2) the intersection of the horizontal line $E = E_1$ with the curve V_2 yields then the final temperature T_2 . If the curves are as drawn, $T_2 < T_1$.

Alternatively, and somewhat more directly, one can use the knowledge of $E(T, V)$ shown in Fig. 5 · 9 · 2 to draw curves of T versus V for various values of the energy E . On such a plot (illustrated in Fig. 5 · 9 · 3) one knows by (5 · 9 · 2) that the initial values of T and V determine a given energy curve, say $E = E_1$, and that as a result of the free expansion one must always end up

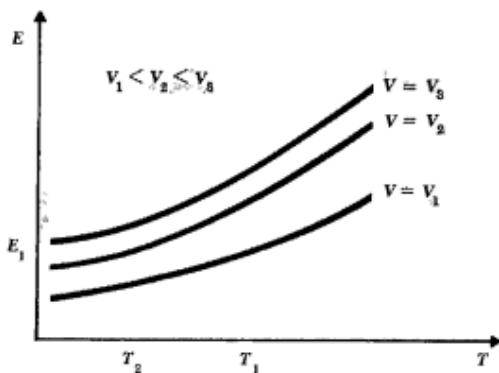
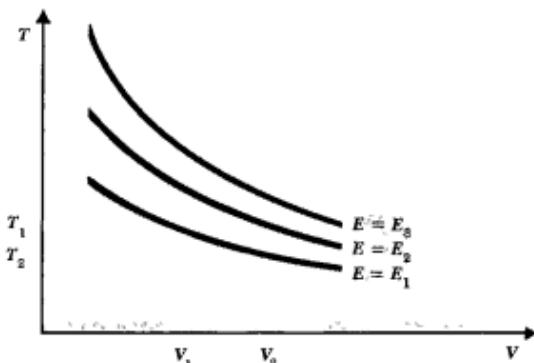


Fig. 5 · 9 · 2 Schematic diagram showing the dependence of the internal energy E of a gas on its temperature T for various values of its volume V .

Fig. 5·9·3 Schematic diagram showing curves of constant internal energy E . Each curve describes corresponding values of T and V which yield the given energy E .



somewhere on this *same* curve. The final temperature can thus be read off immediately from this curve for any value of the final volume V_2 .

Example: van der Waals gas Let us calculate the temperature change in the case of the free expansion of one mole of a van der Waals gas. If we denote the molar internal energy by $\epsilon(T, v)$, the energy-conservation condition (5·9·2) is

$$\epsilon(T_2, v_2) = \epsilon(T_1, v_1)$$

This becomes by (5·8·19)

$$\int_{T_1}^{T_2} c_v(T') dT' - \frac{a}{v_2} = \int_{T_1}^{T_2} c_v(T') dT' - \frac{a}{v_1}$$

Hence $\int_{T_1}^{T_2} c_v(T') dT' - \int_{T_1}^{T_2} c_v(T') dT' = a \left(\frac{1}{v_2} - \frac{1}{v_1} \right)$

or $\int_{T_1}^{T_2} c_v(T') dT' = a \left(\frac{1}{v_2} - \frac{1}{v_1} \right) \quad (5\cdot9\cdot3)$

Over the small temperature range $T_1 < T' < T_2$, any possible temperature dependence of c_v is negligibly small. Thus c_v can be regarded as substantially constant, and (5·9·3) becomes simply

$$c_v(T_2 - T_1) = a \left(\frac{1}{v_2} - \frac{1}{v_1} \right)$$

or $T_2 - T_1 = -\frac{a}{c_v} \left(\frac{1}{v_2} - \frac{1}{v_1} \right) \quad (5\cdot9\cdot4)$

For an expansion where $v_2 > v_1$, or $1/v_1 > 1/v_2$, one gets thus (since $c_v > 0$)

$$T_2 < T_1 \quad (5\cdot9\cdot5)$$

Hence the temperature is reduced as a result of the free expansion.

In principle, it appears that the free expansion of a gas could provide a method of cooling the gas to low temperatures. In practice, a difficulty is encountered because of the appreciable heat capacity C_c of the container. Since its internal energy also changes by an amount $C_c(T_2 - T_1)$, a given volume change of the gas results in a much smaller net temperature change when C_c is finite than when it is zero. (If the container is taken into account, the heat capacity c_v in (5·9·4) must be replaced by the total heat capacity $c_v + C_c$.)

5 · 10 Throttling (or Joule-Thomson) Process

The difficulty associated with the presence of containing walls can be overcome by replacing the single-event free-expansion process just discussed (where this one event must also supply the energy necessary to change the container temperature) with a continuous-flow process (where the temperature of the walls can adjust itself initially and remains unchanged after the steady-state situation has been reached). We now discuss this steady-state experimental arrangement, which was first suggested by Joule and Thomson.

Consider a pipe with thermally insulated walls. A porous plug in the pipe provides a constriction to the flow of the gas. (Alternatively, a valve which is only slightly opened may provide such a constriction.) A continuous stream of gas flows from left to right. The presence of the constriction results in a constant pressure difference being maintained across this constriction. Thus the gas pressure p_1 to the left of the constriction is greater than the gas pressure p_2 to the right of the constriction. Let T_1 denote the temperature of the gas on the left side of the constriction. What then is the gas temperature T_2 on the right side?

Let us analyze the situation. Focus attention on the system consisting of the mass M of gas lying between the dashed planes A and B shown in Fig. 5 · 10 · 2. (We suppose that the planes A and B are chosen so far apart that the volume occupied by the constriction itself is negligible compared to the volume contained between A and B .) At some initial time the plane B coincides with the constriction, and virtually the entire mass M of gas lies to the left of the constriction (see Fig. 5 · 10 · 2a). Then it occupies some volume V_1 corresponding to the pressure p_1 . As this mass M of gas flows down the pipe, the planes A and B which define its geometrical boundaries also move down the pipe. After some time has elapsed, the plane A will have moved so as to coincide with the constriction, and virtually the entire mass M of the gas will lie to the right of the constriction. There it occupies some different volume V_2 corresponding to the lower pressure p_2 . This is the final situation illustrated in Fig. 5 · 10 · 2b.

In the process just described, the difference in internal energy of the mass M of gas between the final situation when it is to the right and the initial

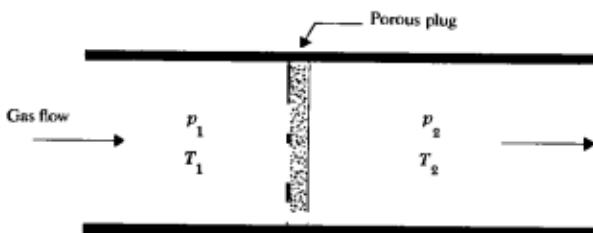


Fig. 5 · 10 · 1 A steady-state throttling process in which a gas is flowing through a porous plug.

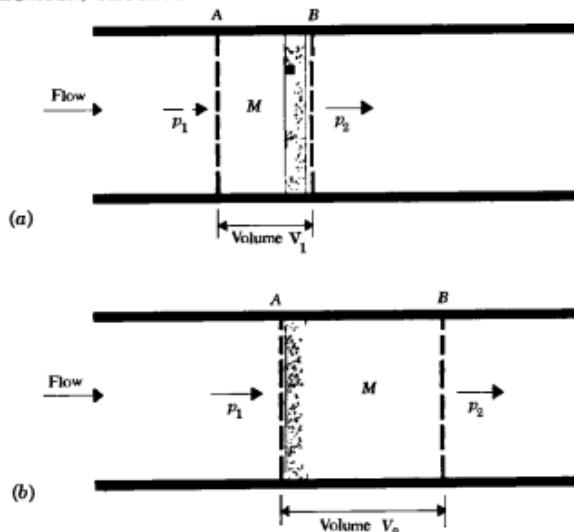


Fig. 5·10·2 Diagram showing a mass M of gas passing through a constriction (a porous plug in this case) (a) before passing through the constriction (b) after passing through the constriction.

situation when it is to the left of the constriction is simply

$$\Delta E = E_2 - E_1 = E(T_2, p_2) - E(T_1, p_1) \quad (5 \cdot 10 \cdot 1)$$

In this process the mass M of gas also does work. Indeed, it does work $p_2 V_2$ in displacing the gas to the right of the constriction by the volume V_2 against the *constant* pressure p_2 . Furthermore, the gas to the left of the constriction does work $p_1 V_1$ on the mass M of gas by displacing it by the volume V_1 with a *constant* pressure p_1 . Hence the *net* work done by the mass M of gas in the process is simply

$$W = p_2 V_2 - p_1 V_1 \quad (5 \cdot 10 \cdot 2)$$

But no heat is absorbed by the mass M of gas in the process we have described. This is not just because the walls are adiabatically insulated so that no heat enters them from the outside; more importantly, after the steady state situation has been established, there is no temperature difference between the walls and the adjacent gas, so that no heat flows from the walls into the gas. Thus

$$Q = 0 \quad (5 \cdot 10 \cdot 3)$$

Application of the first law to the mass M of gas yields then for the process under consideration the relation

$$\Delta E + W = Q = 0 \quad (5 \cdot 10 \cdot 4)$$

By (5 · 10 · 1) and (5 · 10 · 2) this becomes

$$(E_2 - E_1) + (p_2 V_2 - p_1 V_1) = 0 \\ \text{or} \quad E_2 + p_2 V_2 = E_1 + p_1 V_1 \quad (5 \cdot 10 \cdot 5)$$

Let us define the quantity

$$H = E + pV \quad (5 \cdot 10 \cdot 6)$$

This is the so-called "enthalpy" already encountered in (5 · 5 · 7). Then (5 · 10 · 5) can be written

$$H_2 = H_1$$

or

► $H(T_2, p_2) = H(T_1, p_1) \quad (5 \cdot 10 \cdot 7)$

Thus we arrive at the result that in a throttling process the gas passes through the constriction in such a way that its enthalpy H remains constant.

Note that (5 · 10 · 7) is analogous to the condition (5 · 9 · 2) for the free expansion case. The difference is that the gas *does* work in the throttling process, so that the enthalpy rather than the internal energy is the quantity which is conserved.

Remark Here again the actual passage of the gas through the constriction involves complicated irreversible nonequilibrium processes. Equilibrium situations prevail only to the left and to the right of the constriction. But a knowledge of the enthalpy function $H(T, p)$ characteristic of equilibrium macrostates of the system is sufficient to predict the outcome of the process.

Suppose that $H(T, p)$ is known as a function of T and p . Then, given T_1 and p_1 and the final pressure p_2 , (5 · 10 · 7) provides an equation to determine the unknown final temperature T_2 . In the case of an ideal gas,

$$H = E + pV = E(T) + \nu RT$$

so that $H = H(T)$ is a function of the temperature only. Then the condition (5 · 10 · 7) implies immediately

$$H(T_2) = H(T_1)$$

so that $T_2 = T_1$. Thus the temperature of an *ideal* gas does *not* change in a throttling process.

In the more general case the method of analysis is similar to that used for the discussion of the free expansion in Sec. 5 · 9. From a knowledge of $H(T, p)$ one can construct curves of T versus p for various fixed values of the enthalpy H (see Fig. 5 · 10 · 3). On such a plot the initial values T_1 and p_1 determine a particular enthalpy curve. By virtue of (5 · 10 · 7) one must end up somewhere on this *same* curve as a result of the throttling process. The final temperature T_2 can then be read off immediately from this curve for any value of the final pressure p_2 .

The curves of Fig. 5 · 10 · 3 do in general exhibit maxima. Thus it is possible to attain, as a result of a throttling process with $p_2 < p_1$, conditions where

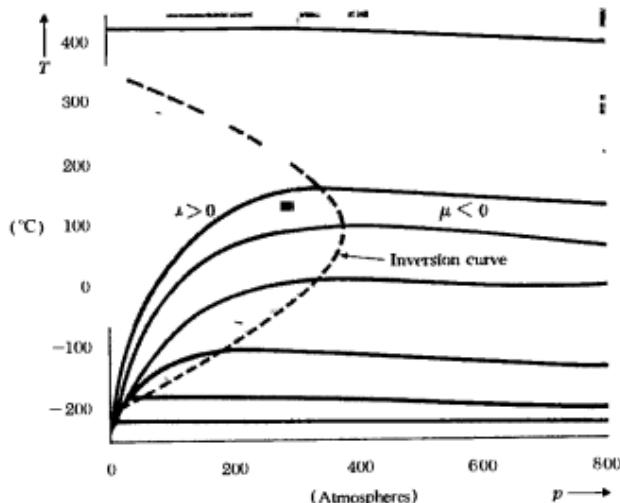


Fig. 5.10-3 Curves of constant enthalpy H in the pT plane of a gas. The numerical values are for nitrogen (N_2). The dashed line is the inversion curve.

the temperature T either increases, decreases, or remains the same. A significant parameter in this context is the slope μ of these curves,

$$\mu = \left(\frac{\partial T}{\partial p} \right)_H \quad (5.10-8)$$

called the Joule-Thomson coefficient.* This quantity gives the change of temperature produced in a throttling process involving an infinitesimal pressure differential. For an infinitesimal pressure drop, T will decrease if $\mu > 0$. The condition $\mu = 0$ implies that no temperature change occurs and locates the maxima of the curves in Fig. 5.10-3. The locus of the maxima forms a curve (shown dashed in Fig. 5.10-3) which is called the "inversion curve." It separates on the diagram the region of positive slope μ (where the temperature tends to fall) from the region of negative slope μ (where the temperature tends to rise).

Let us find an expression for μ in terms of readily measured parameters of the gas. By (5.10-7) we are interested in a situation where H is constant. Starting from the fundamental thermodynamic relation

$$dE = T dS - p dV$$

we get for the enthalpy change the result

$$dH = d(E + pV) = T dS + V dp \quad (5.10-9)$$

(This was already obtained in (5.5-6).) In our case, where H is constant, $dH = 0$. Writing (5.10-9) in terms of the quantities T and p which we have

* It is sometimes also called the Joule-Kelvin coefficient. The reason is that the names Thomson and Kelvin refer to the same person, William Thomson, who was raised to the peerage and thus became Lord Kelvin.

used as independent variables in discussing the throttling process, we get

$$0 = T \left[\left(\frac{\partial S}{\partial T} \right)_p dT + \left(\frac{\partial S}{\partial p} \right)_T dp \right] + V dp$$

or

$$C_p dT + \left[T \left(\frac{\partial S}{\partial p} \right)_T + V \right] dp = 0$$

where we have used $C_p = T(\partial S/\partial T)_p$. Using this result, valid under conditions of constant H , to solve for the ratio dT/dp , we get

$$\mu \equiv \left(\frac{\partial T}{\partial p} \right)_H = - \frac{T(\partial S/\partial p)_T + V}{C_p} \quad (5 \cdot 10 \cdot 10)$$

The numerator can be transformed into more convenient form by a Maxwell relation; by (5 · 6 · 5) one has

$$\left(\frac{\partial S}{\partial p} \right)_T = - \left(\frac{\partial V}{\partial T} \right)_p = -V\alpha$$

where α is the coefficient of expansion defined in (5 · 7 · 8). Thus (5 · 10 · 10) becomes

► $\mu = \frac{V}{C_p} (T\alpha - 1) \quad (5 \cdot 10 \cdot 11)$

Of course, this is properly an intensive quantity, since both the volume V and the heat capacity C_p are extensive quantities.

For an ideal gas we found in (5 · 7 · 15) that $\alpha = T^{-1}$. Then $\mu = 0$ and, as mentioned previously, no temperature change results from a throttling process.

More generally, $\mu > 0$ if $\alpha > T^{-1}$, and conversely $\mu < 0$ if $\alpha < T^{-1}$. The locus of points in the pT plane, where α is such that $\alpha = T^{-1}$, gives the inversion curve.

The Joule-Thomson effect constitutes a practical method for cooling gases and is often used in processes designed to liquefy gases. In order to achieve a lower temperature as a result of throttling a gas, it is necessary to work in that region of pressure and temperature where $\mu > 0$; in particular, the initial temperature must be less than the maximum temperature on the inversion curve (see Fig. 5 · 10 · 3). For example, this maximum inversion temperature is 34°K for helium, 202°K for hydrogen, and 625°K for nitrogen. An attempt to throttle helium gas starting from room temperature would thus result in an increase, rather than a decrease, of the gas temperature. To use the Joule-Thomson effect for cooling helium gas to very low temperatures it is then necessary to precool it first to temperatures below 34°K. This can be done by using liquid hydrogen to precool the helium gas. Alternatively, one can keep the helium gas thermally insulated and let it do mechanical work at the expense of its internal energy. After this procedure has resulted in a sufficiently large temperature decrease of the helium gas, the Joule-Thomson effect can be used as the final stage in the cooling process.

Joule-Thomson effect and molecular forces Neither free-expansion nor throttling processes result in a temperature change in the case of an ideal gas.

Both of these processes become interesting only if the gas is not ideal, i.e., when the mutual interaction between molecules is of importance. The equation of state of any gas can be written in the general form of a series

$$p = kT[n + B_2(T)n^2 + B_3(T)n^3 \dots] \quad (5 \cdot 10 \cdot 12)$$

which is an expansion in powers of the number of molecules per unit volume $n \equiv N/V$. The expression (5·10·12) is called the "virial expansion," and the coefficients B_2, B_3, \dots are called virial coefficients. For an ideal gas $B_2 = B_3 = \dots = 0$. If n is not too large, only the first few terms in (5·10·12) are important. The first correction to the ideal gas consists of retaining the term B_2n^2 and neglecting all higher-order terms. In this case (5·10·12) becomes

$$p = \frac{N}{V} kT \left(1 + \frac{N}{V} B_2 \right) \quad (5 \cdot 10 \cdot 13)$$

One can readily make some qualitative statements about the behavior of B_2 as a function of T on the basis of some simple microscopic considerations. The interaction between two gas molecules is weakly attractive when their mutual separation is relatively large, but becomes strongly repulsive when their separation becomes of the order of a molecular diameter.* At low temperatures the mean kinetic energy of a molecule is small. The weak long-range attraction between molecules is then quite significant and tends to make the mean intermolecular separation less than would be the case in the absence of interaction. This attraction thus tends to reduce the gas pressure below that for an ideal gas; i.e., B_2 in (5·10·13) is then negative. But at higher temperatures the mean kinetic energy of a molecule becomes so large that the weak intermolecular attractive potential energy becomes comparatively negligible. In that case it is the strong short-range repulsive interaction between molecules which is most significant. This repulsion tends to increase the gas pressure above that expected for an ideal gas; i.e., B_2 is then positive. These qualitative considerations lead one to expect that B_2 is an increasing function of T , being negative for sufficiently low temperatures and becoming positive at higher temperatures. (These arguments will be made more quantitative in Sec. 10·4; they lead to a curve of $B_2(T)$ versus T of the type shown in Fig. 10·4·1.)

Let us now apply these considerations to a discussion of the Joule-Thomson effect by evaluating (5·10·11). Let us use the equation of state (5·10·13) to express V as a function of T and p . This is most readily done by noting that the term $(N/V)B_2$ is a correction term which is small compared to unity; hence one commits negligible error by replacing the ratio N/V in that term by the value $p/(kT)$ which this ratio assumes in first approximation. Thus (5·10·13) becomes

$$\begin{aligned} p &= \frac{NkT}{V} \left(1 + \frac{p}{kT} B_2 \right) = \frac{N}{V} (kT + pB_2) \\ \text{or } V &= N \left(\frac{kT}{p} + B_2 \right) \end{aligned} \quad (5 \cdot 10 \cdot 14)$$

* Figure 10·3·1 illustrates a curve of potential energy of mutual interaction as a function of intermolecular distance.

Hence (5 · 10 · 11) yields the result

$$\mu = \frac{1}{C_p} \left[T \left(\frac{\partial V}{\partial T} \right)_p - V \right] = \frac{N}{C_p} \left(T \frac{\partial B_2}{\partial T} - B_2 \right) \quad (5 \cdot 10 \cdot 15)$$

The previously discussed temperature dependence of B_2 allows us now to draw some interesting conclusions. Since B_2 is an increasing function of T , the term $T(\partial B_2 / \partial T)$ is positive. At low temperatures where molecular attraction is predominant, B_2 itself is negative; hence (5 · 10 · 15) shows that $\mu > 0$ in this temperature range. But if one goes to sufficiently high temperatures where molecular repulsion becomes predominant, B_2 becomes positive and sufficiently large to make $\mu < 0$ in (5 · 10 · 15). The existence of the inversion curve where $\mu = 0$ reflects, therefore, the competing effects between molecular attraction and repulsion.

HEAT ENGINES AND REFRIGERATORS

5 · 11 Heat engines

Historically, the subject of thermodynamics began with a study of the basic properties of heat engines. Since the subject has not only great technological importance (sufficient to have been responsible for the industrial revolution) but also intrinsic physical interest, we shall devote some time to a discussion of it.

It is very easy to do mechanical work w upon a device M , and then to extract from it an equivalent amount of heat q , which goes to increase the internal energy of some heat reservoir B .* For example, the device M might be a paddle wheel rotated in a liquid by a falling weight, or an electric resistor upon which electrical work is done.

The fundamentally significant question is: To what extent is it possible to proceed in the reverse way, i.e., to build a device (called a "heat engine") which can extract some of the internal energy from a heat reservoir in the form of heat and convert it into macroscopic work? The situation would then be as diagrammed in Fig. 5 · 11 · 3.

It is necessary to keep in mind the following points. The work should not be provided at the expense of the heat engine itself; otherwise one could not continue the process of heat-to-work conversion indefinitely. Thus one wants the engine to be in the same macrostate at the end of the process as it was at the beginning (i.e., to have gone through a cycle) so that it is ready to start again converting more heat into work in the next cycle. Clearly, steam engines and gasoline engines all go through such cycles. Furthermore, the work put out by the heat engine should be capable of simply changing an external parameter of some outside device (e.g., of lifting a weight) without doing it at the expense of affecting the other degrees of freedom (or entropy) of that device. One can

* We use the small letters w and q to denote intrinsically *positive* amounts of work and heat.

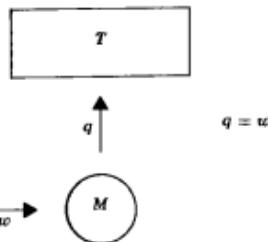


Fig. 5·11·1 Conversion of work w into heat q given off to a heat reservoir at temperature T .

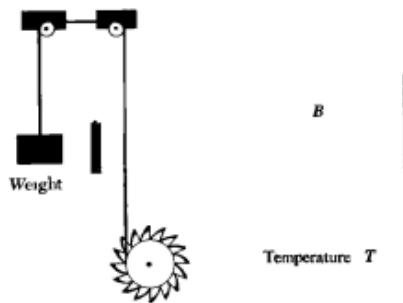


Fig. 5·11·2 A physical illustration showing the conversion of mechanical work into heat.

thus phrase the essential problem of constructing a heat engine in the following way: To what extent is it possible to extract a net amount of energy from one (or more) heat reservoirs, where that energy is randomly distributed over very many degrees of freedom, so as to transform it into energy associated with the single degree of freedom connected with the external parameter of an outside device?

Figure 5·11·3 would then be the prototype of the most desirable type of engine. After a cycle, M is back in the same macrostate as at the beginning, so that its internal energy is the same. Hence the first law of thermodynamics implies that

$$w = q \quad (5\cdot11\cdot1)$$

i.e., to conserve energy, the work put out by the engine must be equal to the heat extracted from the reservoir. One can certainly *not* build an engine which violates this condition.

But one may not be able to construct an engine even when this condition

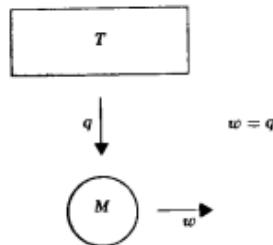


Fig. 5·11·3 A perfect engine.

is satisfied. Thus the engine illustrated in Fig. 5·11·3 is indeed a highly desirable “perfect engine”; i.e., working in a cycle, it extracts heat from a reservoir and performs an equivalent amount of work *without* producing any other effect on the environment. But a perfect engine of this kind is, unfortunately, not realizable. Indeed we know from our discussion of Sec. 3·2 that the conversion of work into heat illustrated in Fig. 5·11·2, or more schematically in Fig. 5·11·1, is an irreversible process in which the distribution of systems over accessible states becomes more random, so that the entropy increases. One cannot, therefore, simply reverse the process as shown in Fig. 5·11·3. In the concrete case of Fig. 5·11·2, one cannot simply expect the heat reservoir *B* to convert its internal energy, randomly distributed over all its degrees of freedom, into a systematic upward motion of the weight. It is, of course, in principle *possible* that this might happen, but from a statistical point of view such an occurrence is *fantastically improbable*.

Equivalently, we can show that an ideal engine of the type illustrated in Fig. 5·11·3 violates the second law of thermodynamics. Indeed, we must require that the total entropy change ΔS of the complete system (consisting of the heat engine, the outside device on which it does work, and the heat reservoir) be such that in a cycle

$$\Delta S \geq 0 \quad (5 \cdot 11 \cdot 2)$$

Now the engine itself returns to its previous state after a cycle; its entropy is thus unchanged after the cycle. Furthermore, we have already pointed out that no entropy change is associated with the outside device on which work is done. On the other hand, the entropy change of the heat reservoir at absolute temperature T_1 is, by (3·6·4), given by $-q/T_1$, since the reservoir absorbs heat ($-q$). Hence (5·11·2) becomes

$$\frac{-q}{T_1} \geq 0$$

or, by (5·11·1),

$$\frac{q}{T_1} = \frac{w}{T_1} \leq 0 \quad (5 \cdot 11 \cdot 3)$$

Since we want the work w done by the engine to be positive, (5·11·3) cannot be satisfied. The inverse process of Fig. 5·11·1 where $w < 0$ is, of course, feasible. There is no objection to converting any amount of work into heat, but the converse is not possible. The second law thus again implies a fundamental irreversibility of natural processes. By (5·11·3) it specifically implies this result:

It is impossible to construct a perfect heat engine. (5 · 11 · 4)

(This statement is sometimes known as Kelvin's formulation of the second law of thermodynamics.)

A perfect heat engine is thus not realizable because it would require the spontaneous occurrence of a process which goes from an initial situation, where a certain amount of energy is distributed randomly over the many degrees of

freedom of a heat reservoir, to a much more special and enormously less probable final situation, where this energy is all associated with the motion of a single degree of freedom capable of performing macroscopic work; in short, because it would require a process where the entropy S decreases. But this kind of process, in which the system consisting of heat reservoir and engine goes to a less random situation, *can* take place if this system is coupled to some other auxiliary system whose degree of randomness (or entropy) is increased in this process by a compensating amount (i.e., by an amount large enough that the *entire* system does increase in randomness). The simplest such auxiliary system is a second heat reservoir at some temperature T_2 lower than T_1 . One then obtains a nonperfect, but realizable, heat engine which not only absorbs heat q_1 from a reservoir at temperature T_1 , but also *rejects* heat to some second reservoir at some lower temperature T_2 . Thus a real engine can be diagrammed as shown in Fig. 5·11·4.

In this case the first law requires that in a cycle

$$q_1 = w + q_2 \quad (5 \cdot 11 \cdot 5)$$

On the other hand, the second law is satisfied if in this cycle the total entropy change of these reservoirs satisfies the inequality

$$\Delta S = \frac{(-q_1)}{T_1} + \frac{q_2}{T_2} \geq 0 \quad (5 \cdot 11 \cdot 6)$$

Equations (5·11·5) and (5·11·6) *can* be satisfied with positive work w performed by the engine on the outside world. By combining these equations one gets

$$\begin{aligned} \frac{-q_1}{T_1} + \frac{q_1 - w}{T_2} &\geq 0 \\ \frac{w}{T_2} &\leq q_1 \left(\frac{1}{T_2} - \frac{1}{T_1} \right) \\ \text{or } \eta = \frac{w}{q_1} &\leq 1 - \frac{T_2}{T_1} = \frac{T_1 - T_2}{T_1} \end{aligned} \quad (5 \cdot 11 \cdot 7)$$

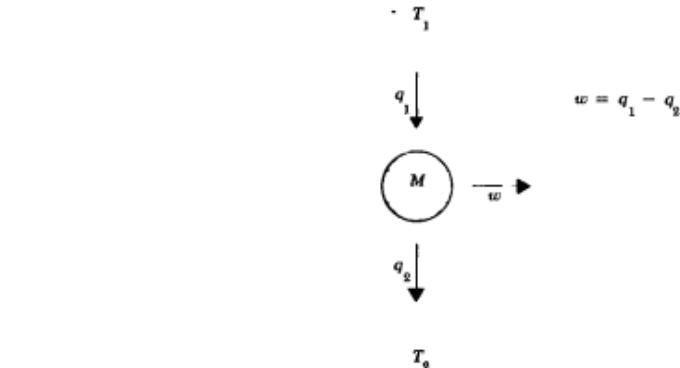


Fig. 5·11·4 A real engine.

For a perfect engine one would have $w/q_1 = 1$. For a real engine this ratio is less than 1, i.e.,

$$\eta \equiv \frac{w}{q_1} = \frac{q_1 - q_2}{q_1} < 1 \quad (5 \cdot 11 \cdot 8)$$

since some heat does not get transformed into work but is instead rejected to some other heat reservoir. The quantity $\eta = w/q_1$ is called the "efficiency" of the engine. Equation (5·11·7) provides us then with a relation for the maximum possible efficiency of an engine operating between two reservoirs of given absolute temperatures. Since the equals sign in the second statement (5·11·6) holds only for a quasi-static process, (5·11·7) implies also that no engine operating between the two given heat reservoirs can have an efficiency greater than that of an engine which operates between the same two reservoirs in a quasi-static manner. Furthermore, (5·11·7) implies that *any* engine which operates between these two reservoirs in a quasi-static manner has the *same* efficiency:

if quasi-static,

$$\eta = \frac{T_1 - T_2}{T_1} \quad (5 \cdot 11 \cdot 9)$$

Carnot engines It is of interest to exhibit explicitly how such an engine operating quasi-statically between two heat reservoirs can be constructed. Such an engine is the simplest conceivable engine and is called a "Carnot engine" (named after Carnot, the French engineer who was the first to examine theoretically the operation of heat engines). Let x denote the external parameter of the engine M ; changes in this parameter give rise to the work performed by the engine. Let the engine initially be in a state where $x = x_a$ and its temperature $T = T_2$, the temperature of the colder heat reservoir. The Carnot engine then goes through a cycle consisting of four steps, all per-

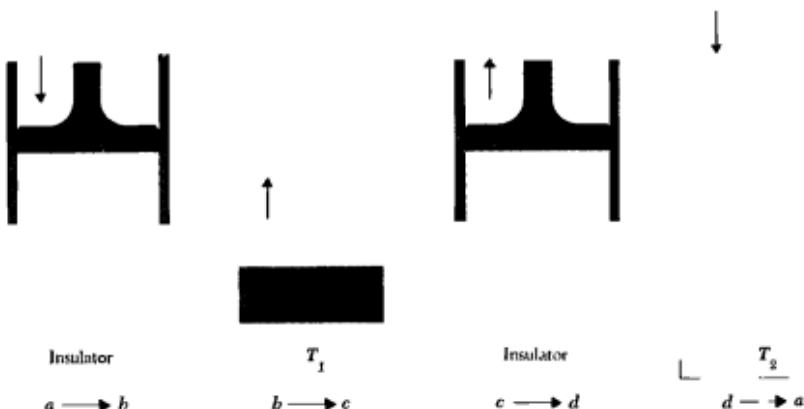


Fig. 5·11·5 The four steps of a Carnot cycle in which a gas is used as the working substance. The external parameter x is the volume V of the gas.

formed in a quasi-static fashion. Label macrostates of the engine by small letters a, b, c, d .

1. $a \rightarrow b$: The engine is *thermally insulated*. Its external parameter is changed slowly until the engine temperature reaches T_1 . Thus $x_a \rightarrow x_b$ such that $T_2 \rightarrow T_1$.

2. $b \rightarrow c$: The engine is now placed in *thermal contact* with the heat reservoir at temperature T_1 . Its external parameter is changed further, the engine remaining at temperature T_1 and absorbing some heat q_1 from the reservoir. Thus $x_b \rightarrow x_c$ such that heat q_1 is absorbed by the engine.

3. $c \rightarrow d$: The engine is again *thermally insulated*. Its external parameter is changed in such a direction that its temperature goes back to T_2 . Thus $x_c \rightarrow x_d$ such that $T_1 \rightarrow T_2$.

4. $d \rightarrow a$: The engine is now placed in *thermal contact* with the heat reservoir at temperature T_2 . Its external parameter is then changed until it returns to its initial value x_a , the engine remaining at temperature T_2 and rejecting some heat q_2 to this reservoir. Thus $x_d \rightarrow x_a$ and heat q_2 is given off by the engine.

The engine is now back in its initial state and the cycle is completed.

Example Let us illustrate a Carnot cycle with a particular kind of system. Take, for example, a gas (not necessarily ideal) contained in a cylinder closed off by a piston. The external parameter is the volume V of the gas. The four steps of the Carnot cycle are illustrated in Fig. 5·11·5. The area enclosed by the quadrilateral figure in Fig. 5·11·6 represents the total work

$$w = \int_a^b p \, dV + \int_b^c p \, dV + \int_c^d p \, dV + \int_d^a p \, dV$$

performed by the engine in a cycle.

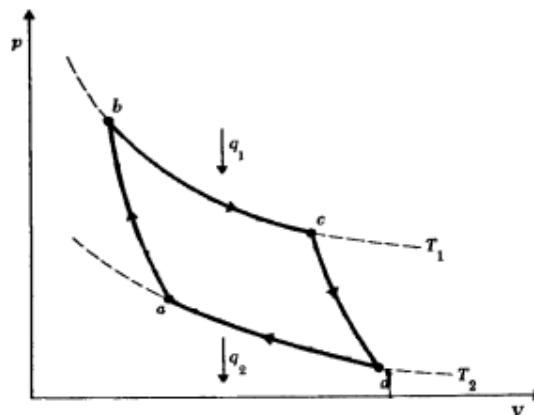


Fig. 5·11·6 The Carnot cycle of Fig. 5·11·5 illustrated on a pV diagram.

Practical engines, such as steam or gasoline engines, are somewhat more complicated than Carnot engines. But, like all heat engines, they cannot be perfect and are well known to have mechanisms (such as condensers or exhausts) by which they reject heat to some low-temperature reservoir, usually the surrounding atmosphere.

5 · 12 Refrigerators

A refrigerator is a device which, operating in a cycle, removes heat from a reservoir at lower absolute temperature and rejects it to a reservoir at higher absolute temperature. It can be represented by the diagram of Fig. 5 · 12 · 1, which is similar to that of Fig. 5 · 11 · 4 except that the directions of all the arrows have been reversed. The first law, applied to the refrigerator of Fig. 5 · 12 · 1, requires that

$$w + q_2 = q_1 \quad (5 \cdot 12 \cdot 1)$$

Since a Carnot engine operates quasi-statically by passing continuously through a series of near-equilibrium states, one could run it equally well quasi-statically in the reverse direction. In this case it would operate like a particularly simple special kind of refrigerator.

Needless to say, Fig. (5 · 12 · 1) does not represent the ideal refrigerator one might like to have. After all, if the two heat reservoirs were placed in thermal contact with each other, some amount of heat q would flow spontaneously from the reservoir at higher temperature T_1 to the reservoir at lower temperature T_2 . The "perfect refrigerator" would just reverse the process, i.e., it would simply remove heat q from the reservoir at lower temperature and transfer it to the reservoir at higher temperature *without* affecting the environment in any other way; i.e., the perfect refrigerator would not require any work

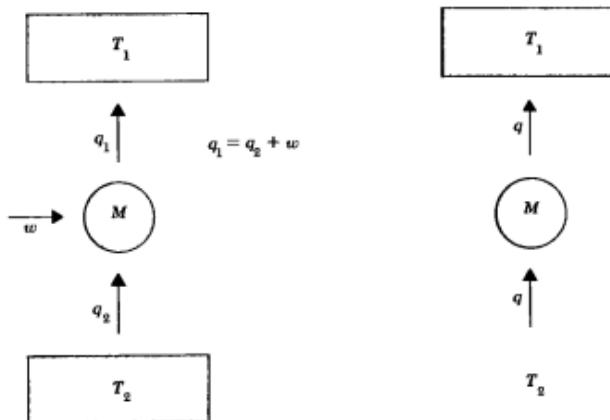


Fig. 5 · 12 · 1 A real refrigerator.

Fig. 5 · 12 · 2 A perfect refrigerator.

to be done on it and would be represented by the diagram of Fig. 5·12·2. But a perfect refrigerator would again violate the second law. Indeed, the total entropy change in Fig. (5·12·2) has to satisfy the inequality

$$\Delta S = \frac{q}{T_1} + \frac{(-q)}{T_2} \geq 0$$

or $q \left(\frac{1}{T_1} - \frac{1}{T_2} \right) \geq 0 \quad (5 \cdot 12 \cdot 2)$

which is impossible for $q > 0$ and $T_1 > T_2$. Thus we arrive at the following statement:

It is impossible to construct a perfect refrigerator. (5·12·3)

(This statement is sometimes known as the Clausius formulation of the second law of thermodynamics.)

This result is, of course, only too familiar. Kitchen refrigerators have a nasty habit of requiring an external source of power.

A real refrigerator is then properly represented by Fig. 5·12·1, where some amount of work w must be done on the refrigerator to make it function. In that case one has, by (5·12·1),

$$q_2 = q_1 - w \quad (5 \cdot 12 \cdot 4)$$

i.e., the heat removed from the colder reservoir is *less* than that given off to the warmer reservoir. The second law imposes then the requirement that

$$\Delta S = \frac{q_1}{T_1} + \frac{(-q_2)}{T_2} \geq 0$$

or $\frac{q_2}{q_1} \leq \frac{T_2}{T_1} \quad (5 \cdot 12 \cdot 5)$

where the equals sign holds only for a refrigerator operating between the two reservoirs in a quasi-static manner.

Remark It can be shown that the Kelvin and Clausius statements of the second law are equivalent and that either one implies that there must exist a function with the properties of entropy. This was the basis of the historical macroscopic approach to classical thermodynamics. The interested reader is referred to the bibliography at the end of this chapter for books developing this point of view.

SUGGESTIONS FOR SUPPLEMENTARY READING

Thermodynamic relations and properties of pure substances

- M. W. Zemansky: "Heat and Thermodynamics," 4th ed., chaps. 6, 11, and 13, McGraw-Hill Book Company, New York, 1957.
 H. B. Callen: "Thermodynamics," chaps. 5 and 6, John Wiley & Sons, Inc., New York, 1960.

Practical engines and refrigerators

- M. W. Zemansky: "Heat and Thermodynamics," 4th ed., secs. 7.2–7.4, 12.1–12.10, McGraw-Hill Book Company, New York, 1957.
- J. K. Roberts and A. R. Miller: "Heat and Thermodynamics," 5th ed., chap. 14, Interscience Publishers, New York, 1960.

Liquefaction of gases

- M. W. Zemansky: "Heat and Thermodynamics," 4th ed., secs. 14.1–14.2, 16.1, McGraw-Hill Book Company, New York, 1957.
- J. K. Roberts and A. R. Miller: "Heat and Thermodynamics," 5th ed., pp. 130–139, Interscience Publishers, New York, 1960.

Completely macroscopic development of thermodynamics based on the Kelvin or Clausius statements of the second law; discussion of Carnot engines

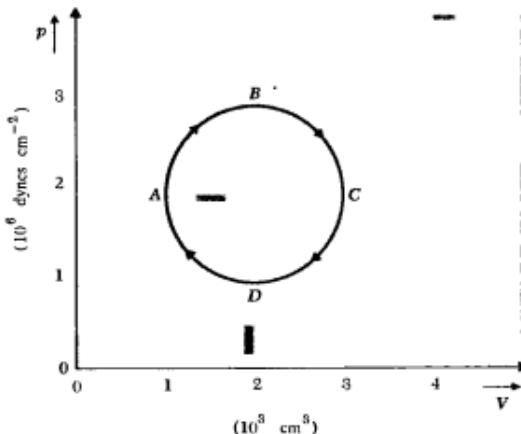
- M. W. Zemansky: "Heat and Thermodynamics," 4th ed., chaps. 7, 9, and 10, McGraw-Hill Book Company, New York, 1957.
- E. Fermi: "Thermodynamics," secs. 7–13, Dover Publications, New York, 1957.

Jacobian method of manipulating partial derivatives

- L. D. Landau and E. M. Lifshitz: "Statistical Physics," sec. 16, Addison-Wesley Publishing Company, Reading, Mass., 1958.
- F. W. Crawford: "Heat, Thermodynamics, and Statistical Physics," secs. 11.13–11.17, Harcourt, Brace and World, New York, 1963.

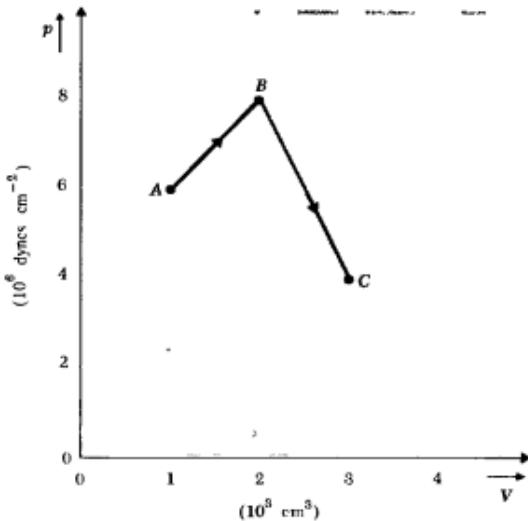
PROBLEMS

- 5.1** An ideal gas has a temperature-independent molar specific heat c_V at constant volume. Let $\gamma \equiv c_p/c_V$ denote the ratio of its specific heats. The gas is thermally insulated and is allowed to expand quasi-statically from an initial volume V_i at temperature T_i to a final volume V_f .
- Use the relation $pV^\gamma = \text{constant}$ to find the final temperature T_f of this gas.
 - Use the fact that the entropy remains constant in this process to find the final temperature T_f .
- 5.2** The molar specific heat at constant volume of a monatomic ideal gas is known to be $\frac{3}{2}R$. Suppose that one mole of such a gas is subjected to a cyclic quasi-static process which appears as a circle on the diagram of pressure p versus volume V shown in the figure. Find the following quantities:
- The net work (in joules) done by the gas in one cycle.
 - The internal energy difference (in joules) of the gas between state C and state A .
 - The heat absorbed (in joules) by the gas in going from A to C via the path ABC of the cycle.



- 5.3 An ideal diatomic gas has a molar internal energy equal to $E = \frac{5}{2}RT$ which depends only on its absolute temperature T . A mole of this gas is taken quasi-statically first from state A to state B , and then from state B to state C along the straight line paths shown in the diagram of pressure p versus volume V .

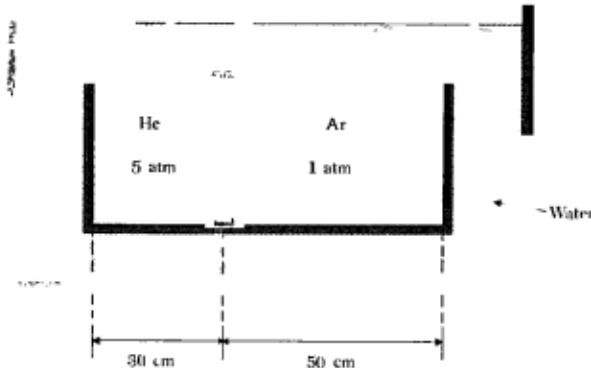
- What is the molar heat capacity at constant volume of this gas?
- What is the work done by the gas in the process $A \rightarrow B \rightarrow C$?
- What is the heat absorbed by the gas in this process?
- What is its change of entropy in this process?



- 5.4 A cylindrical container 80 cm long is separated into two compartments by a thin piston, originally clamped in position 30 cm from the left end. The left compartment is filled with one mole of helium gas at a pressure of 5 atmospheres; the right compartment is filled with argon gas at 1 atmosphere of pressure.

These gases may be considered ideal. The cylinder is submerged in 1 liter of water, and the entire system is initially at the uniform temperature of 25°C. The heat capacities of the cylinder and piston may be neglected. When the piston is unclamped, a new equilibrium situation is ultimately reached with the piston in a new position.

- What is the increase in temperature of the water?
- How far from the left end of the cylinder will the piston come to rest?
- What is the increase of total entropy of the system?



- 5.5 A vertical cylinder contains ν moles of an ideal gas and is closed off by a piston of mass M and area A . The acceleration due to gravity is g . The molar specific heat c_V (at constant volume) of the gas is a constant independent of temperature. The heat capacities of the piston and cylinder are negligibly small and any frictional forces between the piston and the cylinder walls can be neglected. The whole system is *thermally insulated*. Initially, the piston is clamped in position so that the gas has a volume V_0 and a temperature T_0 . The piston is now released and, after some oscillations, comes to rest in a final equilibrium situation corresponding to a larger volume of the gas.
- Does the temperature of the gas increase, decrease, or remain the same?
 - Does the entropy of the gas increase, decrease, or remain the same?
 - Calculate the final temperature of the gas in terms of T_0 , V_0 , and the other parameters mentioned in the statement of the problem.
- 5.6 The following describes a method used to measure the specific heat ratio $\gamma = c_p/c_V$ of a gas. The gas, assumed ideal, is confined within a vertical cylindrical container and supports a freely moving piston of mass m . The piston and cylinder both have the same cross-sectional area A . Atmospheric pressure is p_0 , and when the piston is in equilibrium under the influence of gravity (acceleration g) and the gas pressure, the volume of the gas is V_0 . The piston is now displaced slightly from its equilibrium position and is found to oscillate about this position with frequency ν . The oscillations of the piston are slow enough that the gas always remains in internal equilibrium, but fast enough that the gas cannot exchange heat with the outside. The variations in gas pressure and volume are thus adiabatic. Express γ in terms of m , g , A , p_0 , V_0 , and ν .
- 5.7 Consider the earth's atmosphere as an ideal gas of molecular weight μ in a uniform gravitational field. Let g denote the acceleration due to gravity.

(a) If z denotes the height above sea level, show that the change of atmospheric pressure p with height is given by

$$\frac{dp}{p} = -\frac{\mu g}{RT} dz$$

where T is the absolute temperature at the height z .

(b) If the decrease of pressure in (a) is due to an adiabatic expansion, show that

$$\frac{dp}{p} = \frac{\gamma}{\gamma - 1} \frac{dT}{T}$$

(c) From (a) and (b) calculate dT/dz in degrees per kilometer. Assume the atmosphere to consist mostly of nitrogen (N_2) gas for which $\gamma = 1.4$.

(d) In an isothermal atmosphere at temperature T , express the pressure p at height z in terms of the pressure p_0 at sea level.

(e) If the sea-level pressure and temperature are p_0 and T_0 , respectively, and the atmosphere is regarded as adiabatic as in part (b), find again the pressure p at height z .

5.8 When a sound wave passes through a fluid (liquid or gas), the period of vibration is short compared to the relaxation time necessary for a macroscopically small element of volume of the fluid to exchange energy with the rest of the fluid through heat flow. Hence compressions of such an element of volume can be considered adiabatic.

By analyzing one-dimensional compressions and rarefactions of the system of fluid contained in a slab of thickness dx , show that the pressure $p(x,t)$ in the fluid depends on the position x and the time t so as to satisfy the wave equation

$$\frac{\partial^2 p}{\partial t^2} = u^2 \frac{\partial^2 p}{\partial x^2}$$

where the velocity of sound propagation u is a constant given by $u = (\rho \kappa_s)^{-\frac{1}{2}}$. Here ρ is the equilibrium density of the fluid and κ_s is its *adiabatic compressibility* $\kappa_s = -V^{-1}(\partial V/\partial p)_s$, i.e., its compressibility measured under conditions where the fluid is thermally insulated.

5.9 Refer to the results of the preceding problem.

(a) Calculate the adiabatic compressibility κ_s of an ideal gas in terms of its pressure p and specific heat ratio γ .

(b) Find an expression for the velocity of sound in an ideal gas in terms of γ , its molecular weight μ , and its absolute temperature T .

(c) How does the sound velocity depend on the gas temperature T at a fixed pressure? How does it depend on the gas pressure p at a fixed temperature?

(d) Calculate the velocity of sound in nitrogen (N_2) gas at room temperature and pressure. Take $\gamma = 1.4$.

5.10 Liquid mercury at atmospheric pressure and 0°C (i.e., 273°K) has a molar volume of $14.72 \text{ cm}^3/\text{mole}$ and a specific heat at constant pressure of $c_p = 28.0 \text{ joules mole}^{-1} \text{ deg}^{-1}$. Its coefficient of expansion is $\alpha = 1.81 \times 10^{-4} \text{ deg}^{-1}$, and its compressibility is $\kappa = 3.88 \times 10^{-12} \text{ cm}^2 \text{ dyne}^{-1}$. Find its specific heat c_v at constant volume and the ratio $\gamma \equiv c_p/c_v$.

5.11 Consider an isotropic solid of length L . Its coefficient of linear expansion α_L is defined as $\alpha_L = L^{-1}(\partial L/\partial T)_p$, and is a measure of the change in length of this solid produced by a small change of temperature. By considering an infini-

- tesimal rectangular parallelepiped of this solid, show that the coefficient of volume expansion $\alpha \equiv V^{-1}(\partial V/\partial T)_p$, for this solid is simply equal to $\alpha = 3\alpha_L$.
- 5.12** The following problem arises when experiments are done on solids at high pressures. If the pressure is increased by an amount Δp , this being done under circumstances where the sample is thermally insulated and at a sufficiently slow rate that the process can be regarded as quasi-static, what is the resulting change of temperature ΔT of the sample? If Δp is fairly small, derive an expression for ΔT in terms of Δp , the absolute temperature T of the sample, its specific heat at constant pressure c_p (in ergs g^{-1} deg^{-1}), its density ρ (in g/cm^3), and its volume coefficient of thermal expansion α (in deg^{-1}).
- 5.13** A homogeneous substance at temperature T and pressure p has a molar volume v and a molar specific heat (measured at constant pressure) given by c_p . Its coefficient of volume expansion α is known as a function of temperature. Calculate how c_p depends on the pressure at a given temperature; i.e., calculate $(\partial c_p/\partial p)_T$, expressing the result in terms of T , v , and the properties of α .
- 5.14** In a temperature range near absolute temperature T , the tension force F of a stretched plastic rod is related to its length L by the expression

$$F = aT^2(L - L_0)$$

where a and L_0 are positive constants, L_0 being the unstretched length of the rod. When $L = L_0$, the heat capacity C_L of the rod (measured at constant length) is given by the relation $C_L = bT$, where b is a constant.

(a) Write down the fundamental thermodynamic relation for this system, expressing dS in terms of dE and dL .

(b) The entropy $S(T,L)$ of the rod is a function of T and L . Compute $(\partial S/\partial L)_T$.

(c) Knowing $S(T_0, L_0)$, find $S(T, L)$ at any other temperature T and length L . (It is most convenient to calculate first the change of entropy with temperature at the length L_0 where the heat capacity is known.)

(d) If one starts at $T = T_i$ and $L = L_i$ and stretches the thermally insulated rod quasi-statically until it attains the length L_f , what is the final temperature T_f ? Is T_f larger or smaller than T_i ?

(e) Calculate the heat capacity $C_L(L, T)$ of the rod when its length is L instead of L_0 .

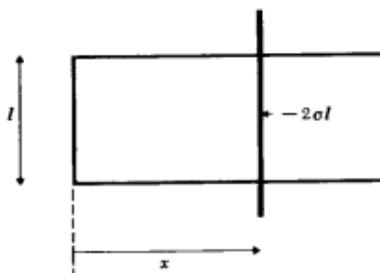
(f) Calculate $S(T, L)$ by writing $S(T, L) - S(T_0, L_0) = [S(T, L) - S(T_0, L)] + [S(T_0, L) - S(T_0, L_0)]$ and using the result of part (e) to compute the first term in the square brackets. Show that the final answer agrees with that found in (c).

- 5.15** The figure illustrates a soap film (shown in gray) supported by a wire frame. Because of surface tension the film exerts a force $2\sigma l$ on the cross wire. This force is in such a direction that it tends to move this wire so as to decrease the area of the film. The quantity σ is called the "surface tension" of the film and the factor 2 occurs because the film has two surfaces. The temperature dependence of σ is given by

$$\sigma = \sigma_0 - \alpha T$$

where σ_0 and α are constants independent of T or x .

(a) Suppose that the distance x (or equivalently, the total film area $2lx$) is the only external parameter of significance in the problem. Write a relation expressing the change dE in mean energy of the film in terms of the heat dQ



absorbed by it and the work done by it in an infinitesimal quasi-static process in which the distance x is changed by an amount dx .

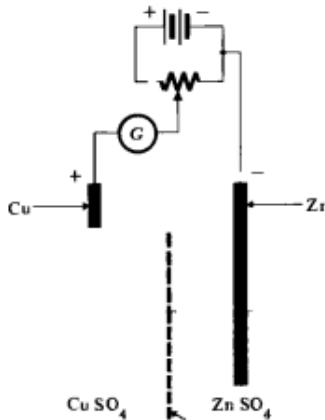
(b) Calculate the change in mean energy $\Delta E = E(x) - E(0)$ of the film when it is stretched at a constant temperature T_0 from a length $x = 0$ to a length x .

(c) Calculate the work $W(0 \rightarrow x)$ done on the film in order to stretch it at this constant temperature from a length $x = 0$ to a length x .

- 5.16** Consider an electrochemical cell of the type illustrated in the figure. The cell can be maintained in equilibrium by connecting a potentiometer across its terminals in such a way that the emf \mathcal{U} produced by the cell is precisely compensated and no net current flows through the outside circuit. The following chemical reaction can take place in the cell:



Suppose that the equilibrium is shifted quasi-statically by an infinitesimal amount so that the reaction proceeds from left to right, dN atoms of copper (Cu) being produced in the process. Then a charge $ze dN$ flows from the Cu to the zinc (Zn) electrode through the outside circuit (where $z = 2$ is the valence of copper) and the cell does an amount of work $\mathcal{U}ze dN$. Expressed in terms of moles, when $d\nu = dN/N_a$ moles of Cu are produced, the charge transferred is $ze N_a d\nu = zf d\nu$ (where $f = N_a e$ is called the Faraday constant), and the work done by the cell is $\mathcal{U}zf d\nu$.



Porous partition

The cell can thus be described in terms of the following independent parameters: its temperature T , the pressure p , and the number of moles ν of Cu metal. The volume change of the material in the cell is negligible. The fundamental thermodynamic relation for this cell becomes then

$$T dS = dE + zfV d\nu \quad (2)$$

Use this relation to find an expression for the change ΔE in the mean energy of the cell at a fixed temperature T and pressure p when one mole of Cu is produced. Show thereby that ΔE (which is the heat of reaction involved in the chemical transformation (1)) can be determined solely by measurements of the emf V of the cell without the necessity of doing any calorimetry.

- 5.17** The equation of state of a gas can be written in the form

$$p = nkT(1 + B_2n)$$

where p is the mean pressure of the gas, T its absolute temperature, $n = N/V$ the number of molecules per unit volume, and $B_2 = B_2(T)$ is the second virial coefficient. The discussion of Sec. 5.10 showed that B_2 is an increasing function of the temperature.

Find how the mean internal energy E of this gas depends on its volume V , i.e., find an expression for $(\partial E/\partial V)_T$. Is it positive or negative?

- 5.18** The free expansion of a gas is a process where the total mean energy E remains constant. In connection with this process, the following quantities are of interest.

(a) What is $(\partial T/\partial V)_E$? Express the result in terms of p , T , $(\partial p/\partial T)_V$, and C_V .

(b) What is $(\partial S/\partial V)_E$? Express the result in terms of p and T .

(c) Using the results (a) and (b), calculate the temperature change $\Delta T = T_2 - T_1$ in a free expansion of a gas from volume V_1 to volume V_2 . Give explicit results for ν moles of a van der Waals gas, assuming C_V to be temperature independent.

- 5.19** The van der Waals equation for 1 mole of gas is given by $(p + av^{-2})(v - b) = RT$. In general, curves of p versus v for various values of T exhibit a maximum and a minimum at the two points where $(\partial p/\partial v)_T = 0$ (the curves are similar to those of Fig. 8.6.1). The maximum and minimum coalesce into a single point on that curve where $(\partial^2 p/\partial v^2)_T = 0$ in addition to $(\partial p/\partial v)_T = 0$. This point is called the "critical point" of the substance and its temperature, pressure, and molar volume are denoted by T_c , p_c , and v_c , respectively.

(a) Express a and b in terms of T_c and v_c .

(b) Express p_c in terms of T_c and v_c .

(c) Write the van der Waals equation in terms of the reduced dimensionless variables

$$T' = \frac{T}{T_c}, \quad v' = \frac{v}{v_c}, \quad p' = \frac{p}{p_c}$$

This form should involve neither a nor b .

- 5.20** Find the inversion curve for a van der Waals gas. Express your result in terms of the reduced variables p' and T' , and find p' as a function of T' along the inversion curve. Sketch this curve on a graph of p' versus T' , being sure to indicate quantitatively the intercepts on the T' axis and the location of the maximum pressure on this curve.

- 5.21** The Joule-Kelvin coefficient is given by

$$\mu \equiv \left(\frac{\partial T}{\partial p} \right)_H = \frac{V}{C_p} \left[\frac{T}{V} \left(\frac{\partial V}{\partial T} \right)_p - 1 \right] \quad (1)$$

Since it involves the absolute temperature T , this relation can be used to determine the absolute temperature T .

Consider any readily measurable arbitrary temperature parameter ϑ (e.g., the height of a mercury column). All that is known is that ϑ is some (unknown) function of T ; i.e., $\vartheta = \vartheta(T)$.

(a) Express (1) in terms of the various directly measurable quantities involving the temperature parameter ϑ instead of the absolute temperature T , i.e., in terms of $\mu' \equiv (\partial \vartheta / \partial p)_H$, $C_p' \equiv (dQ/d\vartheta)_p$, $\alpha' \equiv V^{-1}(\partial V / \partial \vartheta)_p$, and the derivative $d\vartheta/dT$.

(b) Show that, by integrating the resulting expression, one can find T for any given value of ϑ if one knows that $\vartheta = \vartheta_0$ when $T = T_0$ (e.g., if one knows the value of $\vartheta = \vartheta_0$ at the triple point where $T_0 = 273.16$).

- 5.22** Refrigeration cycles have been developed for heating buildings. The procedure is to design a device which absorbs heat from the surrounding earth or air outside the house and then delivers heat at a higher temperature to the interior of the building. (Such a device is called a "heat pump.")

(a) If a device is used in this way, operating between the outside absolute temperature T_0 and an interior absolute temperature T_i , what would be the maximum number of kilowatt-hours of heat that could be supplied to the building for every kilowatt-hour of electrical energy needed to operate the device?

(b) Obtain a numerical answer for the case that the outside temperature is 0°C and the interior temperature is 25°C .

- 5.23** Two identical bodies, each characterized by a heat capacity at constant pressure C which is independent of temperature, are used as heat reservoirs for a heat engine. The bodies remain at constant pressure and undergo no change of phase. Initially, their temperatures are T_1 and T_2 , respectively; finally, as a result of the operation of the heat engine, the bodies will attain a common final temperature T_f .

(a) What is the total amount of work W done by the engine? Express the answer in terms of C , T_1 , T_2 , and T_f .

(b) Use arguments based upon entropy considerations to derive an inequality relating T_f to the initial temperatures T_1 and T_2 .

(c) For given initial temperatures T_1 and T_2 , what is the maximum amount of work obtainable from the engine?

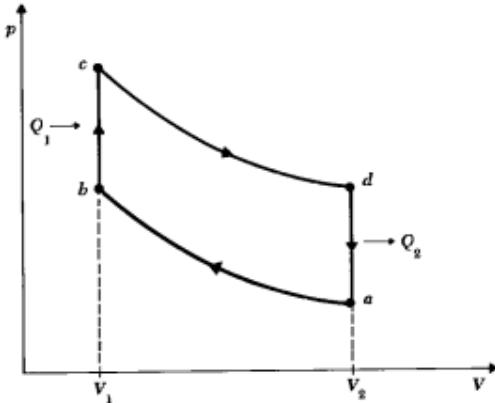
- 5.24** The latent heat of melting of ice is L per unit mass. A bucket contains a mixture of water and ice at the ice point (absolute temperature T_0). It is desired to use a refrigerator in order to freeze an additional mass m of water in the bucket. The heat rejected by the refrigerator goes to warm up a body of constant heat capacity C and, initially, also at temperature T_0 . What is the minimum amount of heat which the refrigerator must reject to this body in the process?

- 5.25** Consider the physical situation illustrated in Fig. 5.11.2. Suppose that under the influence of gravity ($g = 980 \text{ cm sec}^{-2}$) the weight, having a mass $m = 50$ grams, is allowed to descend a distance $L = 1 \text{ cm}$ before coming to rest on a platform. In this process the weight turns the paddle wheel and raises the temperature of the liquid by a slight amount above its original temperature of 25°C .

Calculate the probability that, as a result of a spontaneous fluctuation, the

water gives off its energy to the weight and raises it again so as to restore it to a height of 1 cm or more.

- 5.26 A gasoline engine can be *approximately* represented by the idealized cyclic process *abcd* shown in the accompanying diagram of pressure p versus volume V of the gas in the cylinder. Here $a \rightarrow b$ represents the adiabatic compression of the air-gasoline mixture, $b \rightarrow c$ the rise in pressure at constant volume due to the explosion of the mixture, $c \rightarrow d$ the adiabatic expansion of the mixture during which the engine performs useful work, and $d \rightarrow a$ the final cooling down of the gas at constant volume.



Assume this cycle to be carried out quasi-statically for a fixed amount of ideal gas having a constant specific heat. Denote the specific heat ratio by $\gamma = c_p/c_v$. Calculate the efficiency η (ratio of work performed to heat intake Q_1) for this process, expressing your answer in terms of V_1 , V_2 , and γ .

Basic methods and results of statistical mechanics

6

ONE CAN readily discuss many more important applications based on the macroscopic aspects of the general theory of Chapter 3. But appreciably greater insight and power are gained by considering also the microscopic aspects of the theory. In this chapter we shall therefore turn our attention to the statistical relations summarized in the latter part of Sec. 3·11. Our aim will be (1) to derive general probability statements for a variety of situations of physical interest and (2) to describe practical methods for calculating macroscopic quantities (such as entropies or specific heats) from a knowledge of the purely microscopic properties of a system. In Chapter 7 we shall then apply these methods to a discussion of some important physical situations.

ENSEMBLES REPRESENTATIVE OF SITUATIONS OF PHYSICAL INTEREST

6·1 *Isolated system*

In giving a statistical description of a system, one always has some information available about the physical situation under consideration. The representative statistical ensemble is then constructed in such a way that all the systems in the ensemble satisfy conditions consistent with one's information about the system. Since it is possible to visualize a variety of physical situations, one is led to consider a corresponding number of representative ensembles. We shall describe some of the most important cases in the following sections.

An isolated system represents a situation of fundamental importance, one which we have discussed at length in Chapters 2 and 3. Indeed, whenever one is dealing with a situation where a system A is not isolated but is allowed to interact with some other system A' , it is always possible to reduce the situation to the case of an isolated system by focusing attention on the combined system $A + A'$.

For the sake of simplicity, suppose that the volume V of the system is its only relevant external parameter. An isolated system of this kind consists then of a given number N of particles in a specified volume V , the constant energy of the system being known to lie in some range between E and $E + \delta E$. Probability statements are then made with reference to an ensemble which consists of many such systems, all consisting of this number N of particles in this volume V , and all with their energy lying in the range between E and $E + \delta E$. The fundamental statistical postulate asserts that in an equilibrium situation the system is equally likely to be found in any one of its accessible states. Thus, if the energy of a system in state r is denoted by E_r , the probability P_r of finding the system in state r is given by

$$P_r = \begin{cases} C & \text{if } E < E_r < E + \delta E \\ 0 & \text{otherwise} \end{cases} \quad (6 \cdot 1 \cdot 1)$$

where C is a constant. It can be determined by the normalization condition that $\sum P_r = 1$ when summed over all accessible states in the range between E and $E + \delta E$.

An ensemble representing an isolated system in equilibrium consists then of systems distributed in accordance with (6 · 1 · 1). It is sometimes called a "microcanonical" ensemble.

6 · 2 System in contact with a heat reservoir

We consider the case of a small system A in thermal interaction with a heat reservoir A' . This is the situation already discussed in Sec. 3 · 6 where $A \ll A'$, i.e., where A has many fewer degrees of freedom than A' . The system A may be any relatively small *macroscopic* system. (For example, it may be a bottle of wine immersed in a swimming pool, the pool acting as a heat reservoir.) Sometimes it may also be a distinguishable *microscopic* system which can be clearly identified.* (For example, it may be an atom at some lattice site in a solid, the solid acting as a heat reservoir.) We ask the following question: Under conditions of equilibrium, what is the probability P_r of finding the system A in any *one* particular microstate r of energy E_r ?

This question is immediately answered by the same reasoning as was used in Sec. 3 · 3. We again assume weak interaction between A and A' so that their energies are additive. The energy of A is, of course, not fixed. It is only the total energy of the combined system $A^{(0)} = A + A'$ which has a constant value in some range between $E^{(0)}$ and $E^{(0)} + \delta E$. The conservation of energy can then be written as

$$E_r + E' = E^{(0)} \quad (6 \cdot 2 \cdot 1)$$

where E' denotes the energy of the reservoir A' . When A has an energy E_r ,

* The qualifying remark is introduced because it may not always be possible to label the identity of an individual atomic particle in a quantum mechanical description.

the reservoir A' must then have an energy near $E' = E^{(0)} - E_r$. Hence, if A is in the *one* definite state r , the number of states accessible to the combined system $A^{(0)}$ is just the number of states $\Omega'(E^{(0)} - E_r)$ accessible to A' when its energy lies in a range δE near the value $E' = E^{(0)} - E_r$. But, according to the fundamental statistical postulate, the probability of occurrence in the ensemble of a situation where A is in state r is simply proportional to the number of states accessible to $A^{(0)}$ under these conditions. Hence

$$\blacktriangleright \quad P_r = C' \Omega'(E^{(0)} - E_r) \quad (6 \cdot 2 \cdot 2)$$

where C' is a constant of proportionality independent of r . As usual, it can be determined from the normalization condition for probabilities, i.e.,

$$\sum_r P_r = 1 \quad (6 \cdot 2 \cdot 3)$$

where the sum extends over all possible states of A irrespective of energy.

Up to now, our discussion has been completely general. Let us now make use of the fact that A is a very much smaller system than A' . Then $E_r \ll E^{(0)}$ and (6·2·2) can be approximated by expanding the slowly varying logarithm of $\Omega'(E')$ about the value $E' = E^{(0)}$. Thus

$$\ln \Omega'(E^{(0)} - E_r) = \ln \Omega'(E^{(0)}) - \left[\frac{\partial \ln \Omega'}{\partial E'} \right]_0 E_r \dots \quad (6 \cdot 2 \cdot 4)$$

Since A' acts as a heat reservoir, $E_r \ll E^{(0)}$ and higher-order terms in the expansion can be neglected. The derivative

$$\left[\frac{\partial \ln \Omega'}{\partial E'} \right]_0 \equiv \beta \quad (6 \cdot 2 \cdot 5)$$

is evaluated at the fixed energy $E' = E^{(0)}$ and is thus a constant independent of the energy E_r of A . By (3·3·10) it is just the constant temperature parameter $\beta = (kT)^{-1}$ characterizing the *heat reservoir* A' . (Physically, this means that the reservoir A' is so large compared to A that its temperature remains unaffected by whatever small amount of energy it gives to A .) Hence (6·2·4) becomes

$$\begin{aligned} \ln \Omega'(E^{(0)} - E_r) &= \ln \Omega'(E^{(0)}) - \beta E_r \\ \text{or} \quad \Omega'(E^{(0)} - E_r) &= \Omega'(E^{(0)}) e^{-\beta E_r} \end{aligned} \quad (6 \cdot 2 \cdot 6)$$

Since $\Omega'(E^{(0)})$ is just a constant independent of r , (6·2·2) becomes then simply

$$\blacktriangleright \quad P_r = C e^{-\beta E_r}, \quad (6 \cdot 2 \cdot 7)$$

where C is some constant of proportionality independent of r . Using the normalization condition (6·2·3), C is determined by the relation

$$C^{-1} = \sum_r e^{-\beta E_r}$$

so that (6·2·7) can also be written explicitly in the form

$$P_r = \frac{e^{-\beta E_r}}{\sum e^{-\beta E_r}} \quad (6 \cdot 2 \cdot 8)$$

Let us discuss the results (6·2·2) or (6·2·7) more fully. If A is known to be in a definite one of its states, the reservoir can be in any one of the large number $\Omega'(E^{(0)} - E_r)$ of states accessible to it. Remember that ordinarily the number of states $\Omega'(E')$ accessible to the reservoir is a very rapidly increasing function of its energy (i.e., β in (6·2·5) is positive). Thus, if A is in a state r where its energy E_r is higher, the conservation of energy for the total system implies that the energy of the reservoir is correspondingly lower so that the number of states accessible to the reservoir is markedly reduced. The probability of encountering this situation in the ensemble is accordingly very much less. The exponential dependence of P_r on E_r in (6·2·7) just expresses this state of affairs in mathematical terms.

Example A simple numerical illustration is provided by Fig. 6·2·1 where the bar graphs show the number of states accessible to A and A' for various values of their respective energies. Assume that the total energy of the combined system is known to be 1007. Suppose that A is in one of its states, call it r , of energy 6. Then the energy of the reservoir A' must be 1001 so that it can be in any one of 400,000 possible states. In an ensemble

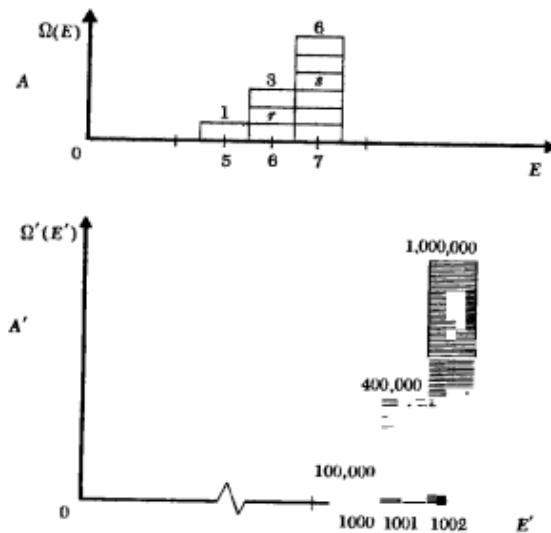


Fig. 6·2·1 Schematic illustration (not drawn to scale) showing the number of states accessible to a system A and to a heat reservoir A' as a function of their respective energies. (The energy scale is in terms of an arbitrary unit.)

consisting of many systems $A^{(0)}$, there will then be 400,000 different possible examples of systems for which A is in state r . Suppose, however, that we consider the situation where A is in a state, such as s , where its energy is 7. Here the reservoir must have energy 1000 so that only 100,000 states are accessible to it. The ensemble contains then merely 100,000 different possible examples of systems for which A is in state s .

The probability (6.2.7) is a very general result and is of fundamental importance in statistical mechanics. The exponential factor $e^{-\beta E_r}$ is called the "Boltzmann factor"; the corresponding probability distribution (6.2.7) is known as the "canonical distribution." An ensemble of systems all of which are in contact with a heat reservoir of known temperature T , i.e., all of which are distributed over states in accordance with (6.2.7), is called a "canonical ensemble."

The fundamental result (6.2.7) gives the probability of finding A in one particular state r of energy E_r . The probability $P(E)$ that A has an energy in a small range between E and $E + \delta E$ is then simply obtained by adding the probabilities for all states whose energy lies in this range; i.e.,

$$P(E) = \sum_r P_r$$

where r is such that $E < E_r < E + \delta E$. But all these states are, by (6.2.7), equally probable and are characterized by essentially the same exponential factor $e^{-\beta E}$; hence one needs simply to multiply the probability of finding A in any one of these states by the number $\Omega(E)$ of its states in this energy range, i.e.,

$$P(E) = C\Omega(E) e^{-\beta E} \quad (6.2.9)$$

To the extent that A itself is a large system (although very much smaller than A'), $\Omega(E)$ is a rapidly increasing function of E . The presence of the rapidly decreasing factor $e^{-\beta E}$ in (6.2.9) results then in a maximum of the product $\Omega(E) e^{-\beta E}$. The larger A is, the sharper is this maximum in $P(E)$; i.e., the more

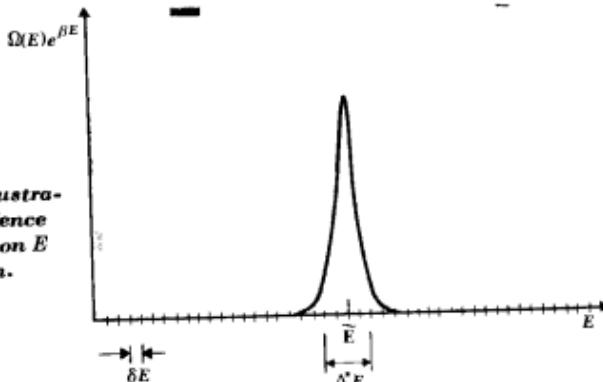


Fig. 6.2.2 Schematic illustration showing the dependence of the function $\Omega(E)e^{-\beta E}$ on E for a macroscopic system.

rapidly $\Omega(E)$ increases with E , the sharper this maximum becomes. Thus we arrive again at the conclusions of Sec. 3 · 7. We emphasize, however, that (6 · 2 · 9) is valid no matter how small A is. It may even be a system of atomic size, provided that it can be treated as a distinguishable system satisfying the additivity of energy (6 · 2 · 1).

Once the probability distribution (6 · 2 · 7) is known, various mean values can readily be computed. For example, let y be any quantity assuming the value y_r in state r of the system A . Then

$$\bar{y} = \frac{\sum_r e^{-\beta E_r} y_r}{\sum_r e^{-\beta E_r}} \quad (6 \cdot 2 \cdot 10)$$

where the summation is over all states r of the system A .

6 · 3 Simple applications of the canonical distribution

The canonical distribution (6 · 2 · 7) yields a host of conclusions. Here we mention only a few illustrative applications where the canonical distribution leads immediately to physically very important results. Most of these will be discussed more fully in Chapter 7.

Paramagnetism Consider a substance which contains N_0 magnetic atoms per unit volume and which is placed in an external magnetic field \mathbf{H} . Assume that each atom has spin $\frac{1}{2}$ (corresponding to one unpaired electron) and an intrinsic magnetic moment μ . In a quantum-mechanical description the magnetic moment of each atom can then point either parallel or antiparallel to the external field \mathbf{H} . If the substance is at absolute temperature T , what is the mean magnetic moment $\bar{\mu}_H$ (in the direction of \mathbf{H}) of such an atom? We assume that each atom interacts only weakly with the other atoms and with the other degrees of freedom of the substance. It is then permissible to focus attention on a single atom as the small system under consideration and to regard all the other atoms and other degrees of freedom as constituting a heat reservoir.*

Each atom can be in two possible states: the state (+) where its spin points up (i.e., parallel to \mathbf{H}) and the state (−) where its spin points down (i.e., antiparallel to \mathbf{H}). Let us discuss these states in turn.

In the (+) state, the atomic magnetic moment μ is parallel to \mathbf{H} so that $\mu_H = \mu$. The corresponding magnetic energy of the atom is then $\epsilon_+ = -\mu H$.

* This assumes that it is possible to identify a single atom unambiguously, an assumption which is justified if the atoms are localized at definite lattice sites of a solid or if they form a dilute gas where the atoms are widely separated. In a concentrated gas the assumption might break down. It would then be necessary to adopt a point of view (which is always permissible, although more complicated) which considers the *entire* gas of atoms as a small microscopic system in contact with a heat reservoir provided by other degrees of freedom.

The probability of finding the atom in this state is thus

$$P_+ = C e^{-\beta \epsilon_+} = C e^{\beta \mu H} \quad (6 \cdot 3 \cdot 1)$$

where C is a constant of proportionality and $\beta = (kT)^{-1}$. This is the state of lower energy (if μ is positive) and is thus the state in which the atom is more likely to be found.

In the $(-)$ state, \mathbf{u} is antiparallel to \mathbf{H} so that $\mu_H = -\mu$. The corresponding energy of the atom is then $\epsilon_- = +\mu H$. The probability of finding the atom in this state is thus

$$P_- = C e^{-\beta \epsilon_-} = C e^{-\beta \mu H} \quad (6 \cdot 3 \cdot 2)$$

This is the state of higher energy (if μ is positive) and is thus the state in which the atom is less likely to be found.

Since the first state where \mathbf{u} is parallel to \mathbf{H} is more probable, it is clear that the mean magnetic moment $\bar{\mu}_H$ must point in the direction of the external field \mathbf{H} . By virtue of (6·3·1) and (6·3·2), the significant parameter in this problem is the quantity

$$y \equiv \beta \mu H = \frac{\mu H}{kT}$$

which measures the ratio of a typical magnetic energy to a typical thermal energy. It is apparent that if T is very large, i.e., if $y \ll 1$, the probability that \mathbf{u} is parallel to \mathbf{H} is almost the same as that of its being antiparallel. In this case \mathbf{u} is almost completely randomly oriented so that $\bar{\mu}_H \approx 0$. On the other hand, if T is very small, i.e., if $y \gg 1$, then it is much more probable that \mathbf{u} is parallel to \mathbf{H} rather than antiparallel to it. In this case $\bar{\mu}_H \approx \mu$.

All these qualitative conclusions can readily be made quantitative by actually calculating the mean value $\bar{\mu}_H$. Thus we have

$$\bar{\mu}_H = \frac{P_{+\mu} + P_{-(-\mu)}}{P_+ + P_-} = \mu \frac{e^{\beta \mu H} - e^{-\beta \mu H}}{e^{\beta \mu H} + e^{-\beta \mu H}}$$

$$\text{or} \quad \bar{\mu}_H = \mu \tanh \frac{\mu H}{kT} \quad (6 \cdot 3 \cdot 3)$$

Here we have used the definition of the hyperbolic tangent

$$\tanh y \equiv \frac{e^y - e^{-y}}{e^y + e^{-y}}$$

The "magnetization" \bar{M}_0 , or mean magnetic moment per unit volume, is then in the direction of H and is given by

$$\bar{M}_0 = N_0 \bar{\mu}_H \quad (6 \cdot 3 \cdot 4)$$

One can easily check that $\bar{\mu}_H$ exhibits the qualitative behavior already discussed. If $y \ll 1$, then $e^y = 1 + y + \dots$ and $e^{-y} = 1 - y + \dots$. Hence

$$\text{for } y \ll 1, \quad \tanh y = \frac{(1 + y + \dots) - (1 - y + \dots)}{2} = y$$

On the other hand, if $y \gg 1$, then $e^y \gg e^{-y}$. Hence

$$\text{for } y \gg 1, \quad \tanh y = 1$$

Thus (6·3·3) leads to the limiting behavior that

$$\text{for } \mu H/kT \ll 1, \quad \bar{\mu}_H = \frac{\mu^2 H}{kT} \quad (6\cdot3\cdot5a)$$

$$\text{for } \mu H/kT \gg 1, \quad \bar{\mu}_H = \mu \quad (6\cdot3\cdot5b)$$

By (6·3·4) and (6·3·5a) it then follows that

$$\text{if } \mu H/kT \ll 1, \quad \bar{M}_0 = \chi H \quad (6\cdot3\cdot6)$$

where χ is a constant of proportionality independent of H . This parameter χ is called the “magnetic susceptibility” of the substance. Equation (6·3·5a) provides an explicit expression for χ in terms of microscopic quantities, i.e.,

$$\chi = \frac{N_0 \mu^2}{kT} \quad (6\cdot3\cdot7)$$

The fact that $\chi \propto T^{-1}$ is known as Curie's law. On the other hand,

$$\text{if } \mu H/kT \gg 1, \quad \bar{M}_0 \rightarrow N_0 \mu \quad (6\cdot3\cdot8)$$

becomes independent of H and equal to the maximum (or “saturation”) magnetization which the substance can exhibit. The complete dependence of the magnetization \bar{M}_0 on temperature T and magnetic field H is shown in Fig. 6·3·1.

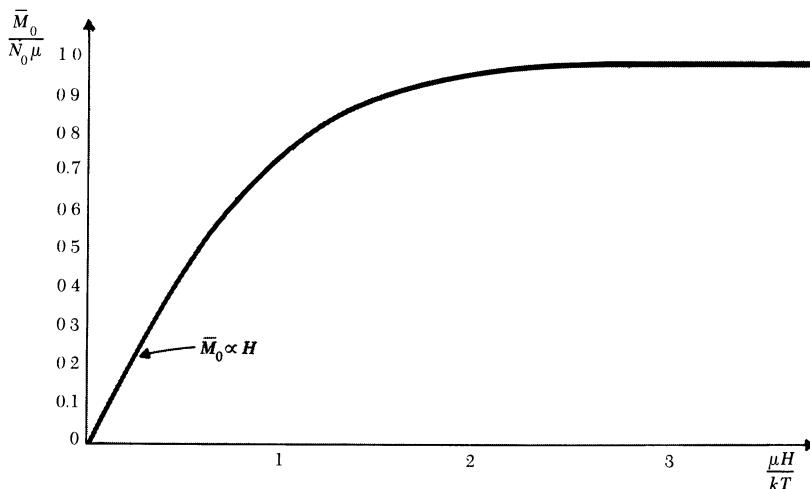


Fig. 6·3·1 Dependence of the magnetization \bar{M}_0 on magnetic field H and temperature T for noninteracting magnetic atoms of spin $\frac{1}{2}$ and magnetic moment μ .

Molecule in an ideal gas Consider a monatomic gas at absolute temperature T confined in a container of volume V . Assume that the number of molecules per unit volume is small enough that the interaction between molecules is very weak; then the total energy of the gas equals the sum of the energies of each molecule. We treat the problem classically so that it is permissible to focus attention on a given distinct molecule (without having to be concerned about the essential indistinguishability of the molecules in the gas). All the remaining molecules can then be regarded as a heat reservoir at temperature T .

The molecule can only be located somewhere inside the container. There its energy is purely kinetic; i.e.,

$$E = \frac{1}{2} m v^2 = \frac{1}{2} \frac{\mathbf{p}^2}{m} \quad (6 \cdot 3 \cdot 9)$$

where m is the mass of the molecule and $\mathbf{v} = \mathbf{p}/m$ is its velocity. If the molecule's position lies in the range between \mathbf{r} and $\mathbf{r} + d\mathbf{r}$ (i.e., if its x coordinate lies between x and $x + dx$, its y coordinate between y and $y + dy$, and its z coordinate between z and $z + dz$) and if its momentum lies in the range between \mathbf{p} and $\mathbf{p} + d\mathbf{p}$ (i.e., if its x component of momentum lies between p_x and $p_x + dp_x$, . . .), then the volume of phase space corresponding to this range of \mathbf{r} and \mathbf{p} is

$$d^3\mathbf{r} d^3\mathbf{p} \equiv (dx dy dz)(dp_x dp_y dp_z) \quad (6 \cdot 3 \cdot 10)$$

To find the probability $P(\mathbf{r}, \mathbf{p}) d^3\mathbf{r} d^3\mathbf{p}$ that the molecule has position lying in the range between \mathbf{r} and $\mathbf{r} + d\mathbf{r}$ and momentum in the range between \mathbf{p} and $\mathbf{p} + d\mathbf{p}$, one need only multiply the number $(d^3\mathbf{r} d^3\mathbf{p})/h_0^3$ of cells in phase space corresponding to this range by the probability that the molecule is found in a particular cell. Thus

$$P(\mathbf{r}, \mathbf{p}) d^3\mathbf{r} d^3\mathbf{p} \propto \left(\frac{d^3\mathbf{r} d^3\mathbf{p}}{h_0^3} \right) e^{-\beta(p^2/2m)} \quad (6 \cdot 3 \cdot 11)$$

where $\beta \equiv (kT)^{-1}$.

Note that the probability density P does not depend on the position \mathbf{r} of the molecule in the box. This reflects merely the fact that, in the absence of external forces, the symmetry of the physical situation is such that there can be no preferred location of a molecule within the box.

To find the probability $P(\mathbf{p}) d^3\mathbf{p}$ that a molecule has momentum lying in the range between \mathbf{p} and $\mathbf{p} + d\mathbf{p}$, irrespective of its location \mathbf{r} , one need only sum the probability (6·3·11) over all possible positions \mathbf{r} , i.e., integrate it over the volume of the container

$$P(\mathbf{p}) d^3\mathbf{p} = \int_{(r)} P(\mathbf{r}, \mathbf{p}) d^3\mathbf{r} d^3\mathbf{p} \propto e^{-\beta(p^2/2m)} d^3\mathbf{p} \quad (6 \cdot 3 \cdot 12)$$

One can equally well express this in terms of the velocity $\mathbf{v} = \mathbf{p}/m$. The probability $P'(\mathbf{v}) d^3\mathbf{v}$ that a molecule has a velocity between \mathbf{v} and $\mathbf{v} + d\mathbf{v}$ is

then

$$P'(\mathbf{v}) d^3\mathbf{v} = P(\mathbf{p}) d^3\mathbf{p} = C e^{-\beta m \mathbf{v}^2/2} d^3\mathbf{v} \quad (6 \cdot 3 \cdot 13)$$

where C is a constant of proportionality which can be determined by the normalization condition that the integral of the probability (6·3·13) over all possible velocities of the molecule must be equal to unity. The result (6·3·13) is the famous “Maxwell distribution” of molecular velocities.

Molecule in an ideal gas in the presence of gravity Consider the situation of the preceding example, but suppose now that a uniform gravitational field acts in the $-z$ direction. Then, instead of (6·3·9), the energy of a molecule in the gas becomes

$$E = \frac{\mathbf{p}^2}{2m} + mgz \quad (6 \cdot 3 \cdot 14)$$

where g is the constant acceleration due to gravity. Analogously to (6·3·11), one then has

$$\begin{aligned} P(\mathbf{r}, \mathbf{p}) d^3\mathbf{r} d^3\mathbf{p} &\propto \frac{d^3\mathbf{r} d^3\mathbf{p}}{h_0^3} e^{-\beta[(\mathbf{p}^2/2m) + mgz]} \\ &\propto d^3\mathbf{r} d^3\mathbf{p} e^{-\beta(\mathbf{p}^2/2m)} e^{-\beta mgz} \end{aligned} \quad (6 \cdot 3 \cdot 15)$$

The probability now does depend on the z coordinate of the molecule. The probability $P(\mathbf{p}) d^3\mathbf{p}$ that a molecule has momentum in the range between \mathbf{p} and $\mathbf{p} + d\mathbf{p}$, irrespective of its location, is given as before by

$$P(\mathbf{p}) d^3\mathbf{p} = \int_{(\mathbf{r})} P(\mathbf{r}, \mathbf{p}) d^3\mathbf{r} d^3\mathbf{p} \quad (6 \cdot 3 \cdot 16)$$

where the integration over \mathbf{r} extends over the volume V of the container. Since (6·3·15) factors into the product of two exponentials, (6·3·16) becomes simply

$$P(\mathbf{p}) d^3\mathbf{p} = C e^{-\beta(\mathbf{p}^2/2m)} d^3\mathbf{p} \quad (6 \cdot 3 \cdot 17)$$

where C is a constant of proportionality. This means that the momentum distribution function, and thus also the velocity distribution function, is exactly the same as that obtained in (6·3·12) in the absence of a gravitational field.

Finally we can find the probability $P(z) dz$ that a molecule is located at a height between z and $z + dz$, irrespective of its momentum or x and y position components. This is found from (6·3·15) by integration:

$$P(z) dz = \int_{(x,y)} \int_{(\mathbf{p})} P(\mathbf{r}, \mathbf{p}) d^3\mathbf{r} d^3\mathbf{p} \quad (6 \cdot 3 \cdot 18)$$

where one integrates over all momenta (from $-\infty$ to $+\infty$ for each momentum component) and over all possible x and y values lying within the container, (i.e., over the cross-sectional area of the container). Again, since (6·3·15) factors into a product of exponentials, (6·3·18) becomes simply

$$P(z) dz = C' e^{-\beta mgz} dz \quad (6 \cdot 3 \cdot 19)$$

where C' is a constant of proportionality. This implies that

$$P(z) = P(0) e^{-mgz/kT} \quad (6 \cdot 3 \cdot 20)$$

i.e., the probability of finding a molecule at height z decreases exponentially with the height. The result (6·3·20) is sometimes called the "law of atmospheres," since it would describe the density variation of the air near the surface of the earth if the atmosphere were at a constant temperature (which it is *not*).

6 · 4 System with specified mean energy

Another situation of physical interest is that where a system A consists of a fixed number N of particles in a given volume V , but where the only information available about the energy of the system is its *mean* energy \bar{E} . This is a very common situation. Suppose, for example, that a system A is brought to some final macrostate as a result of interaction with other macroscopic systems. Then the measurement of the macroscopic work done or the heat absorbed in the process does not tell one the energy of each system in the ensemble, but provides information only about the *mean* energy of the final macrostate of A .

A system A with specified *mean* energy \bar{E} is also described by a canonical distribution. For, if such a system were placed in thermal contact with a heat reservoir at some temperature β , the mean energy of the system would be determined. Thus a proper choice of β would guarantee that the mean energy of the system assumes the specified value \bar{E} .

A more direct argument is readily given. Denote the energy of the system A in state r by E_r . Suppose that the statistical ensemble consists of a very large number a of such systems, a_r of which are in state r . Then the information available to us is that

$$\frac{1}{a} \sum_s a_s E_s = \bar{E} \quad (6 \cdot 4 \cdot 1)$$

equals the specified mean energy. Thus it follows that

$$\Sigma a_s E_s = a\bar{E} = \text{constant}$$

This implies that the situation is equivalent to one where a fixed total amount of energy $a\bar{E}$ is to be distributed over all the systems in the ensemble, each such system being equally likely to be in any one state. If a system in the ensemble is in state r , the remaining $(a - 1)$ systems must then have a combined energy $(a\bar{E} - E_r)$. These $(a - 1)$ systems can be distributed over some very large number $\Phi(E')$ of accessible states if their combined energy is E' . If the one system under consideration is in state r , the remaining $(a - 1)$ systems can then be with equal probability in any of the $\Phi(a\bar{E} - E_r)$ states accessible to them. Since $E_r \ll a\bar{E}$, the mathematical problem is here exactly the same as that of Sec. 6·2, dealing with a system in thermal contact with a heat reservoir, except that the role of energy reservoir is now played not by any physical heat

reservoir of specified temperature parameter β , but by the totality of all the other systems in the ensemble. Accordingly, one gets again the canonical distribution

$$P_r \propto e^{-\beta E_r} \quad (6 \cdot 4 \cdot 2)$$

The parameter $\beta = (\partial \ln \Phi / \partial E')$ does not here have any immediate physical significance in terms of the temperature of a real heat bath. Rather, it is to be determined by the condition that the mean energy calculated with the distribution (6·4·2) is indeed equal to the specified mean value \bar{E} , i.e., by the condition

$$\frac{\sum_r e^{-\beta E_r} E_r}{\sum_r e^{-\beta E_r}} = \bar{E} \quad (6 \cdot 4 \cdot 3)$$

In short, when one is dealing with a system in contact with a heat reservoir of temperature $\beta = (kT)^{-1}$, the canonical distribution (6·4·2) is valid and the mean energy \bar{E} can be calculated by (6·4·3) from the known value of β . If one is dealing with a system of specified mean energy \bar{E} , the canonical distribution (6·4·2) is again valid, but the parameter β is to be calculated by (6·4·3) from the known value of \bar{E} .

6 · 5 Calculation of mean values in a canonical ensemble

When a system A is in thermal contact with a heat reservoir as in Sec. 6·2, or when only its mean energy is known as in Sec. 6·4, the systems in the representative statistical ensemble are distributed over their accessible states in accordance with the canonical distribution

$$P_r = C e^{-\beta E_r} = \frac{e^{-\beta E_r}}{\sum_r e^{-\beta E_r}} \quad (6 \cdot 5 \cdot 1)$$

In these physical situations the energy of the system is not precisely specified and the calculation of important mean values becomes particularly simple.

By (6·5·1) the mean energy is given by

$$\bar{E} = \frac{\sum_r e^{-\beta E_r} E_r}{\sum_r e^{-\beta E_r}} \quad (6 \cdot 5 \cdot 2)$$

where the sums are over *all* accessible states r of the system, irrespective of their energy. The relation (6·5·2) can be reduced to much simpler form by noting that the sum in the numerator can be readily expressed in terms of the sum appearing in the denominator. Thus

$$\sum_r e^{-\beta E_r} E_r = - \sum_r \frac{\partial}{\partial \beta} (e^{-\beta E_r}) = - \frac{\partial}{\partial \beta} Z$$

where

$$\blacktriangleright \quad Z \equiv \sum_r e^{-\beta E_r} \quad (6 \cdot 5 \cdot 3)$$

is just the sum in the denominator of (6·5·2). Hence one obtains

$$\blacktriangleright \quad \bar{E} = - \frac{1}{Z} \frac{\partial Z}{\partial \beta} = - \frac{\partial \ln Z}{\partial \beta} \quad (6 \cdot 5 \cdot 4)$$

The quantity Z defined in (6·5·3) is called the "sum over states" or the "partition function." (The letter Z is used because the German name is "Zustandsumme.")*

The canonical distribution implies a distribution of systems over possible energies; the resulting dispersion of the energy is also readily computed. We can use the general statistical relation (1·3·10), i.e.,

$$\overline{(\Delta E)^2} \equiv \overline{(E - \bar{E})^2} = \overline{E^2 - 2\bar{E}E + \bar{E}^2} = \overline{E^2} - \bar{E}^2 \quad (6 \cdot 5 \cdot 5)$$

Here

$$\overline{E^2} = \frac{\sum_r e^{-\beta E_r} E_r^2}{\sum_r e^{-\beta E_r}} \quad (6 \cdot 5 \cdot 6)$$

$$\text{But } \sum_r e^{-\beta E_r} E_r^2 = - \frac{\partial}{\partial \beta} \left(\sum_r e^{-\beta E_r} E_r \right) = \left(- \frac{\partial}{\partial \beta} \right)^2 \left(\sum_r e^{-\beta E_r} \right)$$

Hence (6·5·6) becomes

$$\overline{E^2} = \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2} \quad (6 \cdot 5 \cdot 7)$$

This can be written in a form involving the mean energy \bar{E} of (6·5·4). Thus

$$\overline{E^2} = \frac{\partial}{\partial \beta} \left(\frac{1}{Z} \frac{\partial Z}{\partial \beta} \right) + \frac{1}{Z^2} \left(\frac{\partial Z}{\partial \beta} \right)^2 = - \frac{\partial \bar{E}}{\partial \beta} + \bar{E}^2$$

Hence (6·5·5) yields

$$\blacktriangleright \quad \overline{(\Delta E)^2} = - \frac{\partial \bar{E}}{\partial \beta} = \frac{\partial^2 \ln Z}{\partial \beta^2} \quad (6 \cdot 5 \cdot 8)$$

Since $(\Delta E)^2$ can never be negative, it follows that $\partial \bar{E} / \partial \beta \leq 0$ (or equivalently, that $\partial \bar{E} / \partial T \geq 0$). These results agree with those of (3·7·15) and (3·7·16).

Suppose that the system is characterized by a single external parameter x . (The generalization of all results to the case when there are several such parameters will be immediate.) Consider a quasi-static change of the external parameter from x to $x + dx$. In this process the energy of the system in state r changes by the amount

$$\Delta_x E_r = \frac{\partial E_r}{\partial x} dx$$

The macroscopic work dW done by the system as a result of this parameter

* Note that since there are in general very many states of the same energy, the sum Z contains very many terms which are equal.

change is then, corresponding to (2·9·5), given by

$$dW = \frac{\sum_r e^{-\beta E_r} \left(-\frac{\partial E_r}{\partial x} dx \right)}{\sum_r e^{-\beta E_r}} \quad (6 \cdot 5 \cdot 9)$$

where the mean value has been calculated with the canonical distribution (6·5·1). Once again the numerator can be written in terms of Z . Thus

$$\sum_r e^{-\beta E_r} \frac{\partial E_r}{\partial x} = -\frac{1}{\beta} \frac{\partial}{\partial x} \left(\sum_r e^{-\beta E_r} \right) = -\frac{1}{\beta} \frac{\partial Z}{\partial x}$$

and (6·5·9) becomes

$$dW = \frac{1}{\beta Z} \frac{\partial Z}{\partial x} dx = \frac{1}{\beta} \frac{\partial \ln Z}{\partial x} dx \quad (6 \cdot 5 \cdot 10)$$

Since one can express dW in terms of the mean generalized force X

$$dW = \bar{X} dx, \quad \bar{X} \equiv -\frac{\partial \bar{E}_r}{\partial x}$$

it follows also, by (6·5·10), that

► $\bar{X} = \frac{1}{\beta} \frac{\partial \ln Z}{\partial x} \quad (6 \cdot 5 \cdot 11)$

For example, if $x = V$, the volume of the system, Eq. (6·5·11) provides an expression for its mean pressure. That is,

$$dW = \bar{p} dV = \frac{1}{\beta} \frac{\partial \ln Z}{\partial V} dV$$

or

► $\bar{p} = \frac{1}{\beta} \frac{\partial \ln Z}{\partial V} \quad (6 \cdot 5 \cdot 12)$

Now Z is a function of β and V (since the energies E_r depend on V). Hence (6·5·12) is an equation relating \bar{p} to $T = (k\beta)^{-1}$ and V , i.e., it gives the equation of state of the system.

6 · 6 Connection with thermodynamics

Note that all the important physical quantities can be expressed completely in terms of $\ln Z$.* In particular, the fact that both the mean energy \bar{E} and the work dW are expressible in terms of $\ln Z$ implies immediately the intimate connection between $d\bar{E}$ and dW which is the content of the second law of thermodynamics. To show this explicitly, we recall that Z in (6·5·3) is a function of both

* The situation is completely analogous to that encountered in (3·12·1), where all physical quantities could be expressed in terms of $\ln \Omega$. The physical consequence (the validity of the second law in the form (6·6·4)) is the same in both cases.

β and x , since $E_r = E_r(x)$. Hence one has $Z = Z(\beta, x)$ and can write for a small change of this quantity

$$d \ln Z = \frac{\partial \ln Z}{\partial x} dx + \frac{\partial \ln Z}{\partial \beta} d\beta \quad (6 \cdot 6 \cdot 1)$$

Consider a quasi-static process where x and β change so slowly that the system is always very close to equilibrium and thus always distributed according to the canonical distribution; then (6·6·1) implies, by virtue of (6·5·4) and (6·5·10), the relation

$$d \ln Z = \beta dW - \bar{E} d\beta \quad (6 \cdot 6 \cdot 2)$$

The last term can be rewritten in terms of the change in \bar{E} rather than the change in β . Thus

$$d \ln Z = \beta dW - d(\bar{E}\beta) + \beta d\bar{E}$$

or $d(\ln Z + \beta\bar{E}) = \beta(dW + d\bar{E}) = \beta dQ \quad (6 \cdot 6 \cdot 3)$

where we have used the definition (2·8·3) for the heat dQ absorbed by the system. Equation (6·6·3) shows again that although dQ is not an exact differential, an exact differential results when dQ is multiplied by the temperature parameter β . This is, of course, the content of the second law of thermodynamics, previously derived in (3·9·5) and expressed there in the form

$$dS = \frac{dQ}{T} \quad (6 \cdot 6 \cdot 4)$$

The identification of (6·6·3) and (6·6·4) becomes complete if one puts

$$S = k(\ln Z + \beta\bar{E}) \quad (6 \cdot 6 \cdot 5)$$

It can readily be verified that this result agrees with the general definition $S = k \ln \Omega(\bar{E})$ introduced in (3·3·12) for the entropy of a macroscopic system of mean energy \bar{E} . The partition function (6·5·3) is a sum over all states r , very many of which have the same energy. One can perform the sum by first summing over all the $\Omega(E)$ states in the energy range between E and $E + \delta E$, and then summing over all such possible energy ranges. Thus

$$Z = \sum_r e^{-\beta E_r} = \sum_E \Omega(E) e^{-\beta E} \quad (6 \cdot 6 \cdot 6)$$

The summand here is just proportional to the probability (6·2·9) that the system A has an energy between E and $E + \delta E$. Since $\Omega(E)$ increases very rapidly while $e^{-\beta E}$ decreases very rapidly with increasing E , the summand $\Omega(E)e^{-\beta E}$ exhibits a very sharp maximum at some value \bar{E} of the energy (see Fig. 6·2·2). The mean value of the energy must then be equal to \bar{E} (i.e., $\bar{E} = \bar{E}$), and the summand is only appreciable in some narrow range Δ^*E surrounding \bar{E} . The subsequent argument is similar to that used in (3·7·17). The sum in (6·6·6) must be equal to the value $\Omega(\bar{E})e^{-\beta\bar{E}}$ of the summand at

its maximum multiplied by a number of the order of $(\Delta^*E/\delta E)$, this being the number of energy intervals δE contained in the range Δ^*E . Thus

$$Z = \Omega(\bar{E}) e^{-\beta \bar{E}} \frac{\Delta^*E}{\delta E}$$

and

$$\ln Z = \ln \Omega(\bar{E}) - \beta \bar{E} + \ln \frac{\Delta^*E}{\delta E}$$

But, if the system has f degrees of freedom, the last term on the right is at most of the order of $\ln f$ and is thus utterly negligible compared to the other terms which are of the order of f . Hence

$$\ln Z = \ln \Omega(\bar{E}) - \beta \bar{E} \quad (6 \cdot 6 \cdot 7)$$

so that (6·6·5) reduces indeed to

$$S = k \ln \Omega(\bar{E}) \quad (6 \cdot 6 \cdot 8)$$

Since $k\beta = T^{-1}$, (6·6·5) can be written in the form

$$TS = kT \ln Z + \bar{E}$$

or



$$F \equiv \bar{E} - TS = -kT \ln Z \quad (6 \cdot 6 \cdot 9)$$

Thus $\ln Z$ is very simply related to the Helmholtz free energy F already encountered in (5·5·12). Indeed, the relations (6·5·12) and (6·5·4) expressing \bar{p} and \bar{E} in terms of derivatives of $\ln Z$ are equivalent to the relations (5·5·14) expressing \bar{p} and S in terms of derivatives of F . They express a connection between these macroscopic quantities and the partition function Z , which is calculable from microscopic information about the system. They are thus analogous to the relations (3·12·1) or (3·12·5) which connect T and \bar{E} with the quantity $\ln \Omega$ or S .

Let us examine the partition function (6·5·3) in the limit as $T \rightarrow 0$ or $\beta \rightarrow \infty$. Then the only terms of appreciable magnitude in the sum are those with the lowest possible value of the energy E_r , i.e., the Ω_0 states corresponding to the ground state energy E_0 . Hence

as $T \rightarrow 0$,

$$Z \rightarrow \Omega_0 e^{-\beta E_0}$$

In this limit the mean energy $\bar{E} \rightarrow E_0$, and the entropy S defined in (6·6·5) becomes,

as $T \rightarrow 0$,

$$S \rightarrow k[(\ln \Omega_0 - \beta E_0) + \beta E_0] = k \ln \Omega_0 \quad (6 \cdot 6 \cdot 10)$$

Thus we regain the statement (known as the "third law of thermodynamics") that the entropy has the limiting property already discussed in Sec. 3·10; i.e., the entropy approaches a value (equal to zero in the absence of randomness of nuclear spin orientations) independent of all parameters of the system.

Suppose that one is dealing with a system $A^{(0)}$ consisting of two systems A and A' which are weakly interacting with each other. Let each state of A be denoted by an index r and its corresponding energy by E_r . Similarly, let

each state of A' be denoted by an index s and its corresponding energy by E_s' . A state of the combined system $A^{(0)} = A + A'$ can then be denoted by the pair of indices r,s ; since A and A' interact only weakly, the corresponding energy of this state is simply given by

$$E_{rs}^{(0)} = E_r + E_s' \quad (6 \cdot 6 \cdot 11)$$

The partition function of $A^{(0)}$ is then, by definition,

$$\begin{aligned} Z^{(0)} &= \sum_{r,s} e^{-\beta E_{rs}^{(0)}} \\ &= \sum_{r,s} e^{-\beta(E_r + E_s')} \\ &= \sum_{r,s} e^{-\beta E_r} e^{-\beta E_s'} \\ &= \left(\sum_r e^{-\beta E_r}\right) \left(\sum_s e^{-\beta E_s'}\right) \end{aligned}$$

that is,

$$Z^{(0)} = ZZ' \quad (6 \cdot 6 \cdot 12)$$

$$\text{or} \quad \ln Z^{(0)} = \ln Z + \ln Z' \quad (6 \cdot 6 \cdot 13)$$

where Z and Z' are the partition functions of A and A' , respectively. By virtue of (6·5·4), the respective mean energies of $A^{(0)}$, A , and A' are then related by

$$\bar{E}^{(0)} = \bar{E} + \bar{E}' \quad (6 \cdot 6 \cdot 14)$$

It then also follows that the respective entropies of these systems are, by virtue of the definition (6·6·5), related by

$$S^{(0)} = S + S' \quad (6 \cdot 6 \cdot 15)$$

Hence (6·6·12) or (6·6·13) reflect the obvious fact that the extensive thermodynamic functions of two weakly interacting systems are simply additive.

Suppose, finally, that two systems A and A' are each separately in internal equilibrium with specified mean energies, or equivalently, with specified temperature parameters β and β' , respectively. Then the probability P_r of finding system A in state r and the probability P_s' of finding A' in state s are given by the canonical distributions

$$P_r = \frac{e^{-\beta E_r}}{\sum_r e^{-\beta E_r}} \quad \text{and} \quad P_s' = \frac{e^{-\beta' E_s'}}{\sum_s e^{-\beta' E_s'}} \quad (6 \cdot 6 \cdot 16)$$

If these systems are placed in thermal contact so that they interact only weakly with each other, then their respective probabilities are statistically independent and the probability P_{rs} of finding system A in state r and system s in state s is given by $P_{rs} = P_r P_s'$. Immediately after the systems are brought into thermal contact, it then follows by (6·6·16) that

$$P_{rs} = \frac{e^{-\beta E_r}}{\sum_r e^{-\beta E_r}} \frac{e^{-\beta' E_s'}}{\sum_s e^{-\beta' E_s'}} \quad (6 \cdot 6 \cdot 17)$$

If $\beta = \beta'$, this becomes simply

$$P_{rs} = \frac{e^{-\beta(E_r + E_s')}}{\sum_r \sum_s e^{-\beta(E_r + E_s')}} \quad (6 \cdot 6 \cdot 18)$$

which is the canonical distribution (corresponding to temperature β) characterizing the equilibrium of the combined system $A + A'$ whose energy levels are given by (6 · 6 · 11). Hence the systems A and A' do remain in equilibrium after being joined. On the other hand, if $\beta \neq \beta'$, then (6 · 6 · 17) does not correspond to a canonical distribution of the combined system and thus does not describe an equilibrium situation. Hence a redistribution of systems over states tends to occur until an ultimate equilibrium situation is reached where P_{rs} is given by a canonical distribution of the form (6 · 6 · 18) with some common temperature parameter β . These comments show directly that the parameter β occurring in the canonical distribution has the familiar properties of a temperature.

The discussion of this section makes it apparent that the canonical distribution implies all the thermodynamic relations already familiar from Chapter 3. The particular definition (6 · 6 · 5) of the entropy is actually quite convenient since it involves a knowledge of $\ln Z$ rather than of $\ln \Omega$. But a computation of Z by (6 · 5 · 3) is relatively simple since it involves an unrestricted sum over all states, whereas a computation of $\Omega(E)$ involves the more difficult problem of counting only those states lying between the energies E and $E + \delta E$. The definition (6 · 6 · 5) for the entropy of a system at specified temperature β has the further advantages that it does not depend, even in principle, on the size δE of any arbitrary energy interval; and that it can be used to define the entropy of an arbitrarily small system. These are distinct mathematical advantages, although the physical significance of the original (and, for large systems, equivalent) definition (6 · 6 · 8) of the entropy is more transparent.

***Remark** It is instructive to express the physical quantities of interest directly in terms of the canonical probability P_r of (6 · 5 · 1). By (6 · 5 · 3), one can write P_r in the form

$$P_r = \frac{e^{-\beta E_r}}{Z} \quad (6 \cdot 6 \cdot 19)$$

The mean energy of the system is then given by

$$\bar{E} = \Sigma P_r E_r \quad (6 \cdot 6 \cdot 20)$$

In a general quasi-static process this energy changes because both E_r and P_r change. Thus

$$d\bar{E} = \sum_r (E_r dP_r + P_r dE_r) \quad (6 \cdot 6 \cdot 21)$$

The work done by the system in this process is

$$dW = \sum_r P_r (-dE_r) = - \sum_r P_r dE_r \quad (6 \cdot 6 \cdot 22)$$

In doing work, the energy of each state, occupied with the given probability P_r , is thus simply changed by dE_r by virtue of the change of external parameters.

The heat absorbed in this process is, by definition,

$$\text{so that } dQ = d\bar{E} + dW \\ dQ = \sum_r E_r dP_r, \quad (6 \cdot 6 \cdot 23)$$

In absorbing heat, the energy of each state is thus unaffected, but its probability of occurrence is changed.

The entropy (6 · 6 · 5) can be written

$$\begin{aligned} S &= k \left[\ln Z + \beta \sum_r P_r E_r \right] \\ &= k \left[\ln Z - \sum_r P_r \ln (ZP_r) \right] \\ &= k \left[\ln Z - \ln Z \left(\sum_r P_r \right) - \sum_r P_r \ln P_r \right] \end{aligned}$$

or

► $S = -k \sum_r P_r \ln P_r \quad (6 \cdot 6 \cdot 24)$

since $\sum_r P_r = 1$

APPROXIMATION METHODS

6 · 7 Ensembles used as approximations

Suppose that one is interested in discussing an isolated system with a given number N of particles in a given volume V , the energy of the system being known to lie in the range between E and $E + \delta E$. The physical equilibrium situation is then such that the system is described statistically in terms of a microcanonical ensemble where all states in the given energy range are equally probable. If a parameter y assumes the value y_r in state r , then the mean value of \bar{y} is given by

$$\bar{y} = \frac{\sum_r y_r}{\Omega(E)} \quad (6 \cdot 7 \cdot 1)$$

Here all summations are subject to the condition that one sums *only* over those states for which the energy E_r lies in the small range

$$E < E_r < E + \delta E \quad (6 \cdot 7 \cdot 2)$$

and $\Omega(E)$ is the number of states in this particular range. The calculation of such sums and of $\Omega(E)$ may be quite difficult because of the equation of constraint (6 · 7 · 2). The trouble is that one cannot simply sum indiscriminately

over all states without restriction as we did in Sec. 6·5 in calculating mean values with the canonical distribution. Instead, one must pick out only those particular states which satisfy the restriction (6·7·2). This difficulty can, however, be readily overcome by the use of quite accurate approximation methods.

One way of circumventing the difficulties presented by the condition (6·7·2) is to replace it with the weaker condition that only the *mean* energy \bar{E} of the system is specified, with \bar{E} chosen to be equal to the given energy E . Then the canonical distribution (6·4·2) is applicable and the probability of the system being in any one of its $\Omega(E_1)$ states of energy between E_1 and $E_1 + \delta E_1$ is given by

$$P(E_1) \propto \Omega(E_1) e^{-\beta E_1} \quad (6 \cdot 7 \cdot 3)$$

Since the number of states $\Omega(E_1)$ for a large system is a very rapidly increasing function of E_1 while $e^{-\beta E_1}$ is rapidly decreasing, the expression (6·7·3) has the usual very sharp maximum at the energy $\bar{E} = E$ (see Fig. 6·2·2). Indeed, the sharpness of this maximum can be explicitly calculated by using the canonical distribution to compute the dispersion $(E_1 - \bar{E})^2$ by (6·5·8). The width Δ^*E_1 of the maximum, given by the square root of this dispersion, is very small relative to \bar{E} for a macroscopic system. (By the arguments of (3·7·14), Δ^*E/\bar{E} is ordinarily of the order of $f^{-\frac{1}{2}}$ where f is the number of degrees of freedom.) Thus, even if the energy of the system should be so precisely known that δE in (6·7·2) is very small (say $\delta E/E \approx 10^{-11}$), it is yet true that $\Delta^*E_1 < \delta E$ for a system consisting of a mole of particles. Thus values of the energy E_1 lying outside the range (6·7·2) occur with negligible probability in the canonical distribution. A specification of the mean energy \bar{E} is then almost equivalent to a specification of the total energy E by (6·7·2). Hence one expects that mean values can be computed with negligible error by using the canonical distribution; i.e., instead of (6·7·1) one can write

$$\bar{y} = \frac{\sum_r e^{-\beta E_r} y_r}{\sum_r e^{-\beta E_r}} \quad (6 \cdot 7 \cdot 4)$$

where there appears now no further complicating restriction on the domain of summation, since one sums over *all* states.

The foregoing comments can be phrased in more physical terms. If a macroscopic system A is in contact with a heat reservoir, the relative fluctuations in the energy of A are exceedingly small. Suppose now that A is removed from contact with the heat reservoir and is thermally insulated; then its total energy cannot change at all. But the distinction between this situation and the previous one is so small that it is really utterly irrelevant for most purposes; in particular, the mean values of all physical quantities (e.g., of the mean pressure or the mean magnetic moment of A) remain quite unaffected. Hence it makes no difference whether these mean values are calculated by considering

the system to be isolated so that it has equal probability of being in any one of its states of accurately specified fixed energy, or by considering it to be in contact with a heat reservoir so that it is distributed over all its states in accordance with a canonical distribution. But the latter procedure is mathematically simpler.

Calculating the dispersion $\overline{(y - \bar{y})^2}$ of some quantity y is a much more delicate matter. There is no guarantee that the dispersion is the same when calculated under conditions where E is precisely specified (i.e., $\delta E \rightarrow 0$ in (6·6·2)) or under conditions where only the mean energy \bar{E} is specified. As a matter of fact, one would expect the dispersion to be greater in the second case. In particular, if y were the energy E of the system, its dispersion would vanish in the first case where E is precisely specified, but would not vanish in the second case where only the mean value \bar{E} is specified.

When one is dealing with a macroscopic system of very precisely specified energy, the mathematical difficulties encountered in the evaluation of (6·7·1) can therefore be circumvented to excellent approximation. For purposes of calculating mean values, the situation is quite equivalent to one where the system is described by a canonical distribution with a mean energy corresponding to its actual energy.

*6·8 Mathematical approximation methods

The use of a canonical ensemble as an approximation method for handling the difficulties caused by the restrictive condition (6·7·2) can also be considered as a purely mathematical approximation method. This point of view is instructive both because it makes apparent how to find approximations for related situations, and because it permits one to make estimates of the errors involved.

To calculate physical quantities for an isolated system by the relations (3·12·1), one needs to know the function $\ln \Omega(E)$. Simply counting states is not very difficult if one can simply proceed in any order and add them up one at a time to get $1 + 1 + 1 + 1 + \dots$. But the difficulty is that among all these states one wants to count only those which have an energy E_r lying in the range

$$E < E_r < E + \delta E \quad (6\cdot8\cdot1)$$

Thus the sum to be performed is of the form

$$\Omega(E) = \sum'_r u_r, \quad u_r = 1 \text{ for all } r \quad (6\cdot8\cdot2)$$

where the prime on Σ denotes that the sum is to be performed subject to the restriction (6·8·1).

The basic problem is again that of handling the constraint (6·8·1). There are several ways of doing this conveniently.

Method 1 This is the mathematical analogue of the physical approximation used in the preceding section. By virtue of (6·8·1), the sum (6·8·2) depends on the particular energy E . If the energy of interest were not E , but E_1 , the sum would be quite different. Indeed the sum, i.e., $\Omega(E_1)$, is a very rapidly increasing function of E_1 . We wish to calculate it for the particular value $E_1 = E$. We can exploit the rapidly increasing property of the sum $\Omega(E_1)$ by noting that multiplication by the rapidly decreasing function $e^{-\beta E_1}$ produces a function $\Omega(E_1) e^{-\beta E_1}$ with a very sharp maximum near some value $E_1 = \tilde{E}_1$. Here β is some arbitrary positive parameter which (for the time being) has no connection whatever with temperature. By proper choice of β one can make the maximum occur at the desired value $\tilde{E}_1 = E$; one need only choose β so that

$$\frac{\partial}{\partial E_1} \ln [\Omega(E_1) e^{-\beta E_1}] = \frac{\partial \ln \Omega}{\partial E_1} - \beta = 0 \quad (6 \cdot 8 \cdot 3)$$

when $E_1 = E$.

The sharp maximum property of $\Omega(E_1) e^{-\beta E_1}$ implies that when this quantity is summed indiscriminately over all possible energies E_1 , only those terms in some narrow range Δ^*E near E will contribute appreciably. Thus one selects only those terms of interest; i.e.,

$$\sum_{E_1} \Omega(E_1) e^{-\beta E_1} = \Omega(E) e^{-\beta E} K, \quad K \equiv \frac{\Delta^*E}{\delta E_1}$$

where the sum is expressed in terms of the value of the summand at the maximum, multiplied by the number K of terms in the sum contained in the range Δ^*E (see Fig. 6·2·2). Taking logarithms, one gets

$$\ln \left[\sum_{E_1} \Omega(E_1) e^{-\beta E_1} \right] = \ln \Omega(E) - \beta E$$

since $\ln K$ is utterly negligible compared to the other terms. Hence

$$\blacktriangleright \quad \ln \Omega(E) = \ln Z + \beta E \quad (6 \cdot 8 \cdot 4)$$

where $Z \equiv \sum_{E_1} \Omega(E_1) e^{-\beta E_1} = \sum_r e^{-\beta E_r}$ (6·8·5)

The last form on the right is obtained by summing over all individual states, whereas in the first sum one first groups together all terms with a given energy E and then sums over all energies. The relation (6·8·4) represents the desired approximate evaluation of $\ln \Omega$ in terms of the *unrestricted* sum Z over *all* states.

The parameter β is to be determined by the maximum condition (6·8·3) which is an equation expressing β in terms of E . Thus Z is a function of E through its dependence on β . By (6·8·4) the condition (6·8·3) becomes for $E_1 = E$

$$\left[\frac{\partial \ln Z}{\partial \beta} \frac{\partial \beta}{\partial E} + \left(E \frac{\partial \beta}{\partial E} + \beta \right) \right] - \beta = 0$$

or
$$\frac{\partial \ln Z}{\partial \beta} + E = 0 \quad (6 \cdot 8 \cdot 6)$$

Using (6·8·5), this equation for determining β is simply

$$\frac{\sum_r e^{-\beta E_r} E_r}{\sum_r e^{-\beta E_r}} = E \quad (6 \cdot 8 \cdot 7)$$

It is clear from (6·8·3) that the parameter β introduced in this approximation method is just the temperature of the system. Similarly, the entropy can be calculated by (6·8·4) as

$$S = k \ln \Omega = k(\ln Z + \beta E)$$

where the sum Z defined in (6·8·5) is simply the partition function already encountered in (6·5·3).

Method 2 It is possible to handle the restrictive condition in the sum (6·8·2) in a very straightforward fashion by a method similar to that used in Sec. 1·10. Let us shift the complication introduced by the restriction from the summation to the summand by multiplying each term in the sum by the function $\delta(E_r - E)\delta E$, which is equal to unity whenever E_r lies in a range δE about E , but which vanishes otherwise. By Appendix A·7 the function $\delta(E_r - E)$ is just the Dirac δ function. Then one can write

$$\Omega(E) = \sum_r \delta(E_r - E) \delta E \quad (6 \cdot 8 \cdot 8)$$

where the sum is now over all states *without* any restriction, but where the δ function in the summand guarantees that only those terms in (6·8·8) which satisfy the condition (6·8·1) contribute to the sum.

But at this point one can make use of the simple analytic representation for the δ function given in (A·7·16). Thus

$$\delta(E - E_r) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\beta' e^{i(E-E_r)\beta'} e^{(E-E_r)\beta}$$

or in more compact form

$$\delta(E - E_r) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\beta' e^{(E-E_r)\underline{\beta}} \quad (6 \cdot 8 \cdot 9)$$

where

$$\underline{\beta} \equiv \beta + i\beta' \quad (6 \cdot 8 \cdot 10)$$

is complex, the integration is only over its imaginary part, and β is an arbitrary parameter which can be chosen at will.

The sum (6·8·8) is now readily evaluated. One has simply

$$\Omega(E) = \frac{\delta E}{2\pi} \sum_r \int_{-\infty}^{\infty} d\beta' e^{(E-E_r)\underline{\beta}}$$

or $\Omega(E) = \frac{\delta E}{2\pi} \int_{-\infty}^{\infty} d\beta' e^{\underline{\beta} E} Z(\underline{\beta}) \quad (6 \cdot 8 \cdot 11)$

where $Z(\underline{\beta}) \equiv \sum_r e^{-\underline{\beta} E_r} = \sum_r e^{-(\beta+i\beta') E_r} \quad (6 \cdot 8 \cdot 12)$

This last sum is over all states without restriction and is thus relatively simple to evaluate.

These results are *exact*. We note that if $\beta' = 0$, all the terms in the sum (6·8·12) are positive. On the other hand, if $\beta' \neq 0$, the oscillatory factors $e^{-\beta'E}$ cause the terms in the sum not to add in phase, but to add with more or less random signs in the real and imaginary parts. Since there are so very many terms in the sum, the result is that the absolute value $|e^{\beta E}Z(\beta)|$ is *very* much larger for $\beta' = 0$ than for $\beta' \neq 0$. Because of this very sharp maximum, only the region of integration near $\beta' = 0$ contributes appreciably to the integral (6·8·11). Hence we expect that the integral is very well approximated by

$$\Omega(E) = K' e^{\beta E} Z(\beta) \quad (6 \cdot 8 \cdot 13)$$

where K' is some constant which is certainly small compared to the number of degrees of freedom. Thus

$$\ln \Omega(E) = \beta E + \ln Z(\beta) \quad (6 \cdot 8 \cdot 14)$$

since $\ln K'$ is negligibly small compared to the other terms which are of order f . Thus we regain the result (6·8·4).

It is worth doing the argument leading to (6·8·13) more carefully. Because the integrand in (6·8·11) is only appreciable for $\beta' \approx 0$, one can in the significant domain of integration expand its logarithm in a power series about $\beta' = 0$. Thus

$$\begin{aligned} \ln [e^{\beta E} Z(\beta)] &= \beta E + \ln Z(\beta) \\ &= (\beta + i\beta')E + \ln Z(\beta) + B_1(i\beta') + \frac{1}{2}B_2(i\beta')^2 + \dots \end{aligned}$$

or

$$\ln [e^{\beta E} Z(\beta)] = \beta E + \ln Z(\beta) + i(E + B_1)\beta' - \frac{1}{2}B_2\beta'^2 + \dots \quad (6 \cdot 8 \cdot 15)$$

$$\text{where } B_k = \left[\frac{\partial^k \ln Z}{\partial \beta^k} \right]_{\beta'=0} = \frac{\partial^k \ln Z}{\partial \beta^k} \quad (6 \cdot 8 \cdot 16)$$

Hence

$$e^{\beta E} Z(\beta) = e^{\beta E} Z(\beta) e^{-\frac{1}{2}B_2\beta'^2} e^{i(E+B_1)\beta'} \quad (6 \cdot 8 \cdot 17)$$

The parameter β is still at our disposal, and we can choose it so as to optimize our approximation. Irrespective of the choice of β , we already know (and (6·8·17) shows this explicitly) that $|e^{\beta E}Z(\beta)|$ is always maximum for $\beta' = 0$. We should like the integrand to contribute most significantly to the integral in the immediate vicinity of $\beta' = 0$ where the expansion (6·8·15) is most nearly valid. Because of the oscillatory behavior of the integrand $e^{\beta E}Z(\beta)$ caused by the imaginary part β' , this integrand contributes to the integral most importantly in that region where it oscillates *least* rapidly, i.e., where

$$\frac{\partial}{\partial \beta'} [e^{\beta E} Z(\beta)] = 0$$

so that the integrand is stationary with respect to the phase β' . Choosing the region of least rapid oscillation to lie at $\beta' = 0$ means choosing β so that

$$\begin{aligned} E + B_1 &= 0 \\ \text{or} \quad E + \frac{\partial \ln Z}{\partial \beta} &= 0 \end{aligned} \quad (6 \cdot 8 \cdot 18)$$

Then (6·8·17) reduces to

$$e^{\beta E} Z(\beta) = e^{\beta E} Z(\beta) e^{-\frac{1}{2} B_2 \beta^2} \quad (6 \cdot 8 \cdot 19)$$

The argument which led us to expect a sharp maximum of $e^{\beta E} Z(\beta)$ at $\beta' = 0$ implies that B_2 must be such that $B_2 \gg 1$. Hence (6·8·11) becomes simply

$$\begin{aligned} \Omega(E) &= \frac{\delta E}{2\pi} e^{\beta E} Z(\beta) \int_{-\infty}^{\infty} d\beta' e^{-\frac{1}{2} B_2 \beta'^2} \\ \text{or} \quad \Omega(E) &= e^{\beta E} Z(\beta) \frac{\delta E}{\sqrt{2\pi B_2}} \end{aligned} \quad (6 \cdot 8 \cdot 20)$$

Thus

$$\ln \Omega(E) \approx \beta E + \ln Z$$

These are the results (6·8·13) and (6·8·14). Note also that the condition (6·8·18) which determines β is the same as that of (6·8·6), i.e., it is again equivalent to (6·8·7).*

GENERALIZATIONS AND ALTERNATIVE APPROACHES

*6·9 Grand canonical and other ensembles

System with an indefinite number of particles The discussion of the last few sections can be readily generalized to a variety of other situations. Consider, for example, the case where a system A of fixed volume V is in contact with a large reservoir A' with which it can exchange not only energy, but particles (see Fig. 6·9·1). Then neither the energy E of A nor the number N of particles in A are fixed, but the *total* energy $E^{(0)}$ and the total number of particles $N^{(0)}$ of the combined system $A^{(0)} \equiv A + A'$ are fixed; i.e.,

$$\begin{aligned} E + E' &= E^{(0)} = \text{constant} \\ N + N' &= N^{(0)} = \text{constant} \end{aligned} \quad (6 \cdot 9 \cdot 1)$$

where E' and N' denote the energy and number of particles in the reservoir A' . In this situation one can ask for the probability in the ensemble of finding the

* This method based on the integral (6·8·11) and its approximate evaluation by the method of stationary phase is equivalent to the so-called "Darwin-Fowler" method which employs contour integration in the complex plane and the method of steepest descents. See, for example, R. H. Fowler, "Statistical Mechanics," 2d ed., chap. 2, Cambridge University Press, Cambridge, 1955, or E. Schrödinger, "Statistical Thermodynamics," 2d ed., chap. 6, Cambridge University Press, Cambridge, 1952.

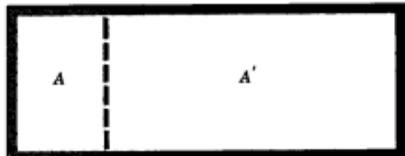


Fig. 6·9·1 A small system A separated from a much larger system A' by a perforated partition. The systems can exchange both energy and particles.

system A in any one particular state r where it contains N_r particles and has an energy E_r .

The argument which answers this question is identical to that of Sec. 6·2. Let $\Omega'(E', N')$ denote the number of states accessible to the reservoir A' when it contains N' particles and has an energy in the range near E' . If A is in the particular state r , the number of states accessible to the combined system $A^{(0)}$ is just the number of states accessible to the reservoir. The probability P_r of finding A in this state is then proportional to this; i.e.,

$$P_r(E_r, N_r) \propto \Omega'(E^{(0)} - E_r, N^{(0)} - N_r) \quad (6\cdot9\cdot2)$$

where we have used the conservation equations (6·9·1). Since A is very small compared to A' , $E_r \ll E^{(0)}$ and $N_r \ll N^{(0)}$. Thus

$$\ln \Omega'(E^{(0)} - E_r, N^{(0)} - N_r) = \ln \Omega'(E^{(0)}, N^{(0)}) - \left[\frac{\partial \ln \Omega}{\partial E'} \right]_0 E_r - \left[\frac{\partial \ln \Omega}{\partial N'} \right]_0 N_r$$

Here the derivatives are evaluated for $E' = E^{(0)}$ and $N' = N^{(0)}$; they are, therefore, constants characterizing the reservoir A' . Denote them by

$$\beta \equiv \left[\frac{\partial \ln \Omega}{\partial E'} \right]_0 \quad \text{and} \quad \alpha \equiv \left[\frac{\partial \ln \Omega}{\partial N'} \right]_0 \quad (6\cdot9\cdot3)$$

Then $\Omega'(E^{(0)} - E_r, N^{(0)} - N_r) = \Omega'(E^{(0)}, N^{(0)}) e^{-\beta E_r - \alpha N_r}$, and

$$\blacktriangleright \qquad P_r \propto e^{-\beta E_r - \alpha N_r} \quad (6\cdot9\cdot4)$$

This is called the "grand canonical" distribution. An ensemble of systems distributed according to this probability distribution is called a "grand canonical ensemble." The parameter β is, by (6·9·3), the temperature parameter of the reservoir; thus $T \equiv (k\beta)^{-1}$ is the absolute temperature of the reservoir. The quantity $\mu \equiv -kT\alpha$ is called the "chemical potential" of the reservoir.

It is obvious from the discussion of Sec. 6·4 that if one considers a physical situation where only the *mean* energy \bar{E} and the *mean* number \bar{N} of particles of a system A are known, the distribution over systems in the ensemble is again described by a grand canonical distribution of the form (6·9·4). But then the parameters β and α no longer characterize any reservoir. Instead, they are to be determined by the conditions that the system A has the specified

mean energy \bar{E} and mean number \bar{N} of particles, i.e., by the equations

$$\bar{E} = \frac{\sum_r e^{-\beta E_r - \alpha N_r} E_r}{\sum_r e^{-\beta E_r - \alpha N_r}}$$

$$\bar{N} = \frac{\sum_r e^{-\beta E_r - \alpha N_r} N_r}{\sum_r e^{-\beta E_r - \alpha N_r}} \quad (6 \cdot 9 \cdot 5)$$

Here the sums are over all possible states of the system A irrespective of its number of particles or of its energy.

When A is a macroscopic system in contact with a reservoir as illustrated in Fig. 6·9·1, it is again clear that the relative fluctuations of its energy about its mean energy \bar{E} , and of its number of particles about its mean number \bar{N} , are very small. Thus the physical properties of A would not be appreciably affected if it were removed from contact with the reservoir so that both its energy and number of particles would be rigorously fixed. Thus, for purposes of calculating mean values of physical quantities, it makes no noticeable difference whether a macroscopic system is isolated, or in contact with a reservoir with which it can only exchange energy, or in contact with a reservoir with which it can exchange both energy and particles. Hence these mean values can equally well be calculated by considering the system to be distributed with equal probability over all states of given energy and number of particles (microcanonical distribution), or to be distributed according to the canonical distribution (6·2·7) over all its states with a given number of particles irrespective of energy, or to be distributed according to the grand canonical distribution (6·9·4) over all its states irrespective of energy and number. In some problems where the constraint of a fixed number of particles is cumbersome, one can thus readily circumvent the complication by approximating the actual situation with one where only the mean number of particles is fixed, i.e., by using the grand canonical distribution (6·9·4). This is sometimes a useful procedure in practical calculations.

System in macroscopic motion Up to now we have always been careful to satisfy the condition of conservation of total energy for an isolated system. But what about other constants of motion like the total linear momentum or total angular momentum? The reason that we have not paid attention to these quantities is that we have always effectively considered the system A of interest to be enclosed in a container A' of very large mass. This container can take up arbitrary amounts of momentum from the system A with negligible effect on the velocity of its center of mass. The system A can thus have arbitrary amounts of momentum, and one need not be concerned about satisfying any momentum conservation conditions for it. What is effectively specified in the problem is then the velocity v_0 of the container A' , and we have chosen $v_0 = 0$ with respect to the laboratory. The system A itself can then

have arbitrary momentum; the condition of equilibrium is only that its *mean* velocity be the same as the specified velocity \mathbf{v}_0 of the container.

These comments show that the system A' acts like a momentum reservoir with a mass M' much larger than that of A . The analogy to the case of energy reservoirs discussed in Sec. 6·2 is apparent. It is, indeed, of some interest to discuss briefly the situation where the combined system $A^{(0)} \equiv A + A'$ is in macroscopic motion with respect to the laboratory. Consider the case where A can exchange both energy and momentum with the much larger system A' . If A is in a state r where its total energy is ϵ_r and its momentum is \mathbf{p}_r , then the conservation conditions for the combined system $A^{(0)}$ of total energy ϵ_0 and momentum \mathbf{p}_0 are

$$\begin{aligned}\epsilon_r + \epsilon' &= \epsilon_0 = \text{constant} \\ \mathbf{p}_r + \mathbf{p}' &= \mathbf{p}_0 = \text{constant}\end{aligned}\quad (6 \cdot 9 \cdot 6)$$

Here ϵ' denotes the total energy and \mathbf{p}' the total momentum of the reservoir A' .

Up to now we have always considered systems whose center of mass is at rest with respect to the laboratory; then the total energy ϵ of a system consists only of the internal energy E of particle motion with respect to the center of mass. In the present problem the situation is different. The number of states $\Omega'(E')$ accessible to A' depends on its internal energy E' with respect to its center of mass. Since the latter moves with velocity \mathbf{p}'/M' , the internal energy of A' differs from its total energy ϵ' by the macroscopic kinetic energy of center-of-mass motion. Thus

$$E' = \epsilon' - \frac{\mathbf{p}'^2}{2M'} \quad (6 \cdot 9 \cdot 7)$$

When A is in state r , it follows by (6·9·6) that the internal energy of A' is

$$\begin{aligned}E' &= \epsilon_0 - \epsilon_r - \frac{1}{2M'} (\mathbf{p}_0 - \mathbf{p}_r)^2 \\ &= \epsilon_0 - \epsilon_r - \frac{\mathbf{p}_0^2}{2M'} + \frac{\mathbf{p}_0 \cdot \mathbf{p}_r}{M'} - \frac{\mathbf{p}_r^2}{2M'} \\ \text{or } E' &\approx \left(\epsilon_0 - \frac{\mathbf{p}_0^2}{2M'} \right) - (\epsilon_r - \mathbf{v}_0 \cdot \mathbf{p}_r)\end{aligned}\quad (6 \cdot 9 \cdot 8)$$

Since M' is very large, we have neglected the term \mathbf{p}_r^2/M' . Also M' is then nearly the mass of the combined system $A^{(0)}$ so that $\mathbf{v}_0 = \mathbf{p}_0/M'$ is the velocity of the center of mass of the total system $A^{(0)}$ (or equivalently, of A').

The probability P_r that A is in state r is

$$P_r \propto \Omega'(E')$$

with E' given by (6·9·8). Expanding $\ln \Omega'(E')$ in the usual way this becomes

$$\blacktriangleright P_r \propto e^{-\beta(\epsilon_r - \mathbf{v}_0 \cdot \mathbf{p}_r)} \quad (6 \cdot 9 \cdot 9)$$

where $\beta = \partial \ln \Omega'/\partial E'$ is the temperature parameter of the reservoir evaluated when its internal energy $E' = \epsilon_0 - p_0^2/2M'$.

Example Consider a molecule A in an ideal gas A' , the center of mass of the whole gas moving with constant velocity v_0 . Suppose that A is in a state with momentum between p and $p + dp$, or velocity between v and $v + dv$, where $p = mv$ and m is the mass of the molecule. Then

$$\epsilon_r - v_0 \cdot p_r = \frac{1}{2}mv^2 - v_0 \cdot mv = \frac{1}{2}m(v - v_0)^2 = \frac{1}{2}mv_0^2$$

Since v_0 is just a constant, it follows by (6.9.9) that the probability of the molecule's velocity being in the range between v and $v + dv$ is simply

$$P(v) d^3v \propto e^{-\frac{1}{2}\theta m(v - v_0)^2} d^3v$$

This is, of course, what one would expect. The molecule has simply a Maxwellian velocity distribution relative to the frame of reference moving with the constant velocity v_0 .

*6.10 Alternative derivation of the canonical distribution

The canonical distribution is so important that it is worth deriving it by an alternative method. Although this derivation is more cumbersome than the one given in Sec. 6.4, it has some instructive features.

We use the notation introduced in Sec. 6.4 and consider a system A of constant specified mean energy \bar{E} . The representative ensemble is supposed to consist of a very large number a of such systems, a_r of which are in state r . Then we know that

$$\sum_r a_r = a \quad (6.10.1)$$

$$\text{while} \quad \frac{1}{a} \sum_r a_r E_r = \bar{E} \quad (6.10.2)$$

The number $\Gamma(a_1, a_2, \dots)$ of distinct possible ways of selecting a total of a distinct systems in such a way that a_1 of them are in state $r = 1$, a_2 in state $r = 2$ etc., is given by the same combinatorial reasoning as that used in Sec. 1.2, i.e., by

$$\Gamma = \frac{a!}{a_1! a_2! a_3! \dots} \quad (6.10.3)$$

$$\text{Thus} \quad \ln \Gamma = \ln a! - \sum_r \ln a_r! \quad (6.10.4)$$

Since the representative ensemble is supposed to consist of a very large number of systems, all the numbers a and a_r are very large so that Stirling's approximation can be used in its simplest form (A.6.2):

$$\ln a_r! = a_r \ln a_r - a_r$$

Hence (6 · 10 · 4) becomes

$$\ln \Gamma = a \ln a - a - \sum_r a_r \ln a_r + \sum_r a_r$$

or $\ln \Gamma = a \ln a - \Sigma a_r \ln a_r \quad (6 \cdot 10 \cdot 5)$

where we have used (6 · 10 · 1). We can now ask: For what distribution of systems over the possible states will the total number Γ of possible ways of achieving this distribution be a maximum? That is, for what set of numbers a_1, a_2, a_3, \dots , subject to conditions (6 · 10 · 1) and (6 · 10 · 2), will Γ (or $\ln \Gamma$) be a maximum?

The condition that $\ln \Gamma$ has an extremum is that for small changes* δa_r of the various numbers there is no change in Γ . Thus we require that

$$\delta \ln \Gamma = - \sum_r (\delta a_r + \ln a_r \delta a_r) = 0 \quad (6 \cdot 10 \cdot 6)$$

subject to the conditions (6 · 10 · 1) and (6 · 10 · 2), i.e.,

$$\sum_r \delta a_r = 0 \quad (6 \cdot 10 \cdot 7)$$

and

$$\sum_r E_r \delta a_r = 0 \quad (6 \cdot 10 \cdot 8)$$

By virtue of (6 · 10 · 7) the condition (6 · 10 · 6) becomes

$$\sum_r \ln a_r \delta a_r = 0 \quad (6 \cdot 10 \cdot 9)$$

In (6 · 10 · 9) all the changes δa_r are not independent of each other since they must satisfy the equations of constraint (6 · 10 · 7) and (6 · 10 · 8). The situation is most expeditiously handled by the method of Lagrange multipliers (see Appendix A · 10). Thus one can multiply (6 · 10 · 7) by a parameter α and (6 · 10 · 8) by a parameter β and then add these equations to (6 · 10 · 9) to obtain

$$\sum_r (\ln a_r + \alpha + \beta E_r) \delta a_r = 0 \quad (6 \cdot 10 \cdot 10)$$

With a proper choice of α and β , all the δa_r in (6 · 10 · 10) can now be regarded as independent. Hence we can conclude that each coefficient of δa_r must separately vanish. If we denote by \tilde{a}_r the value of a_r when Γ is maximum, we have then

$$\ln \tilde{a}_r + \alpha + \beta E_r = 0$$

or $\tilde{a}_r = e^{-\alpha} e^{-\beta E_r} \quad (6 \cdot 10 \cdot 11)$

Here the parameter α is to be determined by the normalization condition

* The numbers a_r are integers, but they are so large that even a change δa_r of many integer units is negligibly small compared to a_r itself and can be considered infinitesimal. Thus differential calculus methods are applicable.

(6·10·1); i.e.,

$$e^{-\alpha} = (\sum e^{-\beta E_r})^{-1}$$

The parameter β is to be determined by the condition (6·10·2), i.e., by the relation

$$\frac{\sum e^{-\beta E_r} E_r}{\sum e^{-\beta E_r}} = \bar{E} \quad (6 \cdot 10 \cdot 12)$$

Putting

► $\tilde{P}_r \equiv \frac{\tilde{a}_r}{a} = \frac{e^{-\beta E_r}}{\sum_r e^{-\beta E_r}}$ (6·10·13)

we have in (6·10·13) regained the canonical distribution (6·4·2) as corresponding to that distribution of systems in the ensemble which makes the number Γ of possible configurations a maximum.

In terms of the probability $P_r = a_r/a$, the expression (6·10·5) for Γ can be written

$$\begin{aligned} \ln \Gamma &= a \ln a - \sum_r a P_r \ln (a P_r) \\ &= a \ln a - a \sum_r P_r (\ln a + \ln P_r) \\ &= a \ln a - a \ln a (\sum_r P_r) - a \sum_r P_r \ln P_r, \end{aligned}$$

or

$$\ln \Gamma = -a \sum_r P_r \ln P_r, \quad (6 \cdot 10 \cdot 14)$$

since $\sum P_r = 1$. (Note that the right side is properly positive, since $0 < P_r < 1$, so that $\ln P_r < 0$.)

Hence the canonical distribution P_r is characterized by the fact that it makes the quantity $-\sum_r P_r \ln P_r$ a maximum subject to a given value $\sum_r P_r E_r = \bar{E}$ of the mean energy. Comparison with the expression (6·6·24) for the entropy in terms of the canonical distribution shows that

$$\ln \Gamma = \frac{a}{k} S \quad (6 \cdot 10 \cdot 15)$$

An increase in $\ln \Gamma$ reflects a more random distribution of systems over the available states, i.e., a loss of specific information about the distribution of systems over the available states. The maximum possible value of Γ gives, by (6·10·15), the entropy of the final equilibrium state.

The quantity $-\ln \Gamma$, i.e., the function $\sum_r P_r \ln P_r$, can be used as a measure of nonrandomness, or information, available about systems in the ensemble. This function plays a key role as a measure of information in problems of communication and general "information theory."*

* See, for example, L. Brillouin, "Science and Information Theory," 2d ed., Academic Press, New York, 1962; or J. R. Pierce, "Symbols, Signals, and Noise," Harper, New York, 1961. Statistical mechanics is considered from the point of view of information theory by E. T. Jaynes in *Phys. Rev.*, vol. 106, p. 620 (1957).

Remark One can easily check that Γ is indeed a *maximum* by expanding (6·10·5) about the value \bar{a}_r . Thus

$$\ln \Gamma = a \ln a - \Sigma (\bar{a}_r + \delta a_r) \ln (\bar{a}_r + \delta a_r) \quad (6 \cdot 10 \cdot 16)$$

But

$$\ln (\bar{a}_r + \delta a_r) = \ln \bar{a}_r + \ln \left(1 + \frac{\delta a_r}{\bar{a}_r} \right) \approx \ln \bar{a}_r + \frac{\delta a_r}{\bar{a}_r} - \frac{1}{2} \left(\frac{\delta a_r}{\bar{a}_r} \right)^2 \dots$$

Then (6·10·16) becomes

$$\ln \Gamma = a \ln a - \sum_r \bar{a}_r \ln \bar{a}_r - \sum_r (1 + \ln \bar{a}_r) \delta a_r - \sum_r \frac{1}{2} \frac{(\delta a_r)^2}{\bar{a}_r}$$

The terms in δa_r vanish, as they must for an extremum, since

$$\Sigma \ln \bar{a}_r \delta a_r = -\Sigma (\alpha + \beta E_r) \delta a_r = 0$$

by (6·10·7) and (6·10·8). Hence one is left with

$$\ln \Gamma = \left(a \ln a - \sum_r \bar{a}_r \ln \bar{a}_r \right) - \frac{1}{2} \sum_r \frac{(\delta a_r)^2}{\bar{a}_r}$$

or

$$\Gamma = \bar{\Gamma} \exp \left[-\frac{1}{2} \sum_r \frac{(\delta a_r)^2}{\bar{a}_r} \right] \quad (6 \cdot 10 \cdot 17)$$

The argument of the exponential function is a very large positive number unless practically all the δa_r vanish. Hence Γ exhibits a very sharp maximum.

SUGGESTIONS FOR SUPPLEMENTARY READING

- C. Kittel: "Elementary Statistical Physics," secs. 11–14, John Wiley & Sons, Inc., New York, 1958.
- T. L. Hill: "An Introduction to Statistical Thermodynamics," chaps. 1 and 2, Addison-Wesley Publishing Company, Reading, Mass., 1960.
- E. Schrödinger: "Statistical Thermodynamics," 2d ed., chaps. 2 and 6, Cambridge University Press, Cambridge, 1952.
- R. Becker: "Theorie der Wärme," secs. 36–41, 46, Springer-Verlag, Berlin, 1955.

PROBLEMS

- 6.1** A simple harmonic one-dimensional oscillator has energy levels given by $E_n = (n + \frac{1}{2})\hbar\omega$, where ω is the characteristic (angular) frequency of the oscillator and where the quantum number n can assume the possible integral values $n = 0, 1, 2, \dots$. Suppose that such an oscillator is in thermal contact with a heat reservoir at temperature T low enough so that $kT/(\hbar\omega) \ll 1$.
- (a) Find the ratio of the probability of the oscillator being in the first excited state to the probability of its being in the ground state.

- (b) Assuming that only the ground state and first excited state are appreciably occupied, find the mean energy of the oscillator as a function of the temperature T .
- 6.2** Consider again the system of Problem 3.2, i.e., N weakly interacting particles, each of spin $\frac{1}{2}$ and magnetic moment μ , located in an external field H . Suppose that this system is in thermal contact with a heat reservoir at the absolute temperature T . Calculate its mean energy \bar{E} as a function of T and H . Compare the result with the answer to Problem 3.2a.
- 6.3** A solid at absolute temperature T is placed in an external magnetic field $H = 30,000$ gauss. The solid contains weakly interacting paramagnetic atoms of spin $\frac{1}{2}$ so that the energy of each atom is $\pm \mu H$.
- (a) If the magnetic moment μ is equal to one Bohr magneton, i.e., $\mu = 0.927 \times 10^{-20}$ ergs/gauss, below what temperature must one cool the solid so that more than 75 percent of the atoms are polarized with their spins parallel to the external magnetic field?
- (b) Suppose that one considered instead a solid which is free of paramagnetic atoms but contains many protons (e.g., paraffin). Each proton has spin $\frac{1}{2}$ and a magnetic moment $\mu = 1.41 \times 10^{-23}$ ergs/gauss. Below what temperature must one cool this solid so that more than 75 percent of the protons have their spins aligned parallel to the external magnetic field?
- 6.4** A sample of mineral oil is placed in an external magnetic field H . Each proton has spin $\frac{1}{2}$ and a magnetic moment μ ; it can, therefore, have two possible energies $\epsilon = \mp \mu H$, corresponding to the two possible orientations of its spin. An applied radio-frequency field can induce transitions between these two energy levels if its frequency ν satisfies the Bohr condition $h\nu = 2\mu H$. The power absorbed from this radiation field is then proportional to the difference in the number of nuclei in these two energy levels. Assume that the protons in the mineral oil are in thermal equilibrium at a temperature T which is so high that $\mu H \ll kT$. How does the absorbed power depend on the temperature T of the sample?
- 6.5** Consider an ideal gas at the absolute temperature T in a uniform gravitational field described by acceleration g . By writing the condition of hydrostatic equilibrium for a slice of the gas located between heights z and $z + dz$, derive an expression for $n(z)$, the number of molecules per cm^3 at height z . Compare this with Eq. (6.3.20), which was derived from statistical mechanics.
- 6.6** A system consists of N weakly interacting particles, each of which can be in either of two states with respective energies ϵ_1 and ϵ_2 , where $\epsilon_1 < \epsilon_2$.
- (a) Without explicit calculation, make a qualitative plot of the mean energy \bar{E} of the system as a function of its temperature T . What is \bar{E} in the limit of very low and very high temperatures? Roughly near what temperature does \bar{E} change from its low to its high temperature limiting values?
- (b) Using the result of (a), make a qualitative plot of the heat capacity C_V (at constant volume) as a function of the temperature T .
- (c) Calculate explicitly the mean energy $\bar{E}(T)$ and heat capacity $C_V(T)$ of this system. Verify that your expressions exhibit the qualitative features discussed in (a) and (b).
- 6.7** The nuclei of atoms in a certain crystalline solid have spin one. According to quantum theory, each nucleus can therefore be in any one of three quantum states labeled by the quantum number m , where $m = 1, 0$, or -1 . This quantum number measures the projection of the nuclear spin along a crystal axis of

the solid. Since the electric charge distribution in the nucleus is not spherically symmetrical, but ellipsoidal, the energy of a nucleus depends on its spin orientation with respect to the internal electric field existing at its location. Thus a nucleus has the same energy $E = \epsilon$ in the state $m = 1$ and the state $m = -1$, compared with an energy $E = 0$ in the state $m = 0$.

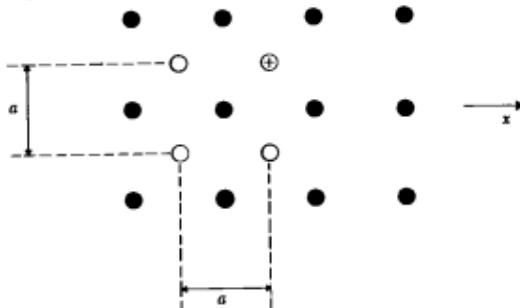
(a) Find an expression, as a function of absolute temperature T , of the nuclear contribution to the molar internal energy of the solid.

(b) Find an expression, as a function of T , of the nuclear contribution to the molar entropy of the solid.

(c) By directly counting the total number of accessible states, calculate the nuclear contribution to the molar entropy of the solid at very low temperatures. Calculate it also at very high temperatures. Show that the expression in part (b) reduces properly to these values as $T \rightarrow 0$ and $T \rightarrow \infty$.

(d) Make a qualitative graph showing the temperature dependence of the nuclear contribution to the molar heat capacity of the solid. Calculate its temperature dependence explicitly. What is its temperature dependence for large values of T ?

- 6.8** The following describes a simple two-dimensional model of a situation of actual physical interest. A solid at absolute temperature T contains N negatively charged impurity ions per cm^3 , these ions replacing some of the ordinary atoms of the solid. The solid as a whole is, of course, electrically neutral. This is so because each negative ion with charge $-e$ has in its vicinity one positive ion with charge $+e$. The positive ion is small and thus free to move between lattice sites. In the absence of an external electric field it will, therefore, be found with equal probability in any one of the four equidistant sites surrounding the stationary negative ion (see diagram; the lattice spacing is a).



If a small electrical field E is applied along the x direction, calculate the electric polarization, i.e., the mean electric dipole moment per unit volume along the x direction.

- 6.9** A wire of radius r_0 is coincident with the axis of a metal cylinder of radius R and length L . The wire is maintained at a positive potential V (statvolts) with respect to the cylinder. The whole system is at some high absolute temperature T . As a result, electrons emitted from the hot metals form a dilute gas filling the cylindrical container and in equilibrium with it. The density of these electrons is so low that their mutual electrostatic interaction can be neglected.

(a) Use Gauss's theorem to obtain an expression for the electrostatic field which exists at points at a radial distance r from the wire ($r_0 < r < R$). The cylinder of length L may be assumed very long so that end effects are negligible.

(b) In thermal equilibrium, the electrons form a gas of variable density which fills the entire space between the wire and cylinder. Using the result of part (a), find the dependence of the electric charge density on the radial distance r .

- 6.10** A dilute solution of macromolecules (large molecules of biological interest) at temperature T is placed in an ultracentrifuge rotating with angular velocity ω . The centripetal acceleration $\omega^2 r$ acting on a particle of mass m may then be replaced by an equivalent centrifugal force $m\omega^2 r$ in the rotating frame of reference.

(a) Find how the relative density $\rho(r)$ of molecules varies with their radial distance r from the axis of rotation.

(b) Show quantitatively how the molecular weight of the macromolecules can be determined if the density ratio ρ_1/ρ_2 at the radii r_1 and r_2 is measured by optical means.

- 6.11** Two atoms of mass m interact with each other by a force derivable from a mutual potential energy of the form

$$U = U_0 \left[\left(\frac{a}{x} \right)^{12} - 2 \left(\frac{a}{x} \right)^6 \right]$$

where x is the separation between the two particles. The particles are in contact with a heat reservoir at a temperature T low enough so that $kT \ll U_0$, but high enough so that classical statistical mechanics is applicable. Calculate the mean separation $\bar{x}(T)$ of the particles and use it to compute the quantity

$$\alpha \equiv \frac{1}{\bar{x}} \frac{\partial \bar{x}}{\partial T}$$

(This illustrates the fundamental procedure for calculating the coefficient of linear expansion of a solid.) Your calculation should make approximations based on the fact that the temperature is fairly low; thus retain only the lowest order terms which yield a value of $\alpha \neq 0$. (Hint: Expand the potential function about its minimum in a power series in x . To evaluate some of the integrals, use approximations similar to those used in evaluating the integral (A·6·12).)

- 6.12** Consider a rectangular box with four walls and a bottom (but no top). The total area of the walls and bottom is A . Find the dimensions of the box which give a maximum volume using

- (a) the methods of straightforward calculus;
 (b) Lagrange multipliers (see Appendix A·10).

- *6.13** Suppose that the expression

$$S = -k \sum_r P_r \ln P_r$$

is accepted as the general definition of the entropy of a system. The following problems illustrate that the entropy so defined has indeed some very interesting properties showing that S is a measure of disorder or randomness in a system.

Imagine that a system A_1 has probability $P_r^{(1)}$ of being found in a state r and a system A_2 has probability $P_s^{(2)}$ of being found in a state s . Then one has

$$S_1 = -k \sum_r P_r^{(1)} \ln P_r^{(1)} \quad \text{and} \quad S_2 = -k \sum_s P_s^{(2)} \ln P_s^{(2)}$$

Each state of the composite system A consisting of A_1 and A_2 can then be labeled by the pair of numbers r, s ; let the probability of A being found in this state be

denoted by P_{rs} . Then its entropy is defined by

$$S = -k \sum_r \sum_s P_{rs} \ln P_{rs}$$

If A_1 and A_2 are weakly interacting so that they are statistically independent, then $P_{rs} = P_r^{(1)}P_s^{(2)}$. Show that under these circumstances the entropy is simply additive, i.e., $S = S_1 + S_2$.

- *6.14 In the preceding problem, assume that A_1 and A_2 are *not* weakly interacting so that $P_{rs} \neq P_r^{(1)}P_s^{(2)}$. One has, of course, the general relations

$$P_r^{(1)} = \sum_s P_{rs} \quad \text{and} \quad P_s^{(2)} = \sum_r P_{rs}$$

Furthermore, all the probabilities are properly normalized so that

$$\sum_r P_r^{(1)} = 1, \quad \sum_s P_s^{(2)} = 1, \quad \sum_r \sum_s P_{rs} = 1$$

(a) Show that

$$S - (S_1 + S_2) = k \sum_{r,s} P_{rs} \ln \left(\frac{P_r^{(1)}P_s^{(2)}}{P_{rs}} \right)$$

(b) By using the inequality of Appendix A·8, $-\ln x \geq -x + 1$, show that

$$S \leq S_1 + S_2$$

where the equals sign holds only if $P_{rs} = P_r^{(1)}P_s^{(2)}$ holds for all r and s . This means that the existence of correlations between the systems, due to the interaction between them, leads to a situation less random than that where the systems are completely independent of each other.

- *6.15 Consider a system distributed over its accessible states r in accordance with an arbitrary probability distribution P_r , and let its entropy be defined by the relation $S = -k \sum_r P_r \ln P_r$. The distribution is properly normalized so that $\sum_r P_r = 1$. Compare this distribution with the canonical distribution

$$P_r^{(0)} = \frac{e^{-\beta E_r}}{Z}, \quad Z \equiv \sum_r e^{-\beta E_r}$$

corresponding to the same mean energy \bar{E} , i.e.,

$$\Sigma P_r^{(0)} E_r = \Sigma P_r E_r = \bar{E}$$

The entropy for this canonical distribution is given by $S_0 = -k \sum_r P_r^{(0)} \ln P_r^{(0)}$.

(a) Show that

$$\begin{aligned} S - S_0 &= k \sum_r [-P_r \ln P_r + P_r \ln P_r^{(0)} - P_r \ln P_r^{(0)} + P_r^{(0)} \ln P_r^{(0)}] \\ &= k \sum_r P_r \ln \frac{P_r^{(0)}}{P_r} \end{aligned}$$

(b) Using again the inequality $\ln x \leq x - 1$ of Appendix A·8, go on to show that $S_0 \geq S$, the equals sign holding only if $P_r = P_r^{(0)}$ for all states r .

This shows (in agreement with the discussion of Sec. 6·10) that, for a specified value of the mean energy, the entropy S is a maximum for the canonical distribution.



Simple applications of statistical mechanics

THE DISCUSSION of the preceding chapter dealt with some detailed microscopic aspects of the general theory of Chapter 3. As a result of this discussion, we have acquired some very powerful tools for calculating the macroscopic properties of any system in equilibrium from a knowledge of its microscopic constituents. The range of applicability of these conceptual tools is very wide indeed. In the present chapter, we shall illustrate their usefulness by discussing some rather simple, but very important, physical situations.

GENERAL METHOD OF APPROACH

7 · 1 Partition functions and their properties

The procedure for calculating macroscopic properties by statistical mechanics is, in principle, exceedingly simple. If the system under consideration is at a specified temperature T , i.e., if it is in thermal contact with some heat reservoir at this temperature, then one need only calculate the partition function Z of (6 · 5 3). Other physical quantities such as \bar{E} , \bar{p} , S , or even dispersions such as $(\Delta E)^2$, can then be immediately obtained from the relations of Sec. 6 · 5 by simply taking suitable derivatives of $\ln Z$. Nor is the situation significantly different if the system is not in contact with a heat reservoir. Even if the system is isolated and has fixed energy, the mean values of the macroscopic parameters of the system are still related to its temperature T as though it were in thermal contact with a heat reservoir of this temperature. Thus the calculation is again reduced to the evaluation of the partition function Z .

Thus one arrives at the near-universal prescription for calculating macroscopic properties by statistical mechanics: evaluate the partition function*

$$Z = \sum_r e^{-\beta E_r} \quad (7 \cdot 1 \cdot 1)$$

* An alternative prescription would, of course, be to evaluate $\Omega(E)$ and then to use relations such as (3 · 12 · 1) to find other quantities. But, for reasons already discussed, a direct calculation of $\Omega(E)$ is in general more difficult than a calculation of Z .

This is an unrestricted sum over all states of the system. If one knows the particles which constitute the system and the interactions between them, it is possible to find the quantum states of this system and to evaluate the sum (7 · 1 · 1). The statistical mechanical problem is then solved. In principle there is no difficulty in formulating the problem, no matter how complex the system may be. The difficulties are reduced to the mathematical ones of carrying out these prescriptions. Thus it is an easy task to find the quantum states and the partition function for an ideal gas of noninteracting atoms; but it is a formidable task to do the same for a liquid where all the molecules interact strongly with each other.

If the system can be treated in the classical approximation, then its energy $E(q_1, \dots, q_f, p_1, \dots, p_f)$ depends on some f generalized coordinates and f momenta. If phase space is subdivided into cells of volume h_0^f , the partition function in Eq. (7 · 1 · 1) can be evaluated by first summing over the number $(dq_1 \cdots dq_f dp_1 \cdots dp_f)/h_0^f$ of cells of phase space which lie in the element of volume $(dq_1 \cdots dq_f dp_1 \cdots dp_f)$ at the point $\{q_1, \dots, q_f, p_1, \dots, p_f\}$ and which have nearly the same energy $E(q_1, \dots, q_f, p_1, \dots, p_f)$; and then summing (or integrating) over all such elements of volume. Thus one obtains in the classical approximation

$$Z = \int \cdots \int e^{-\beta E(q_1, \dots, p_f)} \frac{dq_1 \cdots dp_f}{h_0^f} \quad (7 \cdot 1 \cdot 2)$$

It is worth keeping in mind the following remarks concerning the partition function Z . The first remark pertains to the energy scale used in evaluating Z . The energy of a system is only defined to within an arbitrary additive constant. If one changes by a constant amount ϵ_0 the standard state with respect to which the energy is measured, the energy of each state r becomes $E_r^* = E_r + \epsilon_0$. Correspondingly, the partition function becomes

$$Z^* = \sum_r e^{-\beta(E_r + \epsilon_0)} = e^{-\beta\epsilon_0} \sum_r e^{-\beta E_r} = e^{-\beta\epsilon_0} Z \quad (7 \cdot 1 \cdot 3)$$

or

$$\ln Z^* = \ln Z - \beta\epsilon_0$$

Thus the partition function is also changed. By (6 · 5 · 4) the new mean energy is then given by

$$\bar{E}^* = -\frac{\partial \ln Z^*}{\partial \beta} = -\frac{\partial \ln Z}{\partial \beta} + \epsilon_0 = \bar{E} + \epsilon_0$$

i.e., it is properly shifted by the amount ϵ_0 . On the other hand, the entropy is properly unchanged, since by (6 · 6 · 5)

$$S^* = k(\ln Z^* + \beta\bar{E}^*) = k(\ln Z + \beta\bar{E}) = S$$

Similarly, all expressions for generalized forces (i.e., all equations of state) are unchanged, since they involve only derivatives of $\ln Z$ with respect to an external parameter.

The second remark concerns the decomposition of the partition function for a system A when the latter consists of two parts A' and A'' which interact

only weakly with each other. If the states of A' and A'' are labeled respectively by r and s , then a state of A can be specified by the pair of numbers r,s and its corresponding energy E_{rs} is simply additive, i.e.,

$$E_{rs} = E_r' + E_s'' \quad (7 \cdot 1 \cdot 4)$$

Here A' and A'' may refer to two different distinguishable groups of particles which interact weakly with each other (e.g., He and Ne molecules in an ideal-gas mixture of these two gases). Alternatively, they may refer to two different sets of degrees of freedom of the *same* group of particles (e.g., in a diatomic gas, they may refer to (1) the degrees of freedom describing the translational motion of the centers of mass of the molecules and (2) the degrees of freedom describing the rotation of these molecules about their respective centers of mass).

The important point is only the additivity of the energies in (7 · 1 · 4); for then the partition function Z for the total system A is a sum over all states labeled by rs , i.e.,

$$Z = \sum_{r,s} e^{-\beta(E_r' + E_s'')} = \sum_{r,s} e^{-\beta E_r'} e^{-\beta E_s''} = \left(\sum_r e^{-\beta E_r'} \right) \left(\sum_s e^{-\beta E_s''} \right)$$

Thus

$$Z = Z'Z'' \quad (7 \cdot 1 \cdot 5)$$

and

$$\ln Z = \ln Z' + \ln Z'' \quad (7 \cdot 1 \cdot 6)$$

where Z' and Z'' are the partition functions of A' and A'' , respectively. Thus we have shown that if a system consists of distinct noninteracting parts, the partition function factors into a simple product.* This is a useful result and one which clearly is equally valid when one is dealing with more than two weakly interacting parts.

IDEAL MONATOMIC GAS

7 · 2 Calculation of thermodynamic quantities

Consider a gas consisting of N identical monatomic molecules of mass m enclosed in a container of volume V . Denote the position vector of the i th molecule by \mathbf{r}_i , its momentum by \mathbf{p}_i . Then the total energy of the gas is given by

$$E = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + U(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \quad (7 \cdot 2 \cdot 1)$$

Here the first term on the right represents the total kinetic energy of all the molecules. The term U represents the potential energy of interaction between the molecules. If the gas is sufficiently dilute that the interaction between molecules is negligible, $U \rightarrow 0$ and we obtain the simple case of an ideal gas.

* We already established this result in (6 · 6 · 13) where we showed that it implies the additivity of all extensive thermodynamic functions.

In writing (7·2·1), we assume the constraining condition that all the position vectors \mathbf{r}_i lie inside the volume of the container.

Let us treat the problem classically; the validity of this approximation will be examined in Sec. 7·4. Then we can immediately use (7·1·2) to write the *classical* partition function (denote it by Z')

$$Z' = \int \exp \left\{ -\beta \left[\frac{1}{2m} (\mathbf{p}_1^2 + \cdots + \mathbf{p}_N^2) + U(\mathbf{r}_1, \dots, \mathbf{r}_N) \right] \right\} \frac{d^3\mathbf{r}_1 \cdots d^3\mathbf{r}_N d^3\mathbf{p}_1 \cdots d^3\mathbf{p}_N}{h_0^{3N}}$$

or

$$Z' = \frac{1}{h_0^{3N}} \int e^{-(\beta/2m)\mathbf{p}_1^2} d^3\mathbf{p}_1 \cdots \int e^{-(\beta/2m)\mathbf{p}_N^2} d^3\mathbf{p}_N \int e^{-\beta U(\mathbf{r}_1, \dots, \mathbf{r}_N)} d^3\mathbf{r}_1 \cdots d^3\mathbf{r}_N \quad (7·2·2)$$

where the second expression follows from the first by using the multiplicative property of the exponential function. Since the kinetic energy is a sum of terms, one for each molecule, the corresponding part of the partition function breaks up into a product of N integrals, each identical except for the irrelevant variable of integration, and equal to

$$\int_{-\infty}^{\infty} e^{-(\beta/2m)p^2} d^3\mathbf{p}$$

Since U is *not* in the form of a simple sum of terms for individual molecules, the integral over the coordinates $\mathbf{r}_1, \dots, \mathbf{r}_N$ is very difficult to carry out. This is why the treatment of nonideal gases is complicated. But if the gas is sufficiently dilute to be ideal, then $U = 0$ and the integral becomes trivial; i.e.,

$$\int d^3\mathbf{r}_1 \cdots d^3\mathbf{r}_N = \int d^3\mathbf{r}_1 \int d^3\mathbf{r}_2 \cdots \int d^3\mathbf{r}_N = V^N$$

since each integration extends over the volume of the container. Then Z' factors into a simple product

$$Z' = \xi^N \quad (7·2·3)$$

or $\ln Z' = N \ln \xi \quad (7·2·4)$

where $\xi \equiv \frac{V}{h_0^3} \int_{-\infty}^{\infty} e^{-(\beta/2m)p^2} d^3\mathbf{p} \quad (7·2·5)$

is the partition function for a single molecule.

■ ■ ■

Remark It would be possible to formulate this problem in a slightly different way by not imposing the condition that the position coordinate r of each molecule lie within the container. In that case, to write down an expression for the total energy valid everywhere, one would have to add to (7·2·1) a term

$$U' = \sum_i u(r_i)$$

where $u(r)$ represents the potential energy of a molecule due to the container, i.e.,

$$u(r) = \begin{cases} 0 & \text{if } r \text{ lies inside the container} \\ \infty & \text{if } r \text{ lies outside the container} \end{cases}$$

In this case the partition function (7.2.2) would contain a factor $e^{-\beta U'}$ which would equal unity whenever all molecules are within the container and would equal zero whenever any one molecule is outside the container. Thus the integration over all coordinates without restriction would again immediately reduce to the form (7.2.2) of integration over the volume of the container only.

The integral in (7.2.5) is readily evaluated.

$$\int_{-\infty}^{\infty} e^{-(\beta/2m)p^2} d^3p = \iiint_{-\infty}^{\infty} e^{-(\beta/2m)(p_x^2+p_y^2+p_z^2)} dp_x dp_y dp_z \\ = \int_{-\infty}^{\infty} e^{-(\beta/2m)p_x^2} dp_x \int_{-\infty}^{\infty} e^{-(\beta/2m)p_y^2} dp_y \int_{-\infty}^{\infty} e^{-(\beta/2m)p_z^2} dp_z \\ = \left(\sqrt{\frac{\pi 2m}{\beta}} \right)^3 \quad \text{by (A.4.2)}$$

Hence $\xi = V \left(\frac{2\pi m}{h_0^2 \beta} \right)^{1/2} \quad (7.2.6)$

and $\ln Z' = N \left[\ln V - \frac{3}{2} \ln \beta + \frac{3}{2} \ln \left(\frac{2\pi m}{h_0^2} \right) \right] \quad (7.2.7)$

From this partition function one can immediately calculate a host of other physical quantities. By (6.5.12) the mean gas pressure \bar{p} is given by

$$\bar{p} = \frac{1}{\beta} \frac{\partial \ln Z'}{\partial V} = \frac{1}{\beta} \frac{N}{V} \quad (7.2.8)$$

Thus

► $\bar{p}V = NkT \quad (7.2.8)$

and one regains the equation of state already derived in (3.12.8) under more general conditions (gas not necessarily monatomic).

By (6.5.4), the total mean energy of the gas is

$$\bar{E} = - \frac{\partial}{\partial \beta} \ln Z' = \frac{3}{2} \frac{N}{\beta} = N\bar{\epsilon} \quad (7.2.9)$$

where

► $\bar{\epsilon} = \frac{3}{2}kT \quad (7.2.10)$

is the mean energy per molecule. The heat capacity at constant volume of the gas is then given by

$$C_V = \left(\frac{\partial \bar{E}}{\partial T} \right)_V = \frac{3}{2} Nk = \frac{3}{2} \nu N_A k \quad (7.2.11)$$

where ν is the number of moles and N_a is Avogadro's number. Hence the *molar* specific heat at constant volume of a monatomic gas is

$$c_V = \frac{3}{2}R \quad (7 \cdot 2 \cdot 12)$$

where $R = N_a k$ is the gas constant. These results agree with those already obtained in (5 · 2 · 10) and (5 · 2 · 12).

Remark on fluctuations The fluctuation in the total energy of the gas in contact with a heat reservoir at temperature T can also be readily calculated. By (6 · 5 · 8) the dispersion in energy is given by

$$\overline{(\Delta E)^2} = - \frac{\partial \bar{E}}{\partial \beta}$$

Here the volume V is, of course, kept constant in taking the derivative. Putting $\beta = (kT)^{-1}$, this becomes

$$\overline{(\Delta E)^2} = - \left(\frac{\partial \bar{E}}{\partial T} \right)_V \frac{\partial T}{\partial \beta} = kT^2 \left(\frac{\partial \bar{E}}{\partial T} \right)_V$$

or

$$\overline{(\Delta E)^2} = kT^2 C_V \quad (7 \cdot 2 \cdot 13)$$

Thus the fluctuation in energy of *any* system is quite generally related to its heat capacity at constant volume. In the particular case of a monatomic ideal gas consisting of N molecules one obtains, by (7 · 2 · 11),

$$\overline{(\Delta E)^2} = \frac{3}{2}Nk^2T^2 \quad (7 \cdot 2 \cdot 14)$$

The root-mean-square fluctuation in energy $\Delta^*E = \sqrt{[(\Delta E)^2]}$ can be compared to the mean energy \bar{E} of the gas. Thus

$$\frac{\Delta^*E}{\bar{E}} = \frac{\sqrt{\frac{3}{2}Nk^2T^2}}{\frac{3}{2}NkT} = \sqrt{\frac{2}{3N}} \quad (7 \cdot 2 \cdot 15)$$

This is very small when N is of the order of Avogadro's number.

The entropy of the gas can be calculated by (6 · 6 · 5). Using (7 · 2 · 7) and (7 · 2 · 9) one obtains

$$S = k(\ln Z' + \beta\bar{E}) = Nk \left[\ln V - \frac{3}{2}\ln\beta + \frac{3}{2}\ln\left(\frac{2\pi m}{h_0^2}\right) + \frac{3}{2} \right]$$

or $S = Nk[\ln V + \frac{3}{2}\ln T + \sigma] \quad (7 \cdot 2 \cdot 16)$

where $\sigma = \frac{3}{2}\ln\left(\frac{2\pi m k}{h_0^2}\right) + \frac{3}{2}$

is a constant independent of T , V , or N . This expression for the entropy is, however, *not* correct.

7·3 Gibbs paradox

The challenging statement at the end of the last section suggests that the expression (7·2·16) for the entropy merits some discussion. First, note that our calculation was carried out within the framework of classical mechanics which certainly is not valid at very low temperatures where the system is in the (relatively few) states of very low energy close to the quantum-mechanical ground state and where a quantum description is certainly needed. Hence the circumstance that Eq. (7·2·16) yields $S \rightarrow -\infty$ as $T \rightarrow 0$ (in apparent contradiction to the third law of thermodynamics) is no cause for alarm. In accordance with its classical derivation, (7·2·16) is not expected to be valid at such low temperatures.

Nevertheless, the expression (7·2·16) for S is clearly wrong since it implies that the entropy does not behave properly as an extensive quantity. Quite generally, one must require that all thermodynamic relations remain valid if the size of the whole system under consideration is simply increased by a scale factor α , i.e., if all its extensive parameters are multiplied by the same factor α . In our case, if the independent extensive parameters V and N are multiplied by α , the mean energy \bar{E} in (7·2·9) is indeed properly increased by this same factor, but the entropy S in (7·2·16) is *not* increased by α because of the term $N \ln V$.

Indeed, (7·2·16) asserts that the entropy S of a fixed volume V of gas is simply proportional to the number N of molecules. But this dependence on N is *not* correct, as can readily be seen in the following way. Imagine that a partition is introduced which divides the container into two parts. This is a reversible process which does not affect the distribution of systems over accessible states. Thus the total entropy ought to be the same with, or without, the partition in place; i.e.,

$$S = S' + S'' \quad (7\cdot3\cdot1)$$

where S' and S'' are the entropies of the two parts. But the expression (7·2·16) does *not* yield the simple additivity required by (7·3·1). This is easily verified. Suppose, for example, that the partition divides the gas into two *equal* parts, each containing N' molecules of gas in a volume V' . Then the entropy of each part is given by (7·2·16) as

$$S' = S'' = N'k[\ln V' + \frac{3}{2} \ln T + \sigma]$$

while the entropy of the whole gas without partition is by (7·2·16)



Fig. 7·3·1 A container of gas divided into two equal parts by a partition.

$$S = 2N'k[\ln(2V') + \frac{3}{2}\ln T + \sigma]$$

$$\text{Hence } S - 2S' = 2N'k \ln(2V') - N'k \ln V' = 2N'k \ln 2 \quad (7 \cdot 3 \cdot 2)$$

and is *not* equal to zero as required by (7 · 3 · 1).

This paradox was first discussed by Gibbs and is commonly referred to as the "Gibbs paradox." Something is obviously wrong in our discussion; the question is what. Did we not prove quite generally in (6 · 6 · 15) that the entropies of two weakly interacting systems are additive? How then can we fail to satisfy the condition (7 · 3 · 1)? The answer is quite simple. Our general argument in Sec. 6 · 6 was based on the premise that the external parameters of each subsystem remain the *same*. If we brought the two gases in our example together and left them separated by a partition, then the volume V' of each subsystem would remain the same and their entropies would satisfy the additivity (7 · 3 · 1). But we did more than that—we also removed the partition. In that case Eq. (6 · 6 · 11) is no longer valid because the energies E_r' and E_s'' are both calculated with the volume V' as the external parameter, while for the combined system (with partition removed) the possible states of energy E_{rs} are to be calculated with total volume $2V'$ as the external parameter.

The act of removing the partition has thus very definite physical consequences. Whereas before removal of the partition a molecule of each subsystem could only be found within a volume V' , after the partition is removed it can be located anywhere within the volume $V = 2V'$. If the two subsystems consisted of different gases, the act of removing the partition would lead to diffusion of the molecules throughout the whole volume $2V'$ and consequent random mixing of the different molecules. This is clearly an irreversible process; simply replacing the partition would not unmix the gases. In this case the increase of entropy in (7 · 3 · 2) would make sense as being simply a measure of the irreversible increase of disorder resulting from the mixing of unlike gases.

But if the gases in the subsystems are identical, such an increase of entropy does *not* make physical sense. The root of the difficulty embodied in the Gibbs paradox is that we treated the gas molecules as individually distinguishable, as though interchanging the positions of two like molecules would lead to a physically distinct state of the gas. This is not so. Indeed, if we treated the gas by quantum mechanics (as we shall do in Chapter 9), the molecules would, as a matter of principle, have to be regarded as completely indistinguishable. A calculation of the partition function would then automatically yield the correct result, and the Gibbs paradox would never arise. Our mistake has been to take the classical point of view too seriously. Even though one may be in a temperature and density range where the motion of molecules can be treated to a very good approximation by classical mechanics, one cannot go so far as to disregard the essential indistinguishability of the molecules; one cannot observe and label individual atomic particles as though they were macroscopic billiard balls. If one does want to use the classical approximation, then the indistinguishability of the molecules must be taken into account explicitly in calculating the partition function (7 · 2 · 2). This can be done by noting

that the $N!$ possible permutations of the molecules among themselves do not lead to physically distinct situations, so that the number of distinct states over which one sums in (7·2·2) is too large by a factor of $N!$. The correct partition function Z , which does take into account the essential indistinguishability of the molecules and does not lead to the Gibbs paradox difficulties, is then given by dividing (7·2·3) by this factor, i.e.,

$$Z = \frac{Z'}{N!} = \frac{\xi^N}{N!} \quad (7\cdot3\cdot3)$$

Note that in a *strictly* classical description it would be permissible to consider every particle as distinguishable. If one agrees to consider identical molecules as essentially indistinguishable so as to avoid the Gibbs paradox, then the following question arises: Just how different must molecules be before they should be considered distinguishable (i.e., before their mutual mixing leads to a finite, instead of no, increase of entropy)? In a classical view of nature two molecules could, of course, differ by infinitesimal amounts (e.g., the nuclei of two atoms could have infinitesimally different masses). In a quantum description this troublesome question does not arise because of the quantized discreteness of nature (e.g., the nuclei of two isotopes differ by at least one nucleon mass). Hence the distinction between identical and non-identical molecules is completely unambiguous in a quantum-mechanical description. The Gibbs paradox thus foreshadowed already in the last century conceptual difficulties that were resolved satisfactorily only by the advent of quantum mechanics.

By (7·3·3) one then gets

$$\begin{aligned} \ln Z &= N \ln \xi - \ln N! \\ \text{or} \quad \ln Z &= N \ln \xi - N \ln N + N \end{aligned} \quad (7\cdot3\cdot4)$$

where we have used Stirling's formula. Equation (7·3·4) differs from the corresponding expression (7·2·4) only by the additive term $(-N \ln N + N)$. Since the pressure \bar{p} and energy \bar{E} depend only on *derivatives* of $\ln Z$ with respect to V or β , the previous results (7·2·8) and (7·2·9) for these quantities are unaffected. But the expression for S , which does involve $\ln Z$ itself rather than only its derivatives, is changed by this additive term. Thus (7·3·4) yields, instead of (7·2·16), the result

$$S = kN[\ln V + \frac{3}{2} \ln T + \sigma] + k(-N \ln N + N)$$

or

$$\blacktriangleright \quad S = kN \left[\ln \frac{V}{N} + \frac{3}{2} \ln T + \sigma_0 \right] \quad (7\cdot3\cdot5)$$

$$\text{where } \sigma_0 \equiv \sigma + 1 = \frac{3}{2} \ln \left(\frac{2\pi mk}{h_0^2} \right) + \frac{5}{2} \quad (7\cdot3\cdot6)$$

It is apparent that the extra term involving $\ln N$ avoids the difficulties of the Gibbs paradox. The entropy S in (7·3·5) behaves properly like an

extensive quantity; i.e., it does get multiplied by a factor α if both V and N are multiplied by α .

Since h_0 is an arbitrary constant in the present classical calculation, σ_0 is some arbitrary additive constant in the entropy. Note that (7·3·5) agrees exactly with the entropy expression derived by macroscopic reasoning in (5·4·4). It is only necessary to put $N = \nu N_a$, where ν is the number of moles of gas, and to use the relation (7·2·12), according to which $c_V = \frac{3}{2}N_a k$ for a monatomic ideal gas.

7 · 4 Validity of the classical approximation

We saw that the essential indistinguishability of identical molecules cannot be disregarded even if the motion of the molecules can be treated by classical mechanics. But to what extent is the latter procedure itself valid? That is, to what extent is it permissible to evaluate the partition function (7·2·2) in terms of coordinates r_i and momenta p_i which can be simultaneously specified?

An approximate criterion for the validity for this classical description can be obtained by appealing to the Heisenberg uncertainty principle

$$\Delta q \Delta p \gtrsim \hbar \quad (7 \cdot 4 \cdot 1)$$

This relates the uncertainties Δq and Δp introduced by quantum effects in any attempt at simultaneous specification of a position q and corresponding momentum p of a particle. Suppose that one tries to describe the motion of the gas molecules by classical mechanics. Denote the magnitude of the mean momentum of a molecule by* \bar{p} and the mean separation between molecules by \bar{R} . Then one would certainly expect a classical description to be applicable if

$$\bar{R} \bar{p} \gg \hbar \quad (7 \cdot 4 \cdot 2)$$

when (7·4·1) implies that quantum mechanical effects are not important. Equivalently (7·4·2) expresses the condition that

$$\bar{R} \gg \bar{\lambda} \quad (7 \cdot 4 \cdot 3)$$

i.e., that the mean separation between particles is much greater than their mean de Broglie wavelength

$$\bar{\lambda} = 2\pi \frac{\hbar}{\bar{p}} = \frac{\hbar}{\bar{p}} \quad (7 \cdot 4 \cdot 4)$$

When (7·4·3) is satisfied so that $\bar{R} \gg \bar{\lambda}$, the quantum description ought to be equivalent to the motion of wave packets describing individual particles which move independently in a quasi-classical manner. In the opposite limit, where $\bar{R} \ll \bar{\lambda}$, a state of the whole gas will be shown in Chapter 9 to be described by a single wave function which cannot be decomposed in any simple way; it thus

* The symbol \bar{p} should not be confused with the mean pressure of the gas. We shall denote the latter quantity by capital \bar{P} later in this section.

results in correlations between the motions of the particles even if no forces exist between them.

The mean intermolecular separation \bar{R} can be estimated by imagining each molecule at the center of a little cube of side \bar{R} , these cubes filling the available volume V . Then

$$\begin{aligned} \bar{R}^3 N &= V \\ \text{or} \quad \bar{R} &= \left(\frac{V}{N}\right)^{\frac{1}{3}} \end{aligned} \quad (7 \cdot 4 \cdot 5)$$

The mean momentum \bar{p} can be estimated from the known mean energy $\bar{\epsilon}$ of a molecule in the gas at temperature T . By (7.2.10)

$$\frac{1}{2m} \bar{p}^2 \approx \bar{\epsilon} = \frac{3}{2} kT$$

Thus

$$\bar{p} \approx \sqrt{3mkT}$$

and

$$\bar{\lambda} \approx \frac{\hbar}{\sqrt{3mkT}} \quad (7 \cdot 4 \cdot 6)$$

Hence the condition (7.4.3) becomes

$$\left(\frac{V}{N}\right)^{\frac{1}{3}} \gg \frac{\hbar}{\sqrt{3mkT}} \quad (7 \cdot 4 \cdot 7)$$

This shows that the classical approximation ought to be applicable if the concentration N/V of molecules in the gas is sufficiently small, if the temperature T is sufficiently high, and if the mass of the molecules is not too small.

Numerical estimates Consider, for example, helium (He) gas at room temperature and pressure. Then one has

mean pressure $\bar{P} = 760 \text{ mm Hg} \approx 10^6 \text{ dynes/cm}^2$

temperature $T \approx 300^\circ\text{K}$; hence $kT \approx 4 \times 10^{-14} \text{ ergs}$

molecular mass $m = \frac{4}{6 \times 10^{23}} \approx 7 \times 10^{-24} \text{ grams}$

The equation of state gives

$$\frac{N}{V} = \frac{\bar{P}}{kT} = 2.5 \times 10^{19} \text{ molecules/cm}^3$$

Thus $\bar{R} \approx 34 \times 10^{-8} \text{ cm}$ by (7.4.5)

and $\bar{\lambda} \approx 0.6 \times 10^{-8} \text{ cm}$ by (7.4.6)

Here the criterion (7.4.3) is quite well satisfied, and the classical evaluation of the partition function ought to be a very good approximation if the indistinguishability of the particles is taken into account. Most gases have larger molecular weights and thus smaller de Broglie wavelengths; the criterion (7.4.3) is then even better satisfied.

On the other hand, consider the conduction electrons in a typical metal. In a first approximation, interactions between these electrons can be neglected

so that they can be treated as an ideal gas. But the numerical values of the significant parameters are then quite different. First, the mass of the electron is very small, about 10^{-27} g or 7000 times less than that of the He atom. This makes the de Broglie wavelength of the electron much longer,

$$\bar{\lambda} \approx (0.6 \times 10^{-8}) \sqrt{7000} \approx 50 \times 10^{-8} \text{ cm}$$

In addition, there is about one conduction electron per atom in the metal. Since there is roughly one atom in a cube 2×10^{-8} cm on a side,

$$\bar{R} \approx 2 \times 10^{-8} \text{ cm}$$

This is much smaller than for the He gas case; i.e., the electrons in a metal form a very dense gas. Hence the criterion (7·4·3) is certainly not satisfied. Thus there exists no justification for discussing electrons in a metal by classical statistical mechanics; indeed, a completely quantum-mechanical treatment is essential.

THE EQUIPARTITION THEOREM

7 · 5 Proof of the theorem

In *classical* statistical mechanics there exists a very useful general result which we shall now establish. As usual, the energy of a system is a function of some *f* generalized coordinates q_k and corresponding *f* generalized momenta p_k ; i.e.,

$$E = E(q_1, \dots, q_f, p_1, \dots, p_f) \quad (7 \cdot 5 \cdot 1)$$

The following is a situation of frequent occurrence:

- a. The total energy splits additively into the form

$$E = \epsilon_i(p_i) + E'(q_1, \dots, p_f) \quad (7 \cdot 5 \cdot 2)$$

where ϵ_i involves only the one variable p_i and the remaining part E' does *not* depend on p_i .

- b. The function ϵ_i is quadratic in p_i ; i.e., it is of the form

$$\epsilon_i(p_i) = bp_i^2 \quad (7 \cdot 5 \cdot 3)$$

where b is some constant.

The most common situation is one where p_i is a momentum. The reason is that the kinetic energy is usually a quadratic function of each momentum component, while the potential energy does not involve the momenta.

If in assumptions (a) and (b) the variable were not a momentum p_i but a coordinate q_i satisfying the same two conditions, the theorem we want to establish would be exactly the same.

We ask the question: What is the mean value of ϵ_i in thermal equilibrium if conditions (a) and (b) are satisfied?

If the system is in equilibrium at the absolute temperature $T = (k\beta)^{-1}$, it is distributed in accordance with the canonical distribution; the mean value

$\bar{\epsilon}_i$ is then, by definition, expressible in terms of integrals over all phase space

$$\bar{\epsilon}_i = \frac{\int_{-\infty}^{\infty} e^{-\beta E(q_1, \dots, p_f)} \epsilon_i dq_1 \dots dp_f}{\int_{-\infty}^{\infty} e^{-\beta E(q_1, \dots, p_f)} dq_1 \dots dp_f} \quad (7 \cdot 5 \cdot 4)$$

By condition (a) this becomes

$$\begin{aligned} \bar{\epsilon}_i &= \frac{\int e^{-\beta(\epsilon_i + E')} \epsilon_i dq_1 \dots dp_f}{\int e^{-\beta(\epsilon_i + E')} dq_1 \dots dp_f} \\ &= \frac{\int e^{-\beta \epsilon_i} \epsilon_i dp_i \int e^{-\beta E'} dq_1 \dots dp_f}{\int e^{-\beta \epsilon_i} dp_i \int e^{-\beta E'} dq_1 \dots dp_f} \end{aligned}$$

where we have used the multiplicative property of the exponential function and where the last integrals in both numerator and denominator extend over all terms q and p except p_i . These integrals are equal and thus cancel; hence only the one-dimensional integrals survive:

$$\bar{\epsilon}_i = \frac{\int e^{-\beta \epsilon_i} \epsilon_i dp_i}{\int e^{-\beta \epsilon_i} dp_i} \quad (7 \cdot 5 \cdot 5)$$

This can be simplified further by reducing the integral in the numerator to that in the denominator. Thus

$$\begin{aligned} \bar{\epsilon}_i &= \frac{-\frac{\partial}{\partial \beta} (\int e^{-\beta \epsilon_i} dp_i)}{\int e^{-\beta \epsilon_i} dp_i} \\ \text{or} \quad \bar{\epsilon}_i &= -\frac{\partial}{\partial \beta} \ln \left(\int_{-\infty}^{\infty} e^{-\beta \epsilon_i} dp_i \right) \end{aligned} \quad (7 \cdot 5 \cdot 6)$$

Up to now we have made use only of the assumption (7 · 5 · 2). Let us now use the second assumption (7 · 5 · 3). Then the integral in (7 · 5 · 6) becomes

$$\int_{-\infty}^{\infty} e^{-\beta \epsilon_i} dp_i = \int_{-\infty}^{\infty} e^{-\beta b p_i^2} dp_i = \beta^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{-by^2} dy$$

where we have introduced the variable $y \equiv \beta^{\frac{1}{2}} p_i$. Thus

$$\ln \int_{-\infty}^{\infty} e^{-\beta \epsilon_i} dp_i = -\frac{1}{2} \ln \beta + \ln \int_{-\infty}^{\infty} e^{-by^2} dy$$

But here the integral on the right does *not* involve β at all. Hence (7 · 5 · 6) becomes simply

$$\bar{\epsilon}_i = -\frac{\partial}{\partial \beta} \left(-\frac{1}{2} \ln \beta \right) = \frac{1}{2\beta}$$

or

$$\bar{\epsilon}_i = \frac{1}{2} kT \quad (7 \cdot 5 \cdot 7)$$

Note the great generality of this result and that we obtained it *without* needing to evaluate a single integral.

Equation (7 · 5 · 7) is the so-called “equipartition theorem” of classical statistical mechanics. In words it states that the mean value of each independent quadratic term in the energy is equal to $\frac{1}{2}kT$.

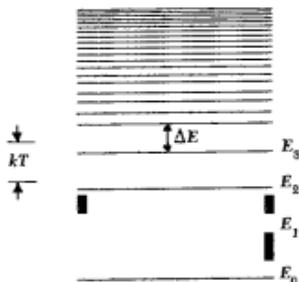


Fig. 7·5·1 Schematic diagram of the energy levels of a system.

It should be emphasized that the equipartition theorem is valid only in *classical* statistical mechanics. In the correct quantum-mechanical description a system has a set of possible energy levels, as indicated in Fig. 7·5·1, where E_0 is the ground-state energy and where at higher energies the levels are, in general, increasingly closely spaced. When the absolute temperature is sufficiently high (and thus the mean energy of the system is sufficiently high) the spacing ΔE between levels around the mean energy \bar{E} is small compared to the thermal energy kT ; i.e., $\Delta E \ll kT$. In this case the fact that there are discrete energy levels is not particularly important, and the classical description (and equipartition theorem where applicable) can be expected to be a good approximation. On the other hand, when the temperature is sufficiently low so that $kT \lesssim \Delta E$, the classical description must certainly break down.

7 · 6 Simple applications

Mean kinetic energy of a molecule in a gas Consider a molecule in a gas (not necessarily an ideal gas) at temperature T . If this molecule has mass m and a center-of-mass momentum $\mathbf{p} = m\mathbf{v}$, its kinetic energy of translation is

$$K = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) \quad (7 \cdot 6 \cdot 1)$$

The kinetic energy of the other molecules does not involve the momentum \mathbf{p} of this particular molecule. The potential energy of interaction between molecules depends only on their position coordinates and thus certainly does not involve \mathbf{p} . Finally, if the molecule is polyatomic, the internal energy of vibration or rotation of its atoms relative to its center of mass also does not involve \mathbf{p} . Hence the essential conditions of the equipartition theorem are satisfied. Since (7·6·1) contains three quadratic terms, the equipartition theorem allows one to conclude immediately that

$$\bar{K} = \frac{3}{2}kT \quad (7 \cdot 6 \cdot 2)$$

if the motion of the center of mass can be treated classically.

For an ideal monatomic gas the *entire* energy is kinetic, so that the mean energy per mole of gas is simply

$$\bar{E} = N_a \left(\frac{3}{2} kT \right) = \frac{3}{2} RT$$

The molar specific heat at constant volume is then

$$c_V = \left(\frac{\partial \bar{E}}{\partial T} \right)_V = \frac{3}{2} R \quad (7 \cdot 6 \cdot 3)$$

Brownian motion Consider a macroscopic particle of mass m immersed in a liquid at temperature T . Let the z axis point in the direction of the gravitational field (if one is present) and focus attention on v_x , the x component of the center-of-mass velocity of the particle. The mean value of v_x must vanish by symmetry; i.e.,

$$\bar{v}_x = 0$$

But it is, of course, not true that v_x itself is always found to vanish if one observes a collection of such particles; velocity fluctuations *do* occur. Indeed, the equipartition theorem can be applied to the center-of-mass energy terms just as in the preceding example; thus one can conclude that

$$\overline{\frac{1}{2}mv_x^2} = \frac{1}{2}kT \quad \text{or} \quad \overline{v_x^2} = \frac{kT}{m}$$

The dispersion $\overline{v_x^2}$ in this velocity component is thus negligibly small when m is large. For example, when the particle is the size of a golf ball, fluctuations in its velocity are essentially unobservable and the particle appears to be at rest. But when m is small (e.g., when the particle has a diameter of about a micron), $\overline{v_x^2}$ becomes appreciable and velocity fluctuations can readily be observed under a microscope. The fact that small particles of this kind perpetually move about in a random manner was first observed by Brown, a botanist, in the last century. The phenomenon is, therefore, called "Brownian motion." It was explained theoretically by Einstein in 1905 on the basis of the intrinsic thermal fluctuations resulting from the interaction of the small particle with the heat bath, i.e., from the random collisions of the particle with the molecules of the liquid. The phenomenon was historically important in helping to gain acceptance for the atomic theory of all matter and for the validity of the statistical description thereof.

Harmonic oscillator Consider a one-dimensional harmonic oscillator which is in equilibrium with a heat reservoir at absolute temperature T . The energy of such an oscillator is given by

$$E = \frac{p^2}{2m} + \frac{1}{2} \kappa_0 x^2 \quad (7 \cdot 6 \cdot 4)$$

where the first term on the right is the kinetic energy involving the momentum p and mass m , and the second term on the right is the potential energy involv-

ing the position coordinate x and spring constant κ_0 . Each of these terms is quadratic in the respective variable. Hence the equipartition theorem leads immediately to the following conclusions, valid in the classical approximation:

$$\text{mean kinetic energy} = \frac{1}{2m} \bar{p^2} = \frac{1}{2} kT$$

$$\text{mean potential energy} = \frac{1}{2} \kappa_0 \bar{x^2} = \frac{1}{2} kT$$

Hence the mean total energy is

$$\bar{E} = \frac{1}{2} kT + \frac{1}{2} kT = kT \quad (7 \cdot 6 \cdot 5)$$

It is instructive to treat this example by quantum mechanics as an illustration of the limits of validity of the classical description. According to quantum mechanics the possible energy levels of the harmonic oscillator are given by

$$E_n = (n + \frac{1}{2})\hbar\omega \quad (7 \cdot 6 \cdot 6)$$

where the possible states of the oscillator are labeled by the quantum number n which can assume all integral values

$$n = 0, 1, 2, 3, \dots$$

Here \hbar is Planck's constant (divided by 2π) and

$$\omega = \sqrt{\frac{\kappa_0}{m}} \quad (7 \cdot 6 \cdot 7)$$

is the classical angular frequency of oscillation of the oscillator. The mean energy of the oscillator is then given by

$$\bar{E} = \frac{\sum_{n=0}^{\infty} e^{-\beta E_n} E_n}{\sum_{n=0}^{\infty} e^{-\beta E_n}} = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\partial}{\partial \beta} \ln Z \quad (7 \cdot 6 \cdot 8)$$

where $Z \equiv \sum_{n=0}^{\infty} e^{-\beta E_n} = \sum_{n=0}^{\infty} e^{-(n+\frac{1}{2})\beta\hbar\omega}$ (7 · 6 · 9)

or $Z = e^{-\frac{1}{2}\beta\hbar\omega} \sum_{n=0}^{\infty} e^{-n\beta\hbar\omega} = e^{-\frac{1}{2}\beta\hbar\omega}(1 + e^{-\beta\hbar\omega} + e^{-2\beta\hbar\omega} + \dots)$

This sum is just an infinite geometric series where each term is obtained from the preceding one as a result of multiplication by $e^{-\beta\hbar\omega}$. The sum can thus immediately be evaluated to give

$$Z = e^{-\frac{1}{2}\beta\hbar\omega} \frac{1}{1 - e^{-\beta\hbar\omega}} \quad (7 \cdot 6 \cdot 10)$$

or $\ln Z = -\frac{1}{2}\beta\hbar\omega - \ln(1 - e^{-\beta\hbar\omega}) \quad (7 \cdot 6 \cdot 11)$

Thus one obtains, by (7 · 6 · 8),

$$\bar{E} = -\frac{\partial}{\partial \beta} \ln Z = -\left(-\frac{1}{2}\hbar\omega - \frac{e^{-\beta\hbar\omega}\hbar\omega}{1 - e^{-\beta\hbar\omega}}\right)$$

or

$$\blacktriangleright \quad \bar{E} = \hbar\omega \left(\frac{1}{2} + \frac{1}{e^{\beta\hbar\omega} - 1} \right) \quad (7 \cdot 6 \cdot 12)$$

Let us now investigate some limiting cases. When

$$\beta\hbar\omega = \frac{\hbar\omega}{kT} \ll 1 \quad (7 \cdot 6 \cdot 13)$$

the temperature is so high that the thermal energy kT is large compared to the separation $\hbar\omega$ between energy levels. Then one expects the classical description to be a good approximation. Indeed, if (7·6·13) is valid, the exponential function can be expanded in Taylor's series so that (7·6·12) becomes

$$\begin{aligned} \bar{E} &= \hbar\omega \left[\frac{1}{2} + \frac{1}{(1 + \beta\hbar\omega + \dots) - 1} \right] \approx \hbar\omega \left[\frac{1}{2} + \frac{1}{\beta\hbar\omega} \right] \\ &\approx \hbar\omega \left[\frac{1}{\beta\hbar\omega} \right] \quad \text{by virtue of (7·6·13)} \end{aligned}$$

or $\bar{E} = \frac{1}{\beta} = kT \quad (7 \cdot 6 \cdot 14)$

in agreement with the classical result (7·6·5).

On the other hand, at low temperatures where

$$\beta\hbar\omega = \frac{\hbar\omega}{kT} \gg 1 \quad (7 \cdot 6 \cdot 15)$$

one has $e^{\beta\hbar\omega} \gg 1$, so that (7·6·12) becomes

$$\bar{E} = \hbar\omega \left(\frac{1}{2} + e^{-\beta\hbar\omega} \right) \quad (7 \cdot 6 \cdot 16)$$

This is quite different from the equipartition result (7·6·5) and approaches properly the ("zero point") energy $\frac{1}{2}\hbar\omega$ of the ground state as $T \rightarrow 0$.

7 · 7 Specific heats of solids

Consider any simple solid with Avogadro's number N_a of atoms per mole. Examples might be copper, gold, aluminum, or diamond. These atoms are free to vibrate about their equilibrium positions. (Such vibrations are called "lattice vibrations.") Each atom is specified by three position coordinates and three momentum coordinates. Since the vibrations are supposed to be small, the potential energy of interaction between atoms can be expanded about their equilibrium positions and is therefore quadratic in the atomic displacements from their equilibrium positions. The net result is that the total energy of lattice vibrations can be written (when expressed in terms of appropriate "normal mode coordinates") in the simple form

$$E = \sum_{i=1}^{3N_a} \left(\frac{p_i^2}{2m} + \frac{1}{2} \kappa_i q_i^2 \right) \quad (7 \cdot 7 \cdot 1)$$

Here the first term is the total kinetic energy involving the $3N_a$ (normal-mode) momenta of the atoms, while the second term is the total potential energy involving their $3N_a$ (normal-mode) coordinates. The coefficients κ_i are positive constants. Thus the total energy is the same as that of $3N_a$ independent one-dimensional harmonic oscillators. If the temperature T is high enough so that classical mechanics is applicable (and room temperature is usually sufficiently high for that), application of the equipartition theorem allows one to conclude immediately that the total mean energy per mole is

$$\bar{E} = 3N_a[(\frac{1}{2}kT) \times 2]$$

or

$$\bar{E} = 3N_akT = 3RT \quad (7 \cdot 7 \cdot 2)$$

Thus the molar specific heat at constant volume becomes

$$\blacktriangleright \quad c_V = \left(\frac{\partial \bar{E}}{\partial T} \right)_V = 3R \quad (7 \cdot 7 \cdot 3)$$

This result asserts that at sufficiently high temperatures all simple solids have the same molar specific heat equal to $3R$ (25 joules mole $^{-1}$ deg $^{-1}$). Historically, the validity of this result was first discovered empirically and is known as the law of Dulong and Petit. Table 7·7·1 lists directly measured values of the molar specific heat c_p at constant pressure for some solids at room temperature. The molar specific heat c_V at constant volume is somewhat less (by about 5 percent, as calculated in the numerical example of Sec. 5·7).

Table 7·7·1 Values* of c_p (joules mole $^{-1}$ deg $^{-1}$) for some solids at $T = 298^\circ\text{K}$

Solid	c_p	Solid	c_p
Copper	24.5	Aluminum	24.4
Silver	25.5	Tin (white)	26.4
Lead	26.4	Sulfur (rhombic)	22.4
Zinc	25.4	Carbon (diamond)	6.1

* "American Institute of Physics Handbook," 2d ed., McGraw-Hill Book Company, New York, 1963, p. 4-48.

Of course, the preceding arguments are not valid for solids at appreciably lower temperatures. Indeed, the third law leads to the general result (5·7·19), which requires that c_V must approach zero as $T \rightarrow 0$. One can obtain an approximate idea of the behavior of c_V at all temperatures by making the crude assumption (first introduced by Einstein) that all atoms in the solid vibrate with the same angular frequency ω . Then $\kappa_i = m\omega^2$ for all terms i in (7·7·1), and the mole of solid is equivalent to an assembly of $3N_a$ independent one-dimensional harmonic oscillators. These can be treated by quantum mechanics so that their total mean energy is just $3N_a$ times that of the single oscillator discussed in (7·6·12); i.e.,

$$\bar{E} = 3N_a\hbar\omega \left(\frac{1}{2} + \frac{1}{e^{\beta\hbar\omega} - 1} \right) \quad (7 \cdot 7 \cdot 4)$$

Hence the molar specific heat of the solid on the basis of this simple Einstein model is given by

$$\begin{aligned} c_V &= \left(\frac{\partial \bar{E}}{\partial T} \right)_V = \left(\frac{\partial \bar{E}}{\partial \beta} \right)_V \frac{\partial \beta}{\partial T} = - \frac{1}{kT^2} \left(\frac{\partial \bar{E}}{\partial \beta} \right)_V \\ &= - \frac{3N_a \hbar \omega}{kT^2} \left[- \frac{e^{\beta \hbar \omega} \hbar \omega}{(e^{\beta \hbar \omega} - 1)^2} \right] \\ \text{or } c_V &= 3R \left(\frac{\Theta_E}{T} \right)^2 \frac{e^{\Theta_E/T}}{(e^{\Theta_E/T} - 1)^2} \end{aligned} \quad (7 \cdot 7 \cdot 5)$$

where $R \equiv N_a k$ and where we have written

$$\beta \hbar \omega = \frac{\hbar \omega}{kT} \equiv \frac{\Theta_E}{T}$$

by introducing the characteristic "Einstein temperature"

$$\Theta_E \equiv \frac{\hbar \omega}{k} \quad (7 \cdot 7 \cdot 6)$$

If the temperature is so high that $kT \gg \hbar \omega$ or $T \gg \Theta_E$, then $\Theta_E/T \ll 1$ and expansion of the exponentials in (7 · 7 · 5) yields again the classical result

$$\text{for } T \gg \Theta_E, \quad c_V \rightarrow 3R \quad (7 \cdot 7 \cdot 7)$$

On the other hand, if the temperature is so low that $kT \ll \hbar \omega$ or $T \ll \Theta_E$ then $\Theta_E/T \gg 1$ and the exponential factor becomes very large compared to unity. The specific heat then becomes quite small; more precisely,

$$\text{for } T \ll \Theta_E, \quad c_V \rightarrow 3R \left(\frac{\Theta_E}{T} \right)^2 e^{-\Theta_E/T} \quad (7 \cdot 7 \cdot 8)$$

Thus the specific heat should approach zero exponentially as $T \rightarrow 0$. Experimentally the specific heat approaches zero more slowly than this, indeed $c_V \propto T^3$ as $T \rightarrow 0$. The reason for this discrepancy is the crude assumption that all atoms vibrate with the same characteristic frequency. In reality this is, of course, not the case (even if all the atoms are identical). The reason is that each atom does not vibrate separately as though it were experiencing a force due to stationary neighboring atoms; instead, there exist many different modes of motion in which various *groups* of atoms oscillate in phase at the same frequency. It is thus necessary to know the many different possible frequencies of these modes of oscillation (i.e., the values of all the coefficients κ_i in (7 · 7 · 1)). This problem will be considered in greater detail in Secs. 10 · 1 and 10 · 2. But it is qualitatively clear that, although T may be quite small, there are always some modes of oscillation (those corresponding to sufficiently large groups of atoms moving together) with a frequency ω so low that $\hbar \omega \ll kT$. These modes still contribute appreciably to the specific heat and thus prevent c_V from decreasing quite as rapidly as indicated by (7 · 7 · 8).

Nevertheless, the very simple Einstein approximation does give a reasonably good description of the specific heats of solids. It also makes clear the existence of a characteristic parameter Θ_E which depends on the properties of

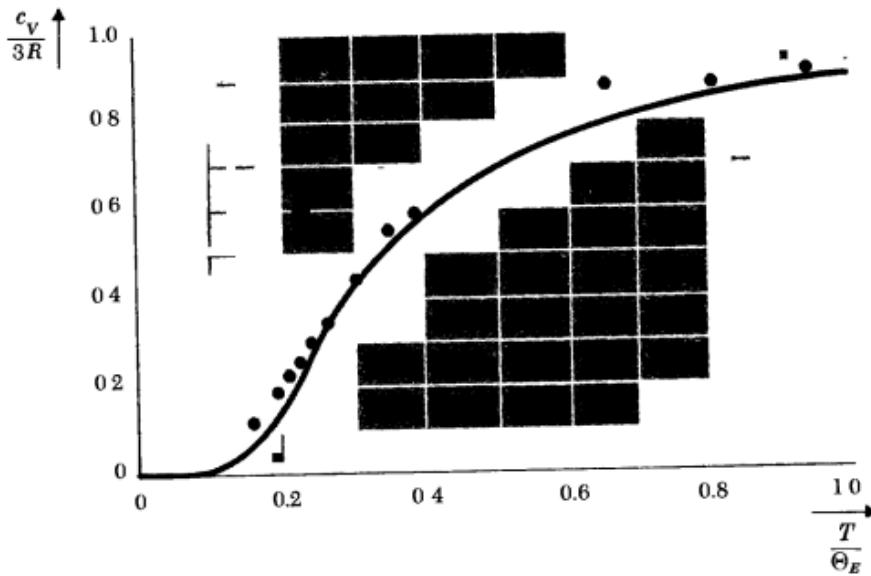


Fig. 7·7·1 Temperature dependence of c_V according to the Einstein model. The points are experimental values of c_V for diamond, the fit to the curve being achieved by choosing $\Theta_E = 1320^\circ\text{K}$ (after A. Einstein, Ann. Physik, vol. 22, p. 186 (1907)).

the solid under consideration. For example, if a solid has atoms of low molecular weight and is hard (i.e., relatively incompressible), this implies that each oscillator has a small mass m and a large spring constant κ_0 (i.e., the spring is stiff). Then (7·6·7) shows that the angular frequency of vibration ω of the atoms is large, or that Θ_E defined in (7·7·6) is large. Thus one must go to higher temperatures before the classical limit $c_V = 3R$ is reached. This explains why a solid such as diamond, which consists of relatively light carbon atoms and is quite hard, has at room temperature a specific heat c_V , which is still considerably smaller than the classical value $3R$ (see Table 7·7·1). Thus for diamond a reasonably good fit with experiment can be obtained by choosing $\Theta_E = 1320^\circ\text{K}$ (see Fig. 7·7·1). For most other solids Θ_E lies closer to $\Theta_E \approx 300^\circ\text{K}$. This corresponds to a frequency of vibration $\omega/2\pi \approx k\Theta_E/(2\pi\hbar)$ of about 6×10^{12} cycles/sec, i.e., to a frequency in the infrared region of the electromagnetic spectrum.

Before the introduction of quantum ideas it was not possible to understand why the molar specific heats of solids should fall below the classical equipartition value $3R$ at low temperatures. In 1907 Einstein's theory clarified the mystery and helped to gain acceptance for the new quantum concepts.

PARAMAGNETISM

7 · 8 General calculation of magnetization

We have considered a simple example of paramagnetism in Sec. 6 · 3. Here we shall discuss the general case of arbitrary spin.

Consider a system consisting of N noninteracting atoms in a substance at absolute temperature T and placed in an external magnetic field \mathbf{H} pointing along the z direction. Then the magnetic energy of an atom can be written as

$$\epsilon = -\mathbf{\mu} \cdot \mathbf{H} \quad (7 \cdot 8 \cdot 1)$$

Here $\mathbf{\mu}$ is the magnetic moment of the atom. It is proportional to the total angular momentum $\hbar \mathbf{J}$ of the atom and is conventionally written in the form

$$\mathbf{\mu} = g\mu_0 \mathbf{J} \quad (7 \cdot 8 \cdot 2)$$

where μ_0 is a standard unit of magnetic moment (usually the Bohr magneton $\mu_0 = e\hbar/(2mc)$, m being the electron mass) and where g is a number of the order of unity, the so-called g factor of the atom.*

Remark Strictly speaking the magnetic field \mathbf{H} used in (7 · 8 · 1) is the local magnetic field acting on the atom. It is not quite the same as the external magnetic field since it includes also the magnetic field produced by all the other atoms. Suitable corrections for the difference can be made by standard arguments of electromagnetic theory. The distinction between external and local field becomes increasingly unimportant when the concentration of magnetic atoms is kept small.

By combining (7 · 8 · 1) and (7 · 8 · 2) one obtains

$$\epsilon = -g\mu_0 \mathbf{J} \cdot \mathbf{H} = -g\mu_0 H J_z \quad (7 \cdot 8 \cdot 3)$$

since \mathbf{H} points in the z direction. In a quantum-mechanical description the values which J_z can assume are discrete and are given by

$$J_z = m$$

where m can take on all values between $-J$ and $+J$ in integral steps; i.e.,

$$m = -J, -J + 1, -J + 2, \dots, J - 1, J \quad (7 \cdot 8 \cdot 4)$$

Thus there are $2J + 1$ possible values of m corresponding to that many possible projections of the angular momentum vector along the z axis. By virtue of

* In the case of atoms having both electronic spin and orbital angular momentum, g would be the Landé g factor.

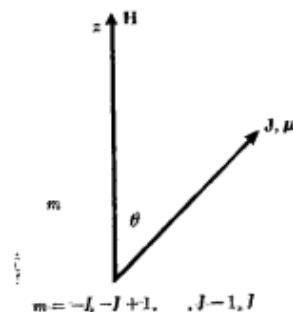


Fig. 7·8·1 Relative orientation of the angular momentum J with respect to H .

(7·8·3), the possible magnetic energies of the atom are then

$$\epsilon_m = -g\mu_0 H m \quad (7\cdot8\cdot5)$$

For example, if $J = \frac{1}{2}$ as would be the case for an atom with a single net electron spin, there are only two possible energies corresponding to $m = \pm \frac{1}{2}$. This was the simple case treated in Sec. 6·3.

The probability P_m that an atom is in a state labeled m is given by

$$P_m \propto e^{-\beta \epsilon_m} = e^{\beta g\mu_0 H m}$$

The z component of its magnetic moment in this state is, by (7·8·2), equal to

$$\mu_z = g\mu_0 m$$

The mean z component of the magnetic moment of an atom is therefore

$$\bar{\mu}_z = \frac{\sum_{m=-J}^J e^{\beta g\mu_0 H m} (g\mu_0 m)}{\sum_{m=-J}^J e^{\beta g\mu_0 H m}} \quad (7\cdot8\cdot6)$$

Here the numerator can conveniently be written as a derivative with respect to the external parameter H , i.e.,

$$\sum_{m=-J}^J e^{\beta g\mu_0 H m} (g\mu_0 m) = \frac{1}{\beta} \frac{\partial Z_a}{\partial H}$$

where

$$Z_a \equiv \sum_{m=-J}^J e^{\beta g\mu_0 H m} \quad (7\cdot8\cdot7)$$

is the partition function of one atom. Hence (7·8·6) becomes*

$$\bar{\mu}_z = \frac{1}{\beta} \frac{1}{Z_a} \frac{\partial Z_a}{\partial H} = \frac{1}{\beta} \frac{\partial \ln Z_a}{\partial H} \quad (7\cdot8\cdot8)$$

* This expression is valid even if the dependence of the energy levels of an atom on H is more complicated than in (7·8·5). See Problem 11.1.

To calculate Z_a , introduce the abbreviation

$$\eta \equiv \beta g\mu_0 H = \frac{g\mu_0 H}{kT} \quad (7 \cdot 8 \cdot 9)$$

which is a dimensionless parameter which measures the ratio of the magnetic energy $g\mu_0 H$, which tends to align the magnetic moment, to the thermal energy kT , which tends to keep it randomly oriented. Thus (7·8·7) becomes

$$Z_a = \sum_{m=-J}^J e^{\eta m} = e^{-\eta J} + e^{-\eta(J-1)} + \dots + e^{\eta J}$$

which is simply a finite geometric series where each term is obtained from the preceding one as a result of multiplication by e^η . This can immediately be summed to give

$$Z_a = \frac{e^{-\eta J} - e^{\eta(J+1)}}{1 - e^\eta}$$

This can be brought to more symmetrical form by multiplying both numerator and denominator by $e^{-\eta/2}$. Then

$$Z_a = \frac{e^{-\eta(J+\frac{1}{2})} - e^{\eta(J+\frac{1}{2})}}{e^{-\frac{1}{2}\eta} - e^{\frac{1}{2}\eta}}$$

or

$$\blacktriangleright Z_a = \frac{\sinh(J + \frac{1}{2})\eta}{\sinh \frac{1}{2}\eta} \quad (7 \cdot 8 \cdot 10)$$

where we have used the definition of the hyperbolic sine

$$\sinh y \equiv \frac{e^y - e^{-y}}{2} \quad (7 \cdot 8 \cdot 11)$$

Thus $\ln Z_a = \ln \sinh(J + \frac{1}{2})\eta - \ln \sinh \frac{1}{2}\eta \quad (7 \cdot 8 \cdot 12)$

By (7·8·8) and (7·8·9) one then obtains

$$\bar{\mu}_z = \frac{1}{\beta} \frac{\partial \ln Z_a}{\partial H} = \frac{1}{\beta} \frac{\partial \ln Z_a}{\partial \eta} \frac{\partial \eta}{\partial H} = g\mu_0 \frac{\partial \ln Z_a}{\partial \eta}$$

Hence

$$\bar{\mu}_z = g\mu_0 \left[\frac{(J + \frac{1}{2}) \cosh(J + \frac{1}{2})\eta}{\sinh(J + \frac{1}{2})\eta} - \frac{\frac{1}{2} \cosh \frac{1}{2}\eta}{\sinh \frac{1}{2}\eta} \right]$$

or

$$\blacktriangleright \bar{\mu}_z = g\mu_0 J B_J(\eta) \quad (7 \cdot 8 \cdot 13)$$

where

$$\blacktriangleright B_J(\eta) \equiv \frac{1}{J} \left[\left(J + \frac{1}{2} \right) \coth \left(J + \frac{1}{2} \right) \eta - \frac{1}{2} \coth \frac{1}{2} \eta \right] \quad (7 \cdot 8 \cdot 14)$$

The function $B_J(\eta)$ thus defined is sometimes called the “Brillouin function.” Let us investigate its limiting behavior for large and small values of the parameter η .

The hyperbolic cotangent is defined as

$$\coth y \equiv \frac{\cosh y}{\sinh y} = \frac{e^y + e^{-y}}{e^y - e^{-y}} \quad (7 \cdot 8 \cdot 15)$$

$$\text{For } y \gg 1, \quad e^{-y} \ll e^y \quad \text{and} \quad \coth y = 1 \quad (7 \cdot 8 \cdot 16)$$

Conversely, for $y \ll 1$, both e^y and e^{-y} can be expanded in power series. Retaining all terms quadratic in y , the result is

$$\begin{aligned} \coth y &= \frac{1 + \frac{1}{2}y^2 + \dots}{y + \frac{1}{6}y^3 + \dots} \\ &= \left(1 + \frac{1}{2}y^2\right) \left[\frac{1}{y} \left(1 + \frac{1}{6}y^2\right)^{-1} \right] \\ &= \frac{1}{y} \left(1 + \frac{1}{2}y^2\right) \left(1 - \frac{1}{6}y^2\right) \\ &= \frac{1}{y} \left(1 + \frac{1}{3}y^2\right) \end{aligned}$$

$$\text{For } y \ll 1, \quad \coth y = \frac{1}{y} + \frac{1}{3}y \quad (7 \cdot 8 \cdot 17)$$

Applying these results to the function $B_J(\eta)$ defined in (7 · 8 · 14) yields

$$\text{for } \eta \gg 1, \quad B_J(\eta) = \frac{1}{J} \left[\left(J + \frac{1}{2} \right) - \frac{1}{2} \right] = 1 \quad (7 \cdot 8 \cdot 18)$$

In the opposite limit where $\eta \ll 1$,

$$\begin{aligned} B_J(\eta) &= \frac{1}{J} \left\{ \left(J + \frac{1}{2} \right) \left[\frac{1}{(J + \frac{1}{2})\eta} + \frac{1}{3} \left(J + \frac{1}{2} \right) \eta \right] - \frac{1}{2} \left[\frac{2}{\eta} + \frac{\eta}{6} \right] \right\} \\ &= \frac{1}{J} \left\{ \frac{1}{3} \left(J + \frac{1}{2} \right)^2 \eta - \frac{1}{12} \eta \right\} \\ &= \frac{\eta}{3J} \left\{ J^2 + J + \frac{1}{4} - \frac{1}{4} \right\} \end{aligned}$$

$$\text{For } \eta \ll 1, \quad B_J(\eta) = \frac{(J + 1)}{3} \eta \quad (7 \cdot 8 \cdot 19)$$

Figure 7 · 8 · 2 shows how $B_J(\eta)$ depends on η for various values of J .

If there are N_0 atoms per unit volume, the mean magnetic moment per unit volume (or magnetization) becomes by (7 · 8 · 13)

$$\blacktriangleright \quad \bar{M}_z = N_0 \bar{\mu}_z = N_0 g \mu_0 J B_J(\eta) \quad (7 \cdot 8 \cdot 20)$$

If $\eta \ll 1$, (7 · 8 · 19) implies that $\bar{M}_z \propto \eta \propto H/T$. One can write this relation

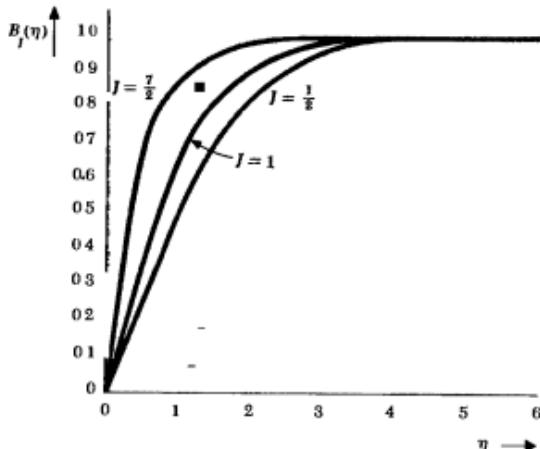


Fig. 7·8·2 Dependence of the Brillouin function $B_J(\eta)$ on its argument η for various values of J .

in the following form:

$$\text{for } g\mu_0 H/kT \ll 1, \quad \bar{M}_z = \chi H \quad (7\cdot8\cdot21)$$

where the constant of proportionality, i.e., the susceptibility χ , is given by

$$\blacktriangleright \quad \chi = N_0 \frac{g^2 \mu_0^2 J(J+1)}{3kT} \quad (7\cdot8\cdot22)$$

Thus $\chi \propto T^{-1}$, a result known as Curie's law. In the other limiting case

$$\text{when } g\mu_0 H/kT \gg 1, \quad \bar{M}_z \rightarrow N_0 g\mu_0 J \quad (7\cdot8\cdot23)$$

One gets then saturation behavior where each atom has the maximum z component of magnetic moment, $g\mu_0 J$, that it can possibly have.

Although the general results (7·8·20) and (7·8·21) are quite important, all the physical ideas are exactly the same as those already discussed in Sec. 6·3 for the special simple case of $J = \frac{1}{2}$. Note that our discussion is equally valid if the total angular momentum \mathbf{J} and magnetic moment $\boldsymbol{\mu}$ of the atom are due to unpaired electrons of the atom (e.g., a gadolinium or iron atom); or if the atom has no unpaired electrons and \mathbf{J} and $\boldsymbol{\mu}$ are due solely to the nucleus of the atom (e.g., a He^3 atom or fluorine F^- ion). The difference is one of magnitudes. In the first case $\boldsymbol{\mu}$ is of the order of a Bohr magneton. But in the second case the magnetic moment is smaller by approximately the ratio of the electron to the nucleon mass; i.e., it is of the order of a nuclear magneton, about 1000 times smaller than the Bohr magneton. Nuclear paramagnetism is thus about 1000 times smaller than electronic paramagnetism. Correspondingly, it requires an absolute temperature about 1000 times smaller to achieve the same extent of preferential nuclear spin orientation along an applied magnetic field as it does to achieve this extent of orientation for an electronic spin (see Problem 6.3).

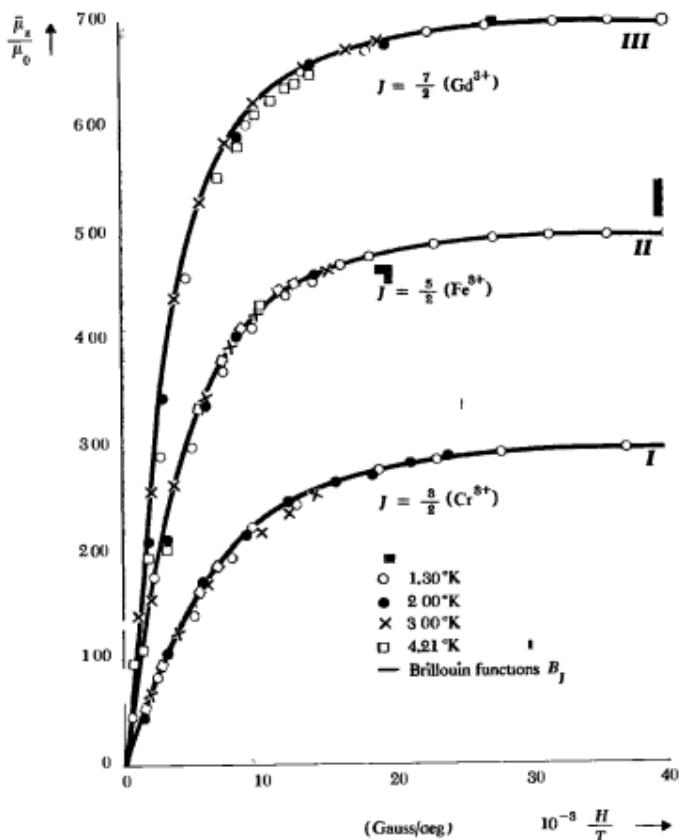


Fig. 7·8·3 Plots of the mean magnetic moment $\bar{\mu}_s$ of an ion (in units of the Bohr magneton μ_0) as a function of H/T . The solid curves are Brillouin functions. The experimental points are those for (I) potassium chromium alum, (II) iron ammonium alum, and (III) gadolinium sulfate octahydrate. In all cases, $J = S$, the total electron spin of the ion, and $g = 2$. Note that at 1.3°K a field of 50,000 gauss is sufficient to produce more than 99.5 percent magnetic saturation (after W. E. Henry, Phys. Rev. vol. 88, p. 561 (1952)).

KINETIC THEORY OF DILUTE GASES IN EQUILIBRIUM

7 · 9 Maxwell velocity distribution

Consider a molecule of mass m in a dilute gas. The gas may possibly consist of several different kinds of molecules; the molecule under consideration may also be polyatomic. Let us denote the position of the center of mass of this molecule by \mathbf{r} and the momentum of its center of mass by \mathbf{p} . If external force fields (e.g., gravity) are neglected, the energy ϵ of this molecule is equal to

$$\epsilon = \frac{\mathbf{p}^2}{2m} + \epsilon^{(\text{int})} \quad (7 \cdot 9 \cdot 1)$$

where the first term on the right is the kinetic energy of the center of mass motion; the second term, which arises only if the molecule is not monatomic, designates the internal energy of rotation and vibration of the atoms with respect to the molecular center of mass. Since the gas is supposed to be sufficiently dilute to be considered ideal, any potential energy of interaction with other molecules is supposed to be negligible; thus ϵ does not depend on \mathbf{r} .

The translational degrees of freedom can be treated classically to an excellent approximation if the gas is dilute and the temperature is not too low; the internal degrees of freedom must usually be treated by quantum mechanics. The state of the molecule can be described by specifying that the position of the center of mass of the molecule lies in the range $(\mathbf{r}; d\mathbf{r})$, i.e., in a volume element of magnitude $d^3\mathbf{r} = dx dy dz$ near the position \mathbf{r} ; that the momentum of its center of mass lies in the range $(\mathbf{p}; d\mathbf{p})$, i.e., within the momentum space volume $d^3\mathbf{p} \equiv dp_x dp_y dp_z$ near the momentum \mathbf{p} ; and that the state of internal motion of the molecule is labeled by some quantum numbers s with corresponding internal energy $\epsilon_s^{(int)}$. This particular molecule is in weak interaction with all the other molecules which act, therefore, as a heat reservoir at the temperature T of the gas. If the gas is sufficiently dilute, it is also permissible to think in classical terms and to focus attention on the particular molecule as a distinguishable entity. Then the molecule satisfies all the conditions of a distinct small system in contact with a heat reservoir and obeys, therefore, the canonical distribution. Hence one obtains for the probability $P_s(\mathbf{r}, \mathbf{p}) d^3\mathbf{r} d^3\mathbf{p}$ of finding the molecule with center-of-mass variables in the ranges $(\mathbf{r}; d\mathbf{r})$ and $(\mathbf{p}; d\mathbf{p})$ and with internal state specified by s the result

$$\begin{aligned} P_s(\mathbf{r}, \mathbf{p}) d^3\mathbf{r} d^3\mathbf{p} &\propto e^{-\beta[\mathbf{p}^2/2m + \epsilon_s^{(int)}]} d^3\mathbf{r} d^3\mathbf{p} \\ &\propto e^{-\beta\mathbf{p}^2/2m} e^{-\beta\epsilon_s^{(int)}} d^3\mathbf{r} d^3\mathbf{p} \end{aligned} \quad (7 \cdot 9 \cdot 2)$$

The probability $P(\mathbf{r}, \mathbf{p}) d^3\mathbf{r} d^3\mathbf{p}$ of finding the molecule with center-of-mass variables in the ranges $(\mathbf{r}; d\mathbf{r})$ and $(\mathbf{p}; d\mathbf{p})$, irrespective of its internal state, is obtained by summing (7 · 9 · 2) over all possible internal states s . The sum over the factor $\exp(-\beta\epsilon_s^{(int)})$ contributes then only a constant of proportionality, so that the result of summing (7 · 9 · 2) is simply

$$P(\mathbf{r}, \mathbf{p}) d^3\mathbf{r} d^3\mathbf{p} \propto e^{-\beta(\mathbf{p}^2/2m)} d^3\mathbf{r} d^3\mathbf{p} \quad (7 \cdot 9 \cdot 3)$$

This is, of course, identical with the result (6 · 3 · 11) derived previously under less general conditions.

If one multiplies the probability (7 · 9 · 3) by the total number N of molecules of the type under consideration, one obtains the mean number of molecules in this position and momentum range. Let us express the result in terms of the velocity $\mathbf{v} = \mathbf{p}/m$ of the molecule's center of mass. For the type of molecule under consideration, we define

$$f(\mathbf{r}, \mathbf{v}) d^3\mathbf{r} d^3\mathbf{v} \equiv \text{the mean number of molecules with center of mass position between } \mathbf{r} \text{ and } \mathbf{r} + d\mathbf{r}, \text{ and velocity between } \mathbf{v} \text{ and } \mathbf{v} + d\mathbf{v}. \quad (7 \cdot 9 \cdot 4)$$

Then (7·9·3) gives

$$f(\mathbf{r}, \mathbf{v}) d^3\mathbf{r} d^3\mathbf{v} = C N e^{-\beta(m\mathbf{v}^2/2)} d^3\mathbf{r} d^3\mathbf{v} \quad (7\cdot9\cdot5)$$

where C is a constant of proportionality which can be determined by the normalization condition

$$\int_{(r)} \int_{(v)} f(\mathbf{r}, \mathbf{v}) d^3\mathbf{r} d^3\mathbf{v} = N \quad (7\cdot9\cdot6)$$

That is, summing over molecules with all possible velocities \mathbf{v} from $-\infty$ to ∞ and with all possible positions \mathbf{r} anywhere in the volume V of the container must yield the *total* number of molecules. Substituting (7·9·5) in (7·9·6) thus gives

$$C \int_{(r)} \int_{(v)} e^{-\beta(m\mathbf{v}^2/2)} d^3\mathbf{v} d^3\mathbf{r} = N \quad (7\cdot9\cdot7)$$

Since f does *not* depend on \mathbf{r} , the integration over this variable yields simply the volume V . The rest of the integration is similar to that leading to (7·6·7). Thus (7·9·7) reduces to

$$CV \left(\int_{-\infty}^{\infty} e^{-\frac{1}{2}\beta m v_x^2} dv_x \right)^3 = CV \left(\frac{2\pi}{\beta m} \right)^{\frac{3}{2}} = N$$

or
$$C = n \left(\frac{\beta m}{2\pi} \right)^{\frac{3}{2}}, \quad n = \frac{N}{V} \quad (7\cdot9\cdot8)$$

where n is the total number of molecules (of this type) per unit volume. Hence (7·9·5) becomes

$$f(\mathbf{v}) d^3\mathbf{r} d^3\mathbf{v} = n \left(\frac{\beta m}{2\pi} \right)^{\frac{3}{2}} e^{-\frac{1}{2}\beta m \mathbf{v}^2} d^3\mathbf{r} d^3\mathbf{v} \quad (7\cdot9\cdot9)$$

or

$$\blacktriangleright f(\mathbf{v}) d^3\mathbf{r} d^3\mathbf{v} = n \left(\frac{m}{2\pi kT} \right)^{\frac{3}{2}} e^{-m\mathbf{v}^2/2kT} d^3\mathbf{r} d^3\mathbf{v} \quad (7\cdot9\cdot10)$$

Here we have omitted the variable \mathbf{r} in the argument of f since f does not depend on \mathbf{r} . This condition must, of course, be true from symmetry considerations since there is no preferred position in space in the absence of external force fields. Furthermore one sees that f depends only on the magnitude of \mathbf{v} ; i.e.,

$$f(\mathbf{v}) = f(v) \quad (7\cdot9\cdot11)$$

where $v = |\mathbf{v}|$. Again this is obvious by symmetry, since there is no preferred direction in a situation where the container, and thus also the center of mass of the whole gas, is considered to be at rest.

If (7·9·10) is divided by the element of volume $d^3\mathbf{r}$, one obtains

$$f(\mathbf{v}) d^3\mathbf{v} = \left. \begin{aligned} & \text{the mean number of molecules per unit volume} \\ & \text{with center-of-mass velocity in the range between } \mathbf{v} \text{ and } \mathbf{v} + d\mathbf{v}. \end{aligned} \right\} \quad (7\cdot9\cdot12)$$

Equation (7·9·10) is the Maxwell velocity distribution for a molecule of a dilute gas in thermal equilibrium.

7 · 10 Related velocity distributions and mean values

Distribution of a component of velocity Various other distributions of physical interest follow immediately from (7 · 9 · 10). For example, one may be interested in the quantity

$$\left. \begin{aligned} g(v_x) dv_x &= \text{the mean number of molecules per unit volume} \\ &\text{with } x \text{ component of velocity in the range between } v_x \text{ and} \\ &v_x + dv_x, \text{ irrespective of the values of their other velocity} \\ &\text{components.} \end{aligned} \right\} \quad (7 \cdot 10 \cdot 1)$$

Clearly, one obtains this by adding up all molecules with x component of velocity in this range; i.e.,

$$g(v_x) dv_x = \int_{(v_y)} \int_{(v_z)} f(v) d^3 v \quad (7 \cdot 10 \cdot 2)$$

where one sums over all possible y and z velocity components of the molecules. By (7 · 9 · 10) this becomes

$$\begin{aligned} g(v_x) dv_x &= n \left(\frac{m}{2\pi kT} \right)^{\frac{1}{2}} \int_{(v_y)} \int_{(v_z)} e^{-(m/2kT)(v_x^2 + v_y^2 + v_z^2)} dv_x dv_x dv_z \\ &= n \left(\frac{m}{2\pi kT} \right)^{\frac{1}{2}} e^{-mv_x^2/2kT} dv_x \int_{-\infty}^{\infty} e^{-(m/2kT)v_y^2} dv_y \int_{-\infty}^{\infty} e^{-(m/2kT)v_z^2} dv_z \\ &= n \left(\frac{m}{2\pi kT} \right)^{\frac{1}{2}} e^{-mv_x^2/2kT} dv_x \left(\sqrt{\frac{2\pi kT}{m}} \right)^2 \end{aligned}$$

or

$$\blacktriangleright \quad g(v_x) dv_x = n \left(\frac{m}{2\pi kT} \right)^{\frac{1}{2}} e^{-mv_x^2/2kT} dv_x \quad (7 \cdot 10 \cdot 3)$$

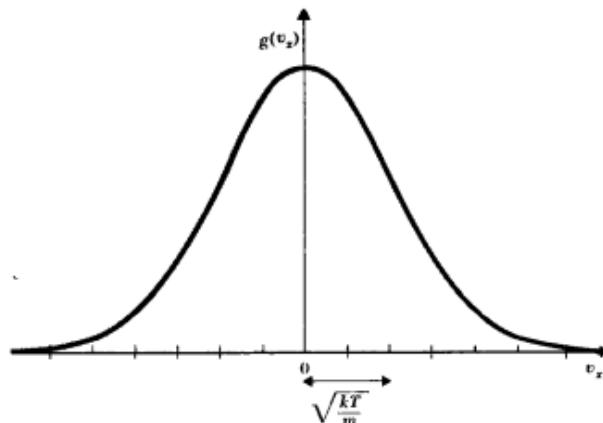


Fig. 7 · 10 · 1 Maxwellian distribution of a molecular velocity component.

Of course, this expression is properly normalized so that

$$\int_{-\infty}^{\infty} g(v_x) dv_x = n \quad (7 \cdot 10 \cdot 4)$$

Equation (7 · 10 · 3) shows that each component of velocity is distributed with a Gaussian distribution *symmetric* about the mean value

$$\bar{v}_x = 0 \quad (7 \cdot 10 \cdot 5)$$

It is physically clear that $\bar{v}_x = 0$ by symmetry, since the x component of velocity of a molecule is as likely to be positive as negative. Mathematically this result follows, since

$$\bar{v}_x = \frac{1}{n} \int_{-\infty}^{\infty} g(v_x) v_x dv_x$$

Here the integrand is an odd function of v_x (i.e., reverses its sign when v_x reverses sign) because $g(v_x)$ is an even function of v_x (i.e., does not reverse its sign under this operation, since it depends only on v_x^2). Thus contributions to the integral from $+v_x$ and $-v_x$ cancel each other. A similar argument shows immediately that

$$\text{if } k \text{ is any odd integer,} \quad \overline{v_x^k} = 0 \quad (7 \cdot 10 \cdot 6)$$

Of course, $\overline{v_x^2}$ is intrinsically positive and is, by virtue of (7 · 10 · 5), the dispersion of v_x . By direct integration, using (7 · 10 · 3), or by recalling the properties of the Gaussian distribution already studied in (1 · 6 · 9), it follows that

$$\overline{v_x^2} = \frac{1}{n} \int_{-\infty}^{\infty} g(v_x) v_x^2 dv_x = \frac{kT}{m} \quad (7 \cdot 10 \cdot 7)$$

The same result follows, of course, also immediately from the equipartition theorem, according to which

$$\overline{\frac{1}{2}mv_x^2} = \frac{1}{2}kT$$

Thus the root-mean-square width of the Gaussian (7 · 10 · 3) is given by $\Delta^* v_x = \sqrt{kT/m}$. The lower the temperature, the narrower will be the width of the distribution function $g(v_x)$.

Needless to say, exactly the same results hold for v_y and v_z , since all velocity components are, by the symmetry of the problem, completely equivalent.

Note also that, since $v^2 = v_x^2 + v_y^2 + v_z^2$, Eq. (7 · 9 · 10) factors so that it can be written in the form

$$\frac{f(\mathbf{v}) d^3\mathbf{v}}{n} = \left[\frac{g(v_x) dv_x}{n} \right] \left[\frac{g(v_y) dv_y}{n} \right] \left[\frac{g(v_z) dv_z}{n} \right]$$

This means that the probability that the velocity lies in the range between \mathbf{v} and $\mathbf{v} + d\mathbf{v}$ is just equal to the product of the probabilities that the velocity components lie in their respective ranges. Thus the individual velocity components behave like statistically independent quantities.

Distribution of speed Another quantity of physical interest is

$$F(v) dv = \text{the mean number of molecules per unit volume} \\ \text{with a speed } v = |\mathbf{v}| \text{ in the range between } v \text{ and } v + dv \quad (7 \cdot 10 \cdot 8)$$

One clearly obtains this quantity by adding up all molecules with speed in this range irrespective of the *direction* of their velocity. Thus

$$F(v) dv = f'f(\mathbf{v}) d^3\mathbf{v}$$

where the integral extends over all velocities, satisfying the condition that

$$v < |\mathbf{v}| < v + dv$$

i.e., over all velocity vectors which terminate in velocity space within a spherical shell of inner radius v and outer radius $v + dv$. Since $f(\mathbf{v})$ depends only on $|\mathbf{v}|$, this integration is just equal to $f(v)$ multiplied by the volume $4\pi v^2 dv$ of this spherical shell. Thus

► $F(v) dv = 4\pi f(v)v^2 dv \quad (7 \cdot 10 \cdot 9)$

Using (7 · 9 · 10), this becomes explicitly

► $F(v) dv = 4\pi n \left(\frac{m}{2\pi kT} \right)^{\frac{3}{2}} v^2 e^{-mv^2/2kT} dv \quad (7 \cdot 10 \cdot 10)$

This is the Maxwell distribution of speeds. Note that it has a maximum for the same reason responsible for the maxima encountered in our general discussion of statistical mechanics. As v increases, the exponential factor *decreases*, but the volume of phase space available to the molecule is proportional to v^2 and *increases*; the net result is a gentle maximum. The expression (7 · 10 · 10) is, of course, properly normalized, i.e.,

$$\int_0^\infty F(v) dv = n \quad (7 \cdot 10 \cdot 11)$$

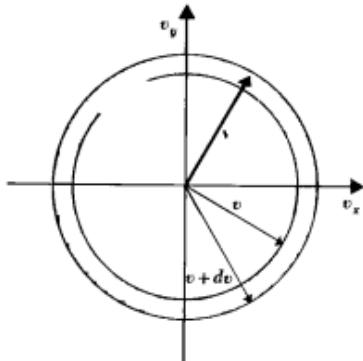


Fig. 7 · 10 · 2 Representation (in two dimensions) of the shell in velocity space containing all molecules with velocity v such that $v < |\mathbf{v}| < v + dv$.

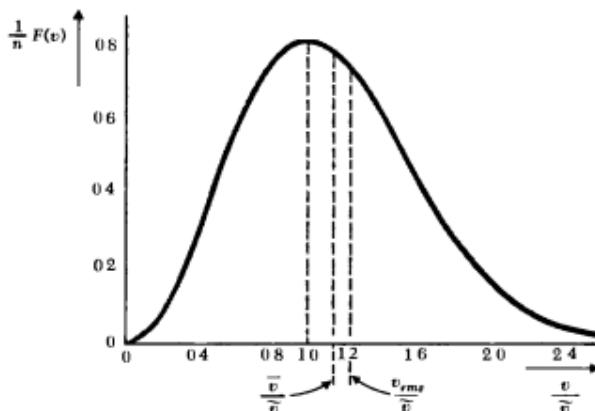


Fig. 7·10·3 Maxwellian distribution of molecular speeds. The speed v is expressed in terms of the speed $v = (2kT/m)^{1/2}$ where F is maximum.

Mean values It is again of interest to calculate some significant mean values. The mean speed is, of course, positive since $v = |v|$ is intrinsically positive. It is given by

$$\bar{v} = \frac{1}{n} \iiint f(v)v d^3v \quad (7 \cdot 10 \cdot 12)$$

where the integration is over all velocities, or equivalently by

$$\bar{v} = \frac{1}{n} \int_0^\infty F(v)v dv$$

where the integration is over all speeds. Thus one gets

$$\begin{aligned} \bar{v} &= \frac{1}{n} \int_0^\infty f(v)v \cdot 4\pi v^2 dv = \frac{4\pi}{n} \int_0^\infty f(v)v^3 dv \\ &= 4\pi \left(\frac{m}{2\pi kT}\right)^{1/2} \int_0^\infty e^{-mv^2/2kT} v^3 dv \quad \text{by (7·9 10)} \\ &= 4\pi \left(\frac{m}{2\pi kT}\right)^{1/2} \cdot \frac{1}{2} \left(\frac{m}{2kT}\right)^{-3/2} \quad \text{by (A 4·6)} \end{aligned}$$

Hence

$$\bar{v} = \sqrt{\frac{8kT}{\pi m}} \quad (7 \cdot 10 \cdot 13)$$

On the other hand, the mean-square speed is given by

$$\bar{v}^2 = \frac{1}{n} \int f(v)v^2 d^3v = \frac{4\pi}{n} \int_0^\infty f(v)v^4 dv \quad (7 \cdot 10 \cdot 14)$$

We could again integrate this by using (A 4·6), but we can save ourselves the work since we know that

$$\overline{\frac{1}{2}mv^2} = \overline{\frac{1}{2}m(v_x^2 + v_y^2 + v_z^2)}$$

Hence by the equipartition theorem (or by using the symmetry argument that

$\overline{v_x^2} = \overline{v_y^2} = \overline{v_z^2}$, so that $\overline{v^2} = 3\overline{v_z^2}$, a quantity already computed in (7·10·7)), we get

$$\frac{1}{2}mv^2 = \frac{3}{2}kT$$

or $\overline{v^2} = \frac{3kT}{m}$ (7·10·15)

The root-mean-square speed is thus

$$v_{\text{rms}} \equiv \sqrt{\overline{v^2}} = \sqrt{\frac{3kT}{m}} \quad (7 \cdot 10 \cdot 16)$$

Finally, one can ask for the most probable speed \tilde{v} of the molecule, i.e., the speed for which $F(v)$ in (7·10·10) is a maximum. This is given by the condition

$$\frac{dF}{dv} = 0$$

i.e., $2v e^{-mv^2/2kT} + v^2 \left(-\frac{m}{kT} v \right) e^{-mv^2/2kT} = 0$

or $v^2 = \frac{2kT}{m}$

Hence \tilde{v} is given by

$$\tilde{v} = \sqrt{\frac{2kT}{m}} \quad (7 \cdot 10 \cdot 17)$$

All these various speeds are proportional to $(kT/m)^{\frac{1}{2}}$. Thus the molecular speed increases when the temperature is raised; and, for a given temperature, a molecule with larger mass has smaller speed. The various speeds we have calculated are such that their ratios

$v_{\text{rms}} : \tilde{v} : \bar{v}$ are proportional to or to	$\left. \begin{array}{l} \sqrt{3} : \sqrt{\frac{8}{\pi}} : \sqrt{2} \\ 1.224 : 1.128 : 1 \end{array} \right\}$	$(7 \cdot 10 \cdot 18)$
--	--	-------------------------

For nitrogen (N_2) gas at room temperature ($300^\circ K$) one finds by (7·10·16), using $m = 28/(6 \times 10^{23})$ g, that

$$v_{\text{rms}} \approx 5 \times 10^4 \text{ cm/sec} \approx 500 \text{ m/sec} \quad (7 \cdot 10 \cdot 19)$$

a number of the order of the velocity of sound in the gas.

7 · 11 Number of molecules striking a surface

It is possible to discuss a number of interesting physical situations by considering the motion of individual molecules in detail. Detailed arguments of this kind constitute the subject matter of what is usually called "kinetic theory." Since we shall for the present restrict ourselves to equilibrium situations, our considerations will be very simple.

Let us focus attention on a dilute gas enclosed in a container and ask the following question: How many molecules per unit time strike a unit area of the wall of this container? This question is very closely related to another question of physical interest: If there is a very small hole in the wall of the container, how many molecules will stream out of this hole per unit time?

Crude calculation To understand the essential features of the situation, it is adequate to adopt a highly simplified approximate point of view. Imagine that the container is a box in the form of a parallelepiped, the area of one end-wall being A . How many molecules per unit time strike this end-wall? Suppose that there are in this gas n molecules per unit volume. Since they all move in random directions, we can say roughly that one-third of them, or $n/3$ molecules per unit volume, have their velocities in a direction predominantly along the z axis (chosen to be normal to the end-wall under consideration as in Fig. 7·11·1). Half of these molecules, i.e., $n/6$ molecules per unit volume, have their velocity in the $+z$ direction so that they will strike the end-wall under consideration. If the mean speed of the molecules is \bar{v} , these molecules cover in an infinitesimal time dt a mean distance $\bar{v} dt$. Hence all those molecules with velocity \bar{v} in the z direction which lie within a distance $\bar{v} dt$ from the end-wall, will, within a time dt , strike the end-wall; those which lie further than a distance $\bar{v} dt$ from the end-wall will not. Thus we arrive at the result that [the number of molecules which strike the end-wall of area A in time dt] is equal to [the number of molecules having velocity \bar{v} in the z direction and contained in the cylinder of volume $A\bar{v} dt$]; i.e., it is given by

$$\left(\frac{n}{6}\right)(A\bar{v} dt) \quad (7\cdot11\cdot1)$$

The total number Φ_0 of molecules which strike unit area of the wall per unit time (i.e., the total molecular "flux") is then given by dividing (7·11·1) by the area A and the time interval dt . Thus

$$\Phi_0 \approx \frac{1}{6}n\bar{v} \quad (7\cdot11\cdot2)$$

We emphasize that this result was obtained by a very crude argument in which we did not consider in any detail the velocity distribution of the molecules, either in magnitude or in direction. Nevertheless, arguments of this

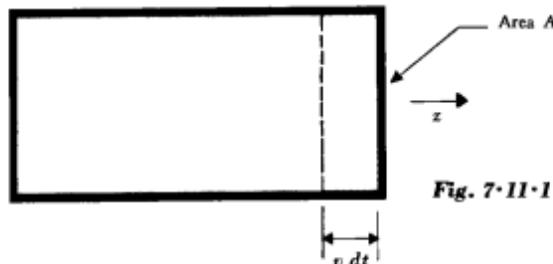


Fig. 7·11·1 Molecules colliding with a wall.

kind are often very useful because they can exhibit the essential features of a phenomenon without the need of exact calculations which are likely to be more laborious. Thus the factor $\frac{1}{6}$ in (7·11·2) is not to be taken very seriously. There ought to be some numerical factor there of the order of $\frac{1}{6}$, but its exact value must certainly depend on the particular way in which various averages are to be computed. (Indeed, the exact calculation will yield a factor $\frac{1}{4}$ instead of $\frac{1}{6}$.) On the other hand, the essential dependence of Φ_0 on n and \bar{v} ought to be correct, i.e., $\Phi_0 \propto n\bar{v}$. Thus (7·11·2) makes the very plausible assertion that Φ_0 is proportionately increased if the concentration of molecules is increased or if their speed is increased.

The dependence of Φ_0 on the temperature T and mean pressure \bar{p} of the gas follows immediately from (7·11·2). The equation of state gives

$$\bar{p} = nkT \quad \text{or} \quad n = \frac{\bar{p}}{kT} \quad (7\cdot11\cdot3)$$

Furthermore, by the equipartition theorem,

$$\frac{1}{2}mv^2 = \frac{3}{2}kT$$

so that $\bar{v} \propto \bar{v}_{\text{rms}} \propto \sqrt{\frac{kT}{m}}$ (7·11·4)

Thus (7·11·2) implies that

$$\Phi_0 \propto \frac{\bar{p}}{\sqrt{mT}} \quad (7\cdot11\cdot5)$$

Exact calculation Consider an element of area dA of the wall of the container. Choose the z axis so as to point along the outward normal of this element of area (see Fig. 7·11·2). Consider first those molecules (in the immediate vicinity of the wall) whose velocity is such that it lies between v and $v + dv$. (That is, the velocity is such that its magnitude lies in the range between v and $v + dv$; its direction, specified by its polar angle θ (with respect to the normal, or z axis) and its azimuthal angle φ , is such that these angles lie between θ and $\theta + d\theta$ and between φ and $\varphi + d\varphi$, respectively.)

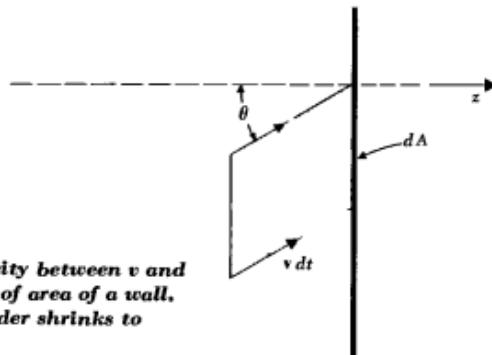


Fig. 7·11·2 Molecules, with velocity between v and $v + dv$, colliding with an element of area of a wall. (Note that the height of the cylinder shrinks to zero as $dt \rightarrow 0$.)

Molecules of this type suffer a displacement $v dt$ in the infinitesimal time interval dt . Hence all molecules of this type which lie within the infinitesimal cylinder of cross-sectional area dA and of a length $v dt$ making an angle θ with the z axis will strike the wall within the time interval dt ; the molecules lying outside this cylinder will not.* The volume of this cylinder is $dA v dt \cos \theta$, while the number of molecules per unit volume in this velocity range is $f(v) d^3v$. Hence [the number of molecules of this type which strike the area dA of the wall in time dt] = $[f(v) d^3v][dA v dt \cos \theta]$. Dividing this by the area dA and the time interval dt , we get for

$$\left. \begin{aligned} \Phi(v) d^3v &\equiv \text{the number of molecules, with velocity between } \\ v \text{ and } v + dv, \text{ which strike a unit area of the wall per unit} \\ \text{time.} \end{aligned} \right\} \quad (7 \cdot 11 \cdot 6)$$

the expression

$$\blacktriangleright \quad \Phi(v) d^3v = d^3v f(v)v \cos \theta \quad (7 \cdot 11 \cdot 7)$$

$$\left. \begin{aligned} \text{Let } \Phi_0 &\equiv \text{the total number of molecules which strike a unit} \\ \text{area of the wall per unit time.} \end{aligned} \right\} \quad (7 \cdot 11 \cdot 8)$$

This is simply given by summing (7 · 11 · 7) over all possible velocities which a molecule can have and which will cause it to collide with the element of area. This means that we have to sum over all possible speeds $0 < v < \infty$, all possible azimuthal angles $0 < \varphi < 2\pi$, and all angles θ in the range $0 < \theta < \pi/2$. (Those molecules for which $\pi/2 < \theta < \pi$ have their velocity directed away from the wall and hence will not collide with it.) In other words, we have to sum over all possible velocities v subject to the restriction that the velocity component $v_z = v \cos \theta > 0$ (since molecules with $v_z < 0$ will not collide with the element of area). Thus we have

$$\Phi_0 = \int_{v_z > 0} d^3v f(v)v \cos \theta \quad (7 \cdot 11 \cdot 9)$$

The results (7 · 11 · 7) and (7 · 11 · 9) are generally valid even if the gas is not in equilibrium (although f might then also be a function of r and t). But if we consider a gas in thermal equilibrium, $f(v) = f(v)$ is only a function of $|v|$. The element of volume in velocity space can be expressed in spherical coordinates

$$d^3v = v^2 dv (\sin \theta d\theta d\varphi)$$

where $\sin \theta d\theta d\varphi = d\Omega$ is just the element of solid angle. Hence (7 · 11 · 9)

* Note that since the length $v dt$ of the cylinder can be considered arbitrarily small, only molecules located in the immediate vicinity of the wall are involved in this argument. Furthermore, since $v dt$ can be made much smaller than the mean distance l traveled by a molecule before it collides with another molecule, collisions between molecules need not be considered in this argument; i.e., any molecule located in the cylinder and traveling toward the wall will indeed strike the wall without being deflected by a collision before it gets there.

becomes

$$\Phi_0 = \int_{v>0} v^2 dv \sin \theta d\theta d\varphi f(v) v \cos \theta \\ = \int_0^\infty f(v) v^3 dv \int_0^{\pi/2} \sin \theta \cos \theta d\theta \int_0^{2\pi} d\varphi$$

The integration over φ gives 2π , while the integral over θ yields the value $\frac{1}{2}$. Hence

$$\Phi_0 = \pi \int_0^\infty f(v) v^3 dv \quad (7 \cdot 11 \cdot 10)$$

This can be expressed in terms of the mean speed already computed in (7 10 12). Thus

$$\bar{v} = \frac{1}{n} \int d^3v f(v) v = \frac{1}{n} \int_0^\infty \int_0^\pi \int_0^{2\pi} (v^2 dv \sin \theta d\theta d\varphi) f(v) v$$

$$\text{or } \bar{v} = \frac{4\pi}{n} \int_0^\infty f(v) v^3 dv \quad (7 \cdot 11 \cdot 11)$$

since the integration over the angles θ and φ is just the total solid angle 4π about a point. Hence (7 11 10) can also be written

$$\blacktriangleright \quad \Phi_0 = \frac{1}{4} n \bar{v} \quad (7 \cdot 11 \cdot 12)$$

This rigorous result can be compared with our previous crude estimate in (7 11 2). We see that the latter was off by a factor of only $\frac{2}{3}$.

The mean speed was already computed from the Maxwell distribution in (7 10 13). Combining this with the equation of state (7 11 3), one obtains for (7 11 12)

$$\Phi_0 = \frac{\bar{p}}{\sqrt{2\pi mkT}} \quad (7 \cdot 11 \cdot 13)$$

7 · 12 Effusion

If a sufficiently small hole (or slit) is made in the wall of the container, the equilibrium of the gas inside the container is disturbed to a negligible extent. In that case the number of molecules which emerge through the small hole is the same as the number of molecules which would strike the area occupied by the hole if the latter were closed off. The process whereby molecules emerge through such a small hole is called "effusion."

One may ask how small the diameter D of the hole (or the width D of the slit) must be so that there is no appreciable effect on the equilibrium state of the gas. The typical dimension against which D is to be compared is the "mean free path" l , i.e., the mean distance which a molecule in the gas travels before it suffers a collision with another molecule. The concept

of mean free path will be discussed in greater detail in Chapter 12. Here it will suffice to make the obvious comment that, at a given temperature, l is inversely proportional to the number of molecules per unit volume. (At room temperature and atmospheric pressure, $l \approx 10^{-8}$ cm in a typical gas.) If $D \ll l$, the hole can be considered as very small. In that case, molecules will now and then emerge from the hole if their velocities happen to be in the right direction. When a few molecules escape through the hole, the remaining molecules in the container are then scarcely affected since l is so large. This is the case of "effusion."

On the other hand, if $D \gg l$, molecules suffer frequent collisions with each other within distances of the order of the hole size. When some molecules emerge through this hole (see Fig. 7·12·1), the molecules behind them are in an appreciably different situation. They no longer continue colliding with the molecules on the right which have just escaped through the hole, but they still suffer constant collisions with the molecules on the left. The net result is that the molecules near the hole experience, by virtue of these continuous molecular impacts, a net force to the right which causes them to acquire a drift velocity in the direction toward the hole. The resultant collective motion of all these molecules moving together is then analogous to the flow of water through the hole of a tank. In this case one has not effusion, but "hydrodynamic flow."

Let us consider the situation when the hole is sufficiently small so that molecules emerge through the hole by effusion. If a vacuum is maintained outside the container, the effusing molecules can be collimated further by additional slits so that one is left with a well-defined "molecular beam." Such molecular beams have been extensively used in experimental physics investigations because they provide one with the possibility of studying individual molecules under circumstances where interactions between them are negligible. The number of molecules which have speed in the range between v and $v + dv$ and which emerge per second from a small hole of area A into a solid angle range $d\Omega$ in the forward direction $\theta = 0$ is given by (7·11·7) as

$$\begin{aligned} A\Phi(v) d^3v &\propto A[f(v)v \cos \theta](v^2 dv d\Omega) \\ &\propto f(v)v^3 dv d\Omega \propto e^{-mv^2/2kT} v^3 dv d\Omega \end{aligned} \quad (7\cdot12\cdot1)$$

Note that this expression involves the factor v^3 , rather than the factor v^2 which occurs in the Maxwellian speed distribution (7·10·10).

Experiments on such a molecular beam can provide a direct test of the Maxwell velocity distribution by checking the prediction (7·12·1). Figure



Fig. 7·12·1 Formation of a molecular beam by effusing molecules.

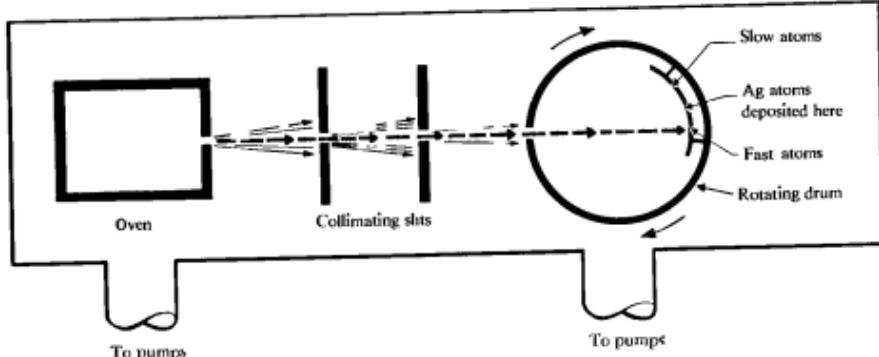


Fig. 7·12·2 A molecular-beam apparatus for studying the velocity distribution of silver (Ag) atoms. The Ag atoms stick to the drum surface upon impact.

7·12·2 shows one experimental arrangement which has been used. Silver (Ag) atoms are produced by evaporation in an oven and emerge through a narrow slit to form a molecular beam. A rotating hollow cylindrical drum, with a slit in it, revolves rapidly about its axis and is located in front of the beam. When molecules enter the slit in the drum, they require different times to reach the opposite side of the drum, a fast molecule requiring less time than a slow one. Since the drum is rotating all the time, the Ag molecules deposited on the opposite inside surface of the drum get spread out on this surface in accordance with their velocity distribution. Thus a measurement of the thickness of the Ag deposit as a function of distance along the drum surface provides a measurement of the molecular velocity distribution.

A more accurate method for determining the velocity distribution involves the use of a velocity selector similar in principle of operation to those used in neutron time-of-flight spectroscopy or for determining the velocity of light (Fizeau wheel). In this method the molecular beam emerges from a hole and is detected by a suitable device at the other end of the apparatus. The velocity selector is placed between the source and the detector and consists, in the simplest case, of a pair of disks mounted on a common axle which can be rotated with known angular velocity. Both disks are identical and have several slots cut along their periphery; thus the rotating disks act as two shutters which are alternately opened and closed. When the disks are properly aligned and not rotating, all molecules can reach the detector by passing through corresponding slots in both disks. But when the disks are rotating, molecules passing through a slot in the first disk can only reach the detector if their velocity is such that [the time of flight required for them to travel to the second disk] is equal to [the time required for the next slot of this disk to rotate to the location of the original slot]. Otherwise they will strike the solid part of the second disk and be stopped. Hence different angular velocities of rotation of the disks allow molecules of different speeds to reach the detector. Measurement of the relative number of molecules arriving there per second then

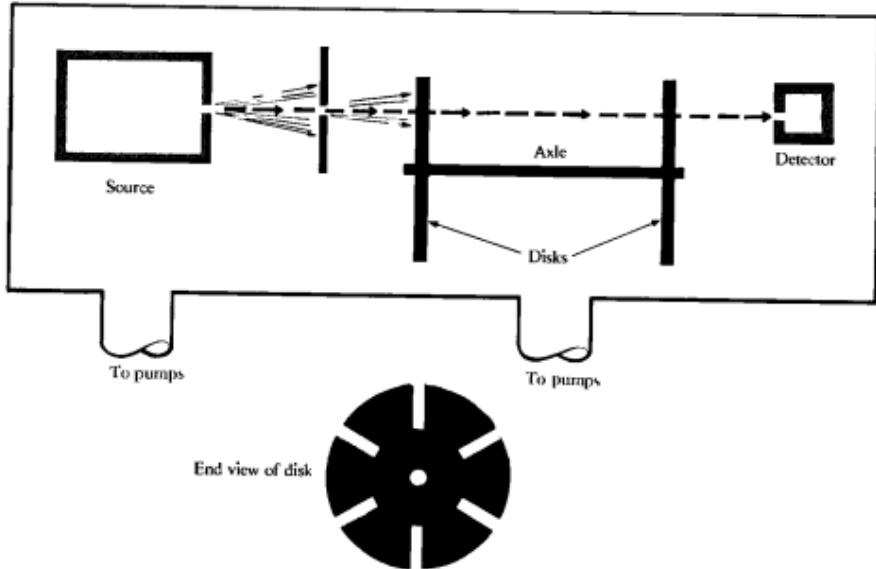


Fig. 7·12·3 Arrangement for studying the molecular velocity distribution by a velocity selector. (A more effective velocity selector results if more than two similar disks are mounted on the same axle.)

allows a direct check of the molecular-velocity distribution. The validity of the Maxwellian distribution has been well confirmed by such experiments.*

Equation (7·11·5) shows that the rate of effusion of a molecule depends on the mass of the molecule, lighter molecules effusing more rapidly than heavier ones. This suggests the application of effusion as a method for the separation of isotopes. Suppose that a container is closed off by a membrane which has very many small holes through which molecules can effuse. If this container is surrounded by a vacuum on the outside and is filled with a gas mixture of two isotopes at some initial time, then the relative concentration of the isotope of larger molecular weight will increase in the container as time goes on. Similarly, the gas pumped off from the surrounding vacuum will be more concentrated in the lighter isotope.†

Another example of interest is illustrated in Fig. 7·12·4. Here a container is divided into two parts by a partition containing a small hole. The container is filled with gas, but one part of the container is maintained at temperature T_1 , the other part at temperature T_2 . One may ask the following

* For recent experimental work on velocity distributions see R. C. Miller and P. Kusch' *J. Chem. Phys.*, vol. 25, p. 860 (1956); also P. M. Marcus and J. H. McFee in I. Estermann (ed.), "Recent Research in Molecular Beams," p. 43, Academic Press, New York, 1959.

† The successful large-scale separation by this method of uranium isotopes (in the form of UF_6 gas) was a crucial step in the development of nuclear fission devices (reactors and bombs) and is described in H. de W. Smyth, "Atomic Energy for Military Purposes," chap. 10, Princeton University Press, Princeton, 1947; or *Rev. Mod. Phys.* vol. 17, p. 430 (1945).

Fig. 7·12·4 A container divided into two parts by a partition containing a small hole. The gas in the two parts is at different temperatures and pressures.



question: What is the relation between the mean gas pressures \bar{p}_1 and \bar{p}_2 in the two parts when the system is in equilibrium, i.e., when a situation is reached where neither \bar{p}_1 or \bar{p}_2 , nor the amount of gas in either part, changes with time? If the linear dimension D of the hole is large ($D \gg l$), then the condition is simply $\bar{p}_2 = \bar{p}_1$; for otherwise the pressure difference would give rise to mass motion of the gas from one side to the other until the pressures on both sides reach equality. But if $D \ll l$, one deals with effusion through the hole rather than with hydrodynamic flow. In this case the equilibrium condition requires that the mass of gas on each side remain constant, i.e., that [the number of molecules which pass per second through the hole from left to right] equals [the number of molecules which pass per second through the hole from right to left]. By (7·11·2) this leads to the simple equality

$$n_1\bar{v}_1 = n_2\bar{v}_2 \quad (7\cdot12\cdot2)$$

By (7·11·5) this condition becomes

$$\frac{\bar{p}_1}{\sqrt{T_1}} = \frac{\bar{p}_2}{\sqrt{T_2}} \quad (7\cdot12\cdot3)$$

Thus the pressures are then not at all equal, but higher gas pressure prevails in the part of the container at higher temperature.

- This discussion has practical consequences in experimental work. Suppose, for example, that it is desired to measure the vapor pressure \bar{p}_v (i.e., the pressure of the vapor in equilibrium with the liquid) of liquid helium at 2°K . The experimental arrangement might be as illustrated in Fig. 7·12·5, the mercury manometer at room temperature being used to make the pressure measurement. A small tube of diameter D connects the manometer to the vapor pressure to be measured, and the difference of mercury levels on the two sides of the manometer measures the pressure difference \bar{p} . Now at 2°K , \bar{p}_v is still fairly large, i.e., the density of helium vapor is large enough that the mean free path l of molecules in the vapor is much less than the diameter D of the connecting tube. Then the pressure \bar{p} read on the manometer is indeed equal to the vapor pressure \bar{p}_v of interest. But suppose that it is desired to measure \bar{p}_v at lower temperatures, say at 0.5°K . Then \bar{p}_v is small and the density of He vapor is so small that l is comparable to, or large with respect to, the diameter D of the connecting tube. If one assumed that the pressure \bar{p} which one reads on the manometer still equals the vapor pressure \bar{p}_v of interest, one would be fooling oneself very badly. So-called "thermomolecular

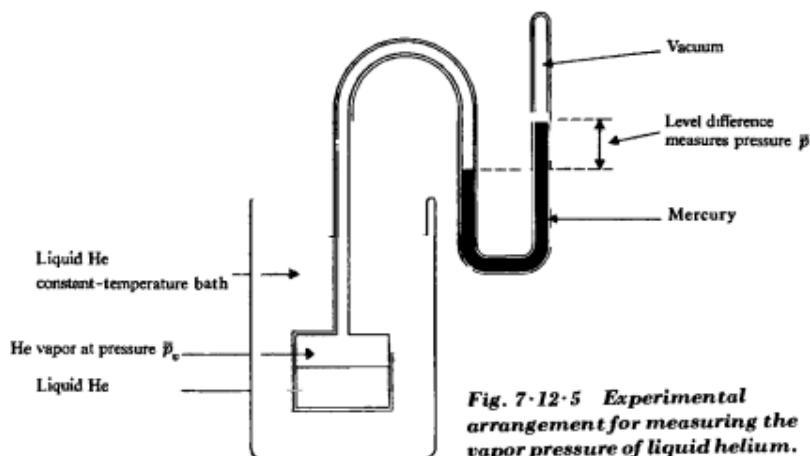


Fig. 7·12·5 Experimental arrangement for measuring the vapor pressure of liquid helium.

corrections" are then necessary to relate \bar{p}_v to the measured pressure \bar{p} . In the limit as $D \ll l$, Eq. (7·12·3) again becomes applicable, so that

$$\frac{\bar{p}_v}{\sqrt{0.5}} = \frac{\bar{p}}{\sqrt{300}}$$

if room temperature is taken to be 300°K. Thus \bar{p} can differ from \bar{p}_v by as much as a factor of $\sqrt{600} \approx 25$, which is a very substantial correction indeed.

7 · 13 Pressure and momentum transfer

It is of interest to consider from a detailed kinetic point of view how a gas exerts a pressure. The basic mechanism is certainly clear: The mean force exerted on a wall of the container is due to the many collisions of molecules with the wall. Let us examine this mechanism in greater detail. We shall again look at the problem in a highly simplified way before doing the exact calculation.

Crude calculation We can give an argument similar to that used at the beginning of Section 7·11. In Fig. 7·11·1 we again imagine that roughly one-third of the molecules move parallel to the z direction. When such a molecule strikes the right end-wall, its kinetic energy remains unchanged. (This must be true, at least, on the average; otherwise one would not have an equilibrium situation.) The magnitude of the momentum of the molecule must then also remain unchanged; i.e., the molecule, approaching the right end-wall with momentum mv in the z direction, must have momentum $-mv$ after it rebounds from the wall. The z component of momentum of the molecule changes then by an amount $\Delta p_z = -2mv$ as a result of the collision with the wall. Correspondingly it follows, by conservation of momentum, that the wall gains in such

a collision an amount of momentum $-\Delta p_z = 2mv$. But the mean force exerted on the wall is, by Newton's laws, just equal to the mean rate of change of momentum of the wall. Hence the mean force on the end-wall can be obtained simply by multiplying [the average momentum $2m\bar{v}$ gained by the wall per collision] by [the mean number of collisions ($\frac{1}{6}n\bar{v}A$) per unit time with the end-wall]. The mean force per unit area, or mean pressure \bar{p} on the wall, is then given by*

$$\bar{p} = \frac{1}{A} (2m\bar{v}) \left(\frac{1}{6} n\bar{v}A \right) = \frac{1}{3} nm\bar{v}^2 \quad (7 \cdot 13 \cdot 1)$$

Exact calculation Suppose that we wish to calculate the mean force \mathbf{F} exerted by the gas on a small element of area dA of the container wall. (See Fig. 7·13·1, where we have chosen the z axis to be normal to the element of area.) Then we must calculate the mean rate of change of momentum of this element of wall, i.e., the mean net momentum delivered to this wall element per unit time by the impinging molecules. If we focus attention on an element of area dA lying inside the gas an infinitesimal distance in front of the wall, then the above calculation is equivalent to finding the mean net molecular momentum which is transported per unit time across this surface from left to right as the molecules cross this surface from both directions.† Let us denote by $\mathbf{G}^{(+)}$ the mean molecular momentum crossing this surface dA per unit time from left to right, and by $\mathbf{G}^{(-)}$ the mean molecular momentum crossing this surface dA per unit time from right to left. Then one has simply

$$\mathbf{F} = \mathbf{G}^{(+)} - \mathbf{G}^{(-)} \quad (7 \cdot 13 \cdot 2)$$

To calculate $\mathbf{G}^{(+)}$, consider the element of surface dA in the gas and focus first attention on those molecules with velocity between \mathbf{v} and $\mathbf{v} + d\mathbf{v}$. (See Fig. 7·13·2, which is similar to Fig. 7·11·2.) The mean number of such molecules which cross this area in an infinitesimal time dt is again the mean number of such molecules contained in the cylinder of volume $|dA \mathbf{v} dt \cos \theta|$; i.e., it is equal to $f(\mathbf{v})d^3\mathbf{v} |dA \mathbf{v} dt \cos \theta|$. By multiplying this number by the momentum

* The symbol \bar{p} stands for mean pressure and should not be confused with the momentum variable p .

† Similarly, and quite equivalently, one can consider an element of area anywhere inside the gas and ask for the mean force which the gas on one side exerts on the gas on the other side. Again this is the same as asking what is the net molecular transport of momentum across this area.

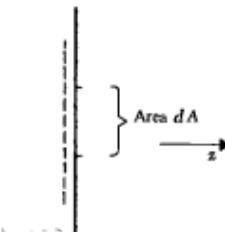


Fig. 7·13·1 An element of area dA of the container wall and a surface of area dA lying inside the gas just in front of the wall.

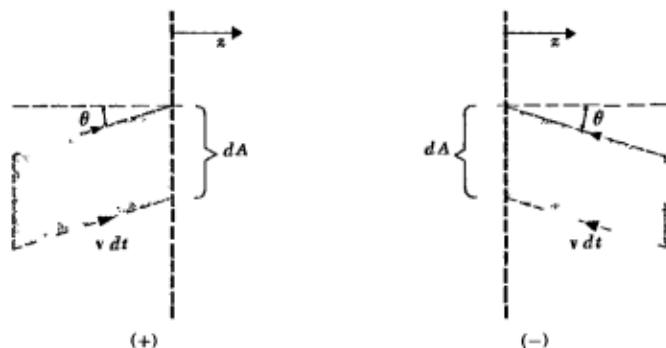


Fig. 7·13·2 Molecules crossing a surface dA in a gas from left to right (+) and from right to left (−). (Note that the height of the cylinders shrinks to zero as $dt \rightarrow 0$.)

mv of each such molecule and dividing by the time dt , one obtains the mean momentum transported across the area dA per unit time by molecules with velocity between v and $v + dv$. By summing over all molecules which cross this area from left to right, i.e., over all molecular velocities with $v_z > 0$, one then gets for the total mean molecular momentum $\mathbf{G}^{(+)}$ transported across this area from left to right the result

$$\mathbf{G}^{(+)} = \int_{v_z > 0} f(\mathbf{v}) d^3\mathbf{v} |dA| v \cos \theta (mv)$$

or
$$\mathbf{G}^{(+)} = dA \int_{v_z > 0} d^3\mathbf{v} f(\mathbf{v}) |v_z|(mv) \quad (7 \cdot 13 \cdot 3)$$

where we have put $v_z = v \cos \theta$ and where the integration is over all velocities for which $v_z > 0$. A similar expression gives the total mean molecular momentum $\mathbf{G}^{(-)}$ transported across this area from right to left, except that the integration must now be over all molecules for which $v_z < 0$. Thus

$$\mathbf{G}^{(-)} = dA \int_{v_z < 0} d^3\mathbf{v} f(\mathbf{v}) |v_z|(mv) \quad (7 \cdot 13 \cdot 4)$$

The force (7·13·2) is then given by the net mean momentum transported across the surface, i.e., by subtracting (7·13·4) from (7·13·3). But in (7·13·3), where the integration is only over positive values of v_z , one can simply put $|v_z| = v_z$ in the integrand. In (7·13·4), where the integration is only over negative values of v_z , one can put $|v_z| = -v_z$ in the integrand. Hence (7·13·3) gives simply

$$\mathbf{F} = \mathbf{G}^{(+)} - \mathbf{G}^{(-)} = dA \int_{v_z > 0} d^3\mathbf{v} f(\mathbf{v}) v_z (mv) + dA \int_{v_z > 0} d^3\mathbf{v} f(\mathbf{v}) v_z (mv)$$

or

$$\mathbf{F} = dA \int d^3\mathbf{v} f(\mathbf{v}) v_z (mv) \quad (7 \cdot 13 \cdot 5)$$

where the two integrals have been combined into a single integral over all possible velocities. Equation (7·13·5) is a very general expression and would

be valid even if the gas is not in equilibrium, i.e., even if f is completely arbitrary.

If the gas is in equilibrium, then $f(\mathbf{v})$ is only a function of $v \equiv |\mathbf{v}|$. Note first that

$$\bar{F}_x = dA \int d^3\mathbf{v} f(\mathbf{v}) v_z v_x = 0 \quad (7 \cdot 13 \cdot 6)$$

since the integrand is odd, having opposite signs for $+v_z$ and $-v_z$. Equation (7 · 13 · 6) expresses the obvious fact that there can be no mean tangential force on the wall in an equilibrium situation. The mean normal force does, of course, not vanish. Measured per unit area it gives the mean pressure, which is thus, by (7 · 13 · 5), equal to

$$\bar{p} = \frac{\bar{F}_z}{dA} = \int d^3\mathbf{v} f(\mathbf{v}) m v_z^2$$

or $\bar{p} = n m \bar{v}^2 \quad (7 \cdot 13 \cdot 7)$

Here we have used the definition

$$\bar{v}^2 \equiv \frac{1}{n} \int d^3\mathbf{v} f(\mathbf{v}) v_z^2$$

By symmetry, $\bar{v}_x^2 = \bar{v}_y^2 = \bar{v}_z^2$ so that

$$\bar{v}^2 = \bar{v}_x^2 + \bar{v}_y^2 + \bar{v}_z^2 = 3\bar{v}_z^2$$

Hence (7 · 13 · 7) can be written equivalently as

► $\bar{p} = \frac{1}{3} n m \bar{v}^2 \quad (7 \cdot 13 \cdot 8)$

This agrees substantially with our crudely derived result (7 · 13 · 1) (except that the averaging is now done carefully so that what appears is \bar{v}^2 rather than \bar{v}^2). Since \bar{v}^2 is related to the mean kinetic energy K of a molecule, (7 · 13 · 8) implies the general relation

$$\bar{p} = \frac{2}{3} n (\frac{1}{2} m \bar{v}^2) = \frac{2}{3} n \bar{K} \quad (7 \cdot 13 \cdot 9)$$

i.e., the mean pressure is just equal to $\frac{2}{3}$ the mean kinetic energy per unit volume of the gas.

Up to now we have not yet made use of the fact that the mean number density of molecules $f(\mathbf{v}) d^3\mathbf{v}$ is given by the Maxwell velocity distribution.* This information allows us to calculate explicitly \bar{v}^2 and is equivalent to using the equipartition theorem result that $\bar{K} = \frac{3}{2} kT$. Then (7 · 13 · 9) becomes

$$\bar{p} = n k T \quad (7 \cdot 13 \cdot 10)$$

so that one regains the equation of state of a classical ideal gas.

* The expressions (7 · 13 · 7) through (7 · 13 · 9) are thus equally valid even if f is given by the quantum-mechanical Fermi-Dirac or Bose-Einstein equilibrium distributions to be discussed in Chapter 9.

SUGGESTIONS FOR SUPPLEMENTARY READING

- J. F. Lee, F. W. Sears, and D. L. Turcotte: "Statistical Thermodynamics," chaps. 3 and 5, Addison-Wesley Publishing Company, Reading, Mass., 1963.
- N. Davidson: "Statistical Mechanics," chaps. 10 and 19, McGraw-Hill Book Company, New York, 1962.
- R. D. Present: "Kinetic Theory of Gases," chaps. 2 and 5, McGraw-Hill Book Company, New York, 1958.

PROBLEMS

- 7.1** Consider a homogeneous mixture of inert monatomic ideal gases at absolute temperature T in a container of volume V . Let there be ν_1 moles of gas 1, ν_2 moles of gas 2, . . . , and ν_k moles of gas k .
- By considering the classical partition function of this system, derive its equation of state, i.e., find an expression for its total mean pressure \bar{p} .
 - How is this total pressure \bar{p} of the gas related to the pressure \bar{p}_i which the i th gas would produce if it alone occupied the entire volume at this temperature?
- 7.2** An ideal monatomic gas of N particles, each of mass m , is in thermal equilibrium at absolute temperature T . The gas is contained in a cubical box of side L , whose top and bottom sides are parallel to the earth's surface. The effect of the earth's uniform gravitational field on the particles should be considered, the acceleration due to gravity being g .
- What is the average kinetic energy of a particle?
 - What is the average potential energy of a particle?
- 7.3** A thermally insulated container is divided by a partition into two compartments, the right compartment having a volume b times as large as the left one. The left compartment contains ν moles of an ideal gas at temperature T and pressure \bar{p} . The right compartment also contains ν moles of an ideal gas at the temperature T . The partition is now removed. Calculate
- the final pressure of the gas mixture in terms of \bar{p} ;
 - the total change of entropy if the gases are different;
 - the total change of entropy if the gases are identical.
- 7.4** A thermally insulated container is divided into two parts by a thermally insulated partition. Both parts contain ideal gases which have equal constant heat capacities c_V . One of these parts contains ν_1 moles of gas at a temperature T_1 and pressure \bar{p}_1 ; the other contains ν_2 moles of gas at a temperature T_2 and pressure \bar{p}_2 . The partition is now removed and the system is allowed to come to equilibrium.
- Find the final pressure.
 - Find the change ΔS of total entropy if the gases are different.
 - Find ΔS if the gases are identical.
- 7.5** A rubber band at absolute temperature T is fastened at one end to a peg, and supports from its other end a weight W . Assume as a simple microscopic model of the rubber band that it consists of a linked polymer chain of N segments joined end to end; each segment has length a and can be oriented either parallel or antiparallel to the vertical direction. Find an expression for the resultant mean length \bar{l} of the rubber band as a function of W . (Neglect the kinetic

energies or weights of the segments themselves, or any interaction between the segments.)

- 7.6 Consider a gas which is *not* ideal so that molecules *do* interact with each other. This gas is in thermal equilibrium at the absolute temperature T . Suppose that the translational degrees of freedom of this gas can be treated classically. What is the mean kinetic energy of (center-of-mass) translation of a molecule in this gas?
- 7.7 Monatomic molecules adsorbed on a surface are free to move on this surface and can be treated as a classical ideal two-dimensional gas. At absolute temperature T , what is the heat capacity per mole of molecules thus adsorbed on a surface of fixed size?
- 7.8 The electrical resistivity ρ of a metal at room temperature is proportional to the probability that an electron is scattered by the vibrating atoms in the lattice, and this probability is in turn proportional to the mean square amplitude of vibration of these atoms. Assuming classical statistics to be valid in this temperature range, what is the dependence of the electrical resistivity ρ on the absolute temperature T ?
- 7.9 A very sensitive spring balance consists of a quartz spring suspended from a fixed support. The spring constant is α , i.e., the restoring force of the spring is $-\alpha x$ if the spring is stretched by an amount x . The balance is at a temperature T in a location where the acceleration due to gravity is g .
- If a very small object of mass M is suspended from the spring, what is the mean resultant elongation \bar{x} of the spring?
 - What is the magnitude $(\bar{x} - \bar{x})^2$ of the thermal fluctuations of the object about its equilibrium position?
 - It becomes impracticable to measure the mass of an object when the fluctuations are so large that $[(\bar{x} - \bar{x})^2]^{\frac{1}{2}} = \bar{x}$. What is the minimum mass M which can be measured with this balance?
- 7.10 A system consists of N very weakly interacting particles at a temperature T sufficiently high so that classical statistical mechanics is applicable. Each particle has mass m and is free to perform one-dimensional oscillations about its equilibrium position. Calculate the heat capacity of this system of particles at this temperature in each of the following cases:
- The force effective in restoring each particle to its equilibrium position is proportional to its displacement x from this position.
 - The restoring force is proportional to x^3 .
- 7.11 Assume the following highly simplified model for calculating the specific heat of graphite, which has a highly anisotropic crystalline layer structure. Each carbon atom in this structure can be regarded as performing simple harmonic oscillations in three dimensions. The restoring forces in directions parallel to a layer are very large; hence the natural frequencies of oscillations in the x and y directions lying within the plane of a layer are both equal to a value $\omega_{||}$ which is so large that $\hbar\omega_{||} \gg 300k$. On the other hand, the restoring force perpendicular to a layer is quite small; hence the frequency of oscillation ω_{\perp} of an atom in the z direction perpendicular to a layer is so small that $\hbar\omega_{\perp} \ll 300k$. On the basis of this model, what is the molar specific heat (at constant volume) of graphite at 300°K?
- 7.12 Consider a solid of compressibility κ . Assume that the atoms in this solid are arranged on a regular cubic lattice, the distance between their nearest neighbors being a . Assume further that a restoring force $-\kappa_0 \Delta a$ acts on a given atom when it is displaced by a distance Δa from its nearest neighbor.

- (a) Use simple reasoning to find an approximate relation between the spring constant κ_0 and the compressibility κ of this solid. (Consider the force needed to decrease the length of one edge of a solid parallelepiped by a small amount.)
- (b) Estimate roughly the order of magnitude of the Einstein temperature Θ_E for copper (atomic weight = 63.5) by assuming that it is a simple cubic structure with density 8.9 g cm^{-3} and compressibility $4.5 \times 10^{-13} \text{ cm}^2 \text{ dyne}^{-1}$.
- 7.13 Show that the general expression (7·8·13) for $\bar{\mu}_z$ becomes identical to the simple expression (6·3·3) in the case where $J = \frac{1}{2}$.
- 7.14 Consider an assembly of N_0 weakly interacting magnetic atoms per unit volume at a temperature T and describe the situation *classically*. Then each magnetic moment μ can make any arbitrary angle θ with respect to a given direction (call it the z direction). In the absence of a magnetic field, the probability that this angle lies between θ and $\theta + d\theta$ is simply proportional to the solid angle $2\pi \sin \theta d\theta$ enclosed in this range. In the presence of a magnetic field H in the z direction, this probability must further be proportional to the Boltzmann factor $e^{-\beta E}$, where E is the magnetic energy of the moment μ making this angle θ with the z axis. Use this result to calculate the classical expression for the mean magnetic moment \bar{M}_z of these N_0 atoms.
- 7.15 Consider the expression (7·8·20) for \bar{M}_z in the limit where the spacing between the magnetic energy levels is small compared to kT , i.e., where $\eta = g\mu J/kT \ll 1$. Assume further that the angle θ between J and the z axis is almost continuous, i.e., that J is so large that the possible values of $\cos \theta = m/J$ are very closely spaced; to be specific assume that J is large enough that $J\eta \gg 1$. Show that in this limit the general expression (7·8·20) for \bar{M}_z does approach the classical expression derived in the preceding problem.
- 7.16 An aqueous solution at room temperature T contains a small concentration of magnetic atoms, each of which has a net spin $\frac{1}{2}$ and a magnetic moment μ . The solution is placed in an external magnetic field H pointing along the z direction. The magnitude of this field is inhomogeneous over the volume of the solution. To be specific, $H = H(z)$ is a monotonic increasing function of z , assuming a value H_1 at the bottom of the solution where $z = z_1$ and a larger value H_2 at the top of the solution where $z = z_2$.
- (a) Let $n_+(z) dz$ denote the mean number of magnetic atoms whose spin points along the z direction and which are located between z and $z + dz$. What is the ratio $n_+(z_2)/n_+(z_1)$?
- (b) Let $n(z) dz$ denote the total mean number of magnetic atoms (of both directions of spin orientation) located between z and $z + dz$. What is the ratio $n(z_2)/n(z_1)$? Is it less than, equal to, or greater than unity?
- (c) Make use of the fact that $\mu H \ll kT$ to simplify the answers to the preceding questions.
- 7.17 What fraction of the molecules of a gas have x components of velocity between $-\tilde{v}$ and $+\tilde{v}$, where \tilde{v} is the most probable speed of the molecules? (Suggestion: consult a table of the error function; see Appendix A·5.)
- 7.18 Use the results of Problem 5.9 to express the velocity of sound in an ideal gas in terms of the most probable speed \tilde{v} of the molecules in the gas and the specific heat ratio $\gamma \equiv c_p/c_v$ of that gas.
- In the case of helium (He) gas, what fraction of the molecules have molecular speeds less than the speed of sound in this gas?
- 7.19 A gas of molecules, each of mass m , is in thermal equilibrium at the absolute temperature T . Denote the velocity of a molecule by v , its three cartesian com-

ponents by v_x , v_y , and v_z , and its speed by v . What are the following mean values:

$$(a) \overline{v_x} \quad (b) \overline{v_x^2} \quad (c) \overline{v^2 v_x}$$

$$(d) \overline{v_x^3 v_y} \quad (e) \overline{(v_x + bv_y)^2} \quad (f) \overline{v_x^2 v_y^2}$$

where b is a constant

(If you need to calculate explicitly any integrals in this problem, you are the kind of person who likes to turn cranks but does not think.)

- 7.20 An ideal monatomic gas is in thermal equilibrium at room temperature T so that the molecular velocity distribution is Maxwellian.

(a) If v denotes the speed of a molecule, calculate $\overline{(1/v)}$. Compare this with $1/\bar{v}$.

(b) Find the mean number of molecules per unit volume whose energy lies in the range between ϵ and $\epsilon + d\epsilon$.

- 7.21 What is the most probable kinetic energy $\tilde{\epsilon}$ of molecules having a Maxwellian velocity distribution? Is it equal to $\frac{1}{2}\tilde{m}\tilde{v}^2$, where \tilde{v} is the most probable speed of the molecules?

- 7.22 A gas of atoms, each of mass m , is maintained at the absolute temperature T inside an enclosure. The atoms emit light which passes (in the x direction) through a window of the enclosure and can then be observed as a spectral line in a spectroscope. A stationary atom would emit light at the sharply defined frequency ν_0 . But, because of the Doppler effect, the frequency of the light observed from an atom having an x component of velocity v_x is not simply equal to the frequency ν_0 , but is approximately given by

$$\nu = \nu_0 \left(1 + \frac{v_x}{c}\right)$$

where c is the velocity of light. As a result, not all of the light arriving at the spectroscope is at the frequency ν_0 ; instead it is characterized by some intensity distribution $I(\nu) d\nu$ specifying the fraction of light intensity lying in the frequency range between ν and $\nu + d\nu$. Calculate

(a) The mean frequency $\bar{\nu}$ of the light observed in the spectroscope.

(b) The root-mean square frequency shift $(\Delta\nu)_{rms} = [\overline{(\nu - \bar{\nu})^2}]^{1/2}$ (measured from the mean frequency) of the light observed in the spectroscope.

(c) The relative intensity distribution $I(\nu) d\nu$ of the light observed in the spectroscope.

- 7.23 In a molecular beam experiment, the source is a tube containing hydrogen at a pressure $\bar{p}_s = 0.15$ mm of mercury and at a temperature $T = 300^\circ\text{K}$. In the tube wall is a slit $20 \text{ mm} \times 0.025 \text{ mm}$, opening into a highly evacuated region. Opposite the source slit and one meter away from it is a second detector slit parallel to the first and of the same size. This slit is in the wall of a small enclosure in which the pressure \bar{p} can be measured.

(a) How many H_2 molecules leave the source slit per second?

(b) How many H_2 molecules arrive at the detector slit per second?

(c) What is the pressure \bar{p}_d in the detector chamber when a steady state has been reached so that \bar{p}_d is independent of time?

- 7.24 A thin-walled vessel of volume V , kept at constant temperature, contains a gas which slowly leaks out through a small hole of area A . The outside pressure is low enough that leakage back into the vessel is negligible. Find the time

- required for the pressure in the vessel to decrease to $1/e$ of its original value. Express your answer in terms of A , V , and the mean molecular speed \bar{v} .
- 7.25 A spherical bulb 10 cm in radius is maintained at room temperature (300°K) except for one square centimeter which is kept at liquid nitrogen temperature (77°K). The bulb contains water vapor originally at a pressure of 0.1 mm of mercury. Assuming that every water molecule striking the cold area condenses and sticks to the surface, estimate the time required for the pressure to decrease to 10^{-6} mm of mercury.
- 7.26 A vessel is closed off by a porous partition through which gases can pass by effusion and then be pumped off to some collecting chamber. The vessel itself is filled with a dilute gas consisting of two types of molecules which differ because they contain two different atomic isotopes and have correspondingly masses m_1 and m_2 . The concentrations of these molecules are c_1 and c_2 , respectively, and are maintained constant inside the vessel by constantly replenishing the supply of gas in it by a steady slow flow of fresh gas through the vessel.
- (a) Let c'_1 and c'_2 denote the concentrations of the two types of molecules in the collecting chamber. What is the ratio c'_2/c'_1 ?
- (b) By using the gas UF₆, one can attempt to separate U²³⁵ from U²³⁸, the first of these isotopes being the one useful in the initiation of nuclear-fission reactions. The molecules in the vessel are then U²³⁸F₆¹⁹ and U²³⁵F₆¹⁹. (The concentrations of these molecules, corresponding to the natural abundance of the two uranium isotopes, are $c_{238} = 99.3$ percent and $c_{235} = 0.7$ percent.) Calculate the corresponding ratio c'_{235}/c'_{238} of the molecules collected after effusion in terms of their original concentration ratio c_{235}/c_{238} .
- 7.27 A container has as one of its walls a membrane containing many small holes. If the container is filled with gas at some moderate pressure p_0 , gas will escape by effusion into the vacuum surrounding the container. It is found that when the container is filled with He gas at room temperature and at pressure p_0 , the pressure will have fallen to $\frac{1}{2}p_0$ after one hour.
- Suppose that the container is filled at room temperature and at total pressure p_0 with a mixture of helium (He) and neon (Ne), the atomic concentrations of both species being 50 percent (i.e., 50 percent of the atoms are He and 50 percent of them are Ne). What will be the ratio $n_{\text{Ne}}/n_{\text{He}}$ of the atomic concentrations of Ne to He after one hour? Express your answer in terms of the atomic weights μ_{Ne} of neon and μ_{He} of helium.
- 7.28 A box of volume V containing an ideal gas of molecular weight μ at temperature T is divided into two equal halves by a partition. Initially the pressure on the left side is $p_1(0)$ and that on the right side is $p_2(0)$. A small hole of area A is now introduced in the partition by opening a valve so that the molecules can effuse through the resulting hole in the (thin) partition.
- (a) Find the pressure $p_1(t)$ of the gas in the left side of the box as a function of time.
- (b) Calculate the change of entropy ΔS of the whole gas after the final equilibrium has been reached.
- 7.29 An enclosure contains a dilute gas at temperature T . Some molecules can escape into a vacuum by effusing through a small hole in one of the walls of the container. Choose the z direction so as to point along the outward normal to the plane of this hole. Let the mass of a molecule be m and the z component of its velocity be denoted by v_z .

(a) What is the mean velocity component \bar{v}_x of a molecule inside the container?

(b) What is the mean velocity component \bar{v}_z of a molecule which has effused into the vacuum?

- 7.30** The molecules of a monatomic ideal gas are escaping by effusion through a small hole in a wall of an enclosure maintained at absolute temperature T .

(a) By physical reasoning (without actual calculation) do you expect the *mean* kinetic energy $\bar{\epsilon}_0$ of a molecule in the effusing beam to be equal to, greater than, or less than the mean kinetic energy $\bar{\epsilon}_i$ of a molecule within the enclosure?

(b) Calculate $\bar{\epsilon}_0$ for a molecule in the effusing beam. Express your answer in terms of $\bar{\epsilon}_i$.

- 7.31** An enclosure contains gas at a pressure \bar{p} and has in one of its walls a small hole of area A through which gas molecules pass into a vacuum by effusion. In this vacuum, directly in front of the hole at a distance L from it, there is suspended a circular disk of radius R . It is oriented so that the normal to its surface points toward the hole (see figure). Assuming that the molecules in the effusing beam get scattered elastically from this disk, calculate the force exerted on the disk by the molecular beam.



Equilibrium between phases or chemical species

8

THE LAST several chapters have elaborated both the macroscopic and the microscopic aspects of the basic theory of Chapter 3. We are thus well prepared to use this theory to discuss a number of important physical situations. Up to now we have dealt almost exclusively with systems consisting of a single "component" (i.e., of a single type of molecule or chemical species) and a single "phase" (i.e., a single spatially homogeneous state of aggregation). But the situations of greatest interest are often more complicated. For example, one may be interested in a single-component system consisting of several phases (e.g., ice and water in equilibrium, or a liquid and its vapor in equilibrium). Alternatively, one may be interested in a single-phase system consisting of several components (e.g., a gas consisting of several types of molecules which may react chemically with each other). Or, in the case of greatest generality, one may be interested in systems consisting of several components in several phases.

In this chapter we shall show how such more complicated systems can be treated by the methods of statistical thermodynamics. Most of our considerations will be independent of any particular microscopic models and will lead to a number of very general results. These yield much valuable insight into many systems of common occurrence. In addition, they are useful both in establishing important relationships between various macroscopic quantities and in furnishing suitable starting points for detailed microscopic calculations.

GENERAL EQUILIBRIUM CONDITIONS

The following sections will amplify the discussion of Sec. 3·1 with the aim of examining equilibrium conditions for systems in various physical situations.

8 · 1 Isolated system

Consider a thermally isolated system A . From our discussion of Sec. 3 · 1, as summarized in the second law of thermodynamics, we know that any spontaneously occurring process is such that the entropy of the system tends to increase. In statistical terms this means that the system tends to approach a situation of larger intrinsic probability. In any such process the spontaneous change of entropy satisfies thus the condition

$$\Delta S \geq 0 \quad (8 \cdot 1 \cdot 1)$$

It follows that if a stable equilibrium situation has been attained where no further spontaneous processes (other than ever-present random fluctuations) can take place, then it is a situation where S is maximum; i.e., it is the most probable situation for the system subject to the given constraints. Hence one can make the following statement.

For a thermally isolated system, the stable equilibrium situation is characterized by the fact that



$$S = \text{maximum} \quad (8 \cdot 1 \cdot 2)$$

This means that if one goes away from the situation where $S = S_{\max}$ is maximum, then for very small departures from this equilibrium situation S does not change ($dS = 0$ for an extremum), but for larger departures S must decrease. That is, the change of entropy $\Delta_m S$ measured from a stable equilibrium situation is such that

$$\Delta_m S \equiv S - S_{\max} \leq 0 \quad (8 \cdot 1 \cdot 3)$$

Example 1 Let us illustrate the situation schematically (see Fig. 8 · 1 · 1). Suppose that the system is characterized by a parameter y (or by several such parameters) which is free to vary. (For example, the system might consist of ice and water, and y might denote the relative concentration of ice.) Then

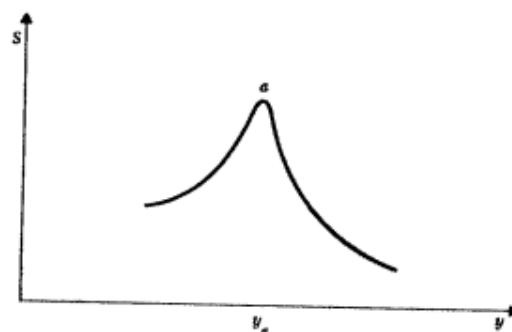


Fig. 8 · 1 · 1 Diagram illustrating the dependence of the entropy S on a parameter y .

the point a on the diagram corresponds to a maximum of S , and the stable equilibrium situation of the system corresponds to an adjustment of the parameter y until it attains the value y_b . Here S attains its absolute maximum so that this corresponds to a situation of absolute stability.

Example 2 A more complicated possible situation is illustrated schematically in Fig. 8·1·2. Here the point a corresponds to a local maximum of S . Thus there is no spontaneous process possible whereby y could move away from the value y_a by relatively small amounts. This value of y corresponds then to a situation of relative equilibrium (or metastable equilibrium). On the other hand, it is possible that by some *major* disturbance the parameter y could attain a value in the region near y_b . Then it will tend to approach the value y_b where the entropy S has its absolute maximum. The point b thus represents a situation of absolutely stable equilibrium.

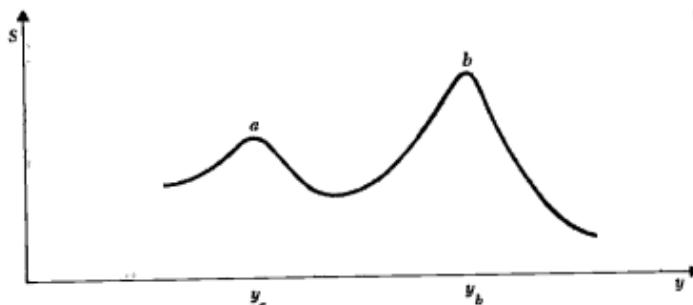


Fig. 8·1·2 Diagram illustrating the dependence of the entropy S on a parameter y when there exists the possibility of metastable equilibrium.

In a thermally isolated system of the kind discussed, the first law of thermodynamics implies that

$$Q = 0 = W + \Delta \bar{E} \quad (8\ 1\cdot 4)$$

or

$$W = -\Delta \bar{E}$$

If the external parameters of the system (e.g., its volume) are kept fixed, then no work gets done and

$$\bar{E} = \text{constant} \quad (8\ 1\cdot 5)$$

while S tends to approach its maximum value in accordance with (8·1·2).

The argument leading to (8·1·2) can be phrased in more explicit statistical terms. Suppose that an isolated system is described by a parameter y (or by several such parameters) so that its total energy is constant. Let $\Omega(y)$ denote the number of states accessible to the system when the parameter has a given value between y and $y + \delta y$ (δy being some fixed small interval); the corresponding entropy of the system is then, by definition, $S(y) = k \ln \Omega(y)$. If the parameter y is free to vary, then the fundamental statistical postulate

asserts that in an equilibrium situation the probability $P(y)$ of finding the system with the parameter between y and $y + \delta y$ is given by

$$\blacktriangleright \quad P(y) \propto \Omega(y) = e^{S(y)/k} \quad (8 \cdot 1 \cdot 6)$$

Equation (8 · 1 · 6) shows explicitly that if y is left free to adjust itself, it will tend to approach a value \bar{y} where $P(y)$ is maximum, i.e., where $S(y)$ is maximum. In equilibrium the relative probability of the occurrence of a fluctuation where $y \neq \bar{y}$ is then given by (8 · 1 · 6) as

$$\frac{P(y)}{P_{\max}} = e^{\Delta_m S/k} \quad (8 \cdot 1 \cdot 7)$$

where $\Delta_m S = S(y) - S_{\max}$.

The relations (8 · 1 · 6) or (8 · 1 · 7) provide more quantitative statements than the assertion (8 · 1 · 2) because they imply not only that the system tends to approach the situation where $S = S_{\max}$, but they also allow one to calculate the probability of occurrence of fluctuations where $S \neq S_{\max}$.

Remark If S depends on a single parameter y , then its maximum occurs for some value $y = \bar{y}$ determined by the condition

$$\frac{\partial S}{\partial y} = 0$$

Expansion of S about its maximum gives then

$$S(y) = S_{\max} + \frac{1}{2} \left(\frac{\partial^2 S}{\partial y^2} \right) (y - \bar{y})^2 +$$

where the second derivative is evaluated at $y = \bar{y}$ and must be negative to guarantee that S is maximum when $y = \bar{y}$. Thus one can write $(\partial^2 S / \partial y^2) = -|\partial^2 S / \partial y^2|$ and obtains by (8 · 1 · 6) the explicit expression

$$P(y) \propto \exp \left[-\frac{1}{2k} \left| \frac{\partial^2 S}{\partial y^2} \right| (y - \bar{y})^2 \right] \quad (8 \cdot 1 \cdot 8)$$

for the probability of fluctuations near the equilibrium situation where $y = \bar{y}$. The fluctuations are thus described by a Gaussian distribution with a dispersion given by

$$\overline{(y - \bar{y})^2} = k \left| \frac{\partial^2 S}{\partial y^2} \right|^{-1}$$

8 · 2 System in contact with a reservoir at constant temperature

Knowing the equilibrium conditions for an isolated system, one can readily deduce similar conditions for other situations of physical interest. For example, much experimental work is done under conditions of constant temperature. Thus we should like to investigate the equilibrium conditions for a

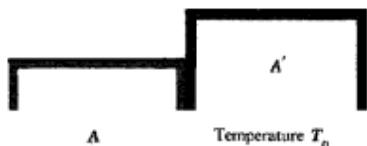


Fig. 8·2·1 A system \$A\$ in contact with a heat reservoir at temperature \$T_0\$.

system \$A\$ in thermal contact with a heat reservoir \$A'\$ which is at the constant absolute temperature \$T_0\$.

The combined system \$A^{(0)}\$ consisting of the system \$A\$ and the heat reservoir \$A'\$ is an isolated system of the type discussed in Sec. 8·1. The entropy \$S^{(0)}\$ of \$A^{(0)}\$ then satisfies the condition (8·1·1) that in any spontaneous process

$$\Delta S^{(0)} \geq 0 \quad (8\cdot2\cdot1)$$

But this condition can readily be written in terms of quantities which refer only to the system \$A\$ of interest. Thus

$$\Delta S^{(0)} = \Delta S + \Delta S' \quad (8\cdot2\cdot2)$$

where \$\Delta S\$ is the entropy change of \$A\$ and \$\Delta S'\$ that of the reservoir \$A'\$ in the process under consideration. But if \$A\$ absorbs heat \$Q\$ from the reservoir \$A'\$ in this process, then \$A'\$ absorbs heat \$(-Q)\$ and suffers a corresponding entropy change

$$\Delta S' = \frac{(-Q)}{T_0}$$

since it remains in internal equilibrium at the constant temperature \$T_0\$. Furthermore, the first law gives

$$Q = \Delta \bar{E} + W$$

where \$\Delta \bar{E}\$ is the internal energy change of \$A\$ and \$W\$ the work done by \$A\$ in the process under consideration. Thus (8·2·2) can be written

$$\Delta S^{(0)} = \Delta S - \frac{Q}{T_0} = \frac{T_0 \Delta S - (\Delta \bar{E} + W)}{T_0} = \frac{\Delta(T_0 S - \bar{E}) - W}{T_0}$$

or $\Delta S^{(0)} = \frac{-\Delta F_0 - W}{T_0} \quad (8\cdot2\cdot3)$

where we have used the fact that \$T_0\$ is a constant and have introduced the definition

$$F_0 \equiv \bar{E} - T_0 S \quad (8\cdot2\cdot4)$$

This reduces to the ordinary Helmholtz free energy \$F = \bar{E} - TS\$ of the system \$A\$ if the latter has a temperature \$T\$ equal to that of the heat reservoir \$A'\$. Of

course, in the general case when A is not in equilibrium with A' , its temperature T is not necessarily equal to T_0 .

The total entropy change $\Delta S^{(0)}$ in (8·2·3) is expressed completely in terms of quantities that refer only to the system A of interest. The fundamental condition (8·2·1) then allows us to draw some interesting conclusions. Since T_0 is in all ordinary cases positive, one gets

$$-\Delta F_0 \geq W \quad (8\cdot2\cdot5)$$

This relation implies that the *maximum* work which can be done by a system in contact with a heat reservoir is given by $(-\Delta F_0)$. (This is the reason for the name "free energy" given to F .) The maximum work corresponds, of course, to the equals sign in (8·2·1) and is obtained when the process used is a quasi-static one (so that A is always in equilibrium with A' and $T = T_0$). Equation (8·2·5) should be compared with the rather different relation (8·1·4) which holds for the work done by an *isolated* system.

If the external parameters of the system A (e.g., its volume) are kept fixed, then $W = 0$ and (8·2·5) yields the condition

$$\Delta F_0 \leq 0 \quad (8\cdot2\cdot6)$$

This equation is analogous to Eq. (8·1·1) for an *isolated* system. It implies that if a system in thermal contact with a heat reservoir, its free energy tends to *decrease*. Thus we arrive at the statement:

If a system, whose external parameters are fixed, is in thermal contact with a heat reservoir, the stable equilibrium situation is characterized by the condition that

$$F_0 = \text{minimum} \quad (8\cdot2\cdot7)$$

This last condition can again be phrased in more explicit statistical terms. Consider the external parameters of A to be fixed so that $W = 0$, and suppose that A is described by some parameter y . The thermodynamic functions of A (e.g., S and \bar{E}) have definite values $S(y_1)$ and $\bar{E}(y_1)$ when y has a given value $y = y_1$. If the parameter changes to any other value y , these functions change by corresponding amounts $\Delta S = S(y) - S(y_1)$ and $\Delta \bar{E} = \bar{E}(y) - \bar{E}(y_1) = Q$. The entropy of the heat reservoir A' also changes since it absorbs some heat, and the corresponding change in the total entropy of $A^{(0)}$ is given by (8·2·3) (with $W = 0$) as

$$\Delta S^{(0)} = - \frac{\Delta F_0}{T_0} \quad (8\cdot2\cdot8)$$

But in an equilibrium situation the probability $P(y)$ that the parameter lies between y and $y + \delta y$ is proportional to the number of states $\Omega^{(0)}(y)$ accessible to the total isolated system $A^{(0)}$ when the parameter lies in this range. Thus one has, analogously to (8·1·6),

$$P(y) \propto \Omega^{(0)}(y) = e^{S^{(0)}(y)/k} \quad (8\cdot2\cdot9)$$

But by (8·2·8)

$$S^{(0)}(y) = S^{(0)}(y_1) - \frac{\Delta F_0}{T_0} = S^{(0)}(y_1) - \frac{F_0(y) - F_0(y_1)}{T_0}$$

Since y_1 is just some arbitrary constant value, the corresponding constant terms can be absorbed in the constant of proportionality of (8·2·9) which then becomes

► $P(y) \propto e^{-F_0(y)/kT_0}$ (8·2·10)

This relation shows directly that the most probable situation is one where F_0 is a minimum. Of course, if \bar{E} is a constant independent of y (as it would be for an isolated system), then $F_0 = \bar{E} - T_0 S(y)$ and (8·1·10) reduces properly to the relation (8·1·6) which involves only the entropy.

Remark The relation (8·2·10) is, of course, akin to the canonical distribution. According to the latter the desired probability is given by

$$P(y) \propto \sum_r e^{-\beta_0 E_r}, \quad \beta_0 = (kT_0)^{-1} \quad (8·2·11)$$

where the sum extends over all states r for which the parameter lies between y and $y + \delta y$. If $\Omega(E; y)$ denotes the number of states for which the parameter lies between y and $y + \delta y$ and whose energy lies between E and $E + \delta E$, the relation (8·2·11) becomes

$$P(y) \propto \sum_E \Omega(E; y) e^{-\beta_0 E}$$

where the sum is now to be taken over all possible energy intervals. The summand has as a function of E the usual very sharp maximum near some value $\bar{E}(y)$ which depends on y and which is equal to the mean energy of the system for this value of y . Hence only terms near this maximum contribute appreciably to the sum and

$$P(y) \propto \Omega(\bar{E}; y) e^{-\beta_0 \bar{E}(y)} = e^{S(y)/k - \beta_0 \bar{E}(y)} = e^{-\beta_1 F_1(y)}$$

8 · 3 System in contact with a reservoir at constant temperature and pressure

Another case of physical interest is that where a system A is maintained under conditions of both constant temperature and constant pressure. This is a situation of frequent occurrence in the laboratory, where one may carry out an experiment in a thermostat at, say, atmospheric pressure. A situation such as this implies that the system A is in thermal contact with a heat reservoir A' which is at a constant temperature T_0 and at a constant pressure p_0 . The system A can exchange heat with the reservoir A' ; but the latter is so large that its temperature T_0 remains unchanged. Similarly, the system A can

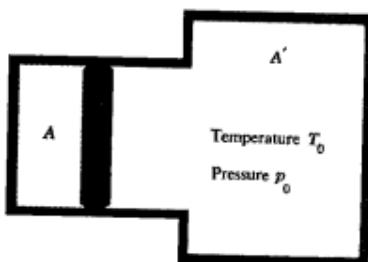


Fig. 8·3·1 A system A in contact with a reservoir at constant temperature T_0 and constant pressure p_0 .

change its volume V at the expense of the reservoir A' , doing work on the reservoir in the process; but again A' is so large that its pressure p_0 remains unaffected by this relatively small volume change.

Remark The system A' may be a single reservoir with which A can interact both by heat transfer and pressure work. Alternatively, A' may be a combination of two reservoirs, one at temperature T_0 with which A can interact only by heat transfer, and the other at pressure p_0 with which A can interact only by pressure work.

The analysis of the equilibrium conditions for a system A under these conditions is very similar to that of the last section. Once again the entropy $S^{(0)}$ of the combined system $A^{(0)} = A + A'$ satisfies the condition that in any spontaneous process

$$\Delta S^{(0)} = \Delta S + \Delta S' \geq 0 \quad (8\cdot3\cdot1)$$

If A absorbs heat Q from A' in this process, then $\Delta S' = -Q/T_0$. But now the first law applied to A gives

$$Q = \Delta \bar{E} + p_0 \Delta V + W^*$$

where $p_0 \Delta V$ is the work done by A against the constant pressure p_0 of the reservoir A' and where W^* denotes any *other* work done by A in the process. (For example, W^* might refer to electric or magnetic work done by A .) Hence one can write

$$\begin{aligned} \Delta S^{(0)} &= \Delta S - \frac{Q}{T_0} = \frac{1}{T_0} [T_0 \Delta S - Q] = \frac{1}{T_0} [T_0 \Delta S - (\Delta \bar{E} + p_0 \Delta V + W^*)] \\ &= \frac{1}{T_0} [\Delta(T_0 S - \bar{E} - p_0 V) - W^*] \end{aligned}$$

or

$$\Delta S^{(0)} = \frac{-\Delta G_0 - W^*}{T_0} \quad (8\cdot3\cdot2)$$

Here we have used the fact that T_0 and p_0 are both constant and have introduced the definition

$$G_0 = \bar{E} - T_0 S + p_0 V \quad (8\cdot3\cdot3)$$

This reduces to the ordinary Gibbs free energy $G = \bar{E} - TS + PV$ for the

system A when the temperature and pressure of the latter are equal to those of the reservoir A' .

The total entropy change $\Delta S^{(0)}$ in (8·3·2) is again expressed completely in terms of quantities which refer only to the system A . The fundamental condition (8·3·1) implies then that

$$-\Delta G_0 \geq W^* \quad (8\cdot3\cdot4)$$

This asserts that the *maximum* work (other than work done on the pressure reservoir) which can be done by the system is given by $(-\Delta G_0)$. (This is the reason that G is also called a “free energy.”) The maximum work corresponds again to the equals sign in (8·3·1) and corresponds to a quasi-static process.

If all the external parameters of A , *except* its volume, are kept fixed, then $W^* = 0$ and (8·3·4) yields the condition

$$\Delta G_0 \leq 0 \quad (8\cdot3\cdot5)$$

Hence one concludes that

If a system is in contact with a reservoir at constant temperature and pressure and if its external parameters are fixed so that it can only do work on the pressure reservoir, then the stable equilibrium situation is characterized by the condition that

► $G_0 = \text{minimum}$ (8·3·6)

This last condition can again be phrased in more explicit statistical terms. The probability that a parameter y of the system $A^{(0)}$ assumes a value between y and $y + \delta y$ is given by

$$P(y) \propto e^{S^{(0)}(y)/k} \quad (8\cdot3\cdot7)$$

But with $W^* = 0$, the change in $S^{(0)}$ due to a parameter change away from some standard value y_1 is by (8·3·2)

$$\Delta S^{(0)} = -\frac{\Delta G_0}{T_0} \quad (8\cdot3\cdot8)$$

so that

$$S^{(0)}(y) = S^{(0)}(y_1) - \frac{G_0(y) - G_0(y_1)}{T_0}$$

Since y_1 is just some constant, (8·3·7) leads to the proportionality

► $P(y) \propto e^{-G_0(y)/kT_0}$ (8·3·9)

This again shows explicitly that the most probable situation is one where G_0 is a minimum and allows calculation of the probability of fluctuations about this equilibrium.

8·4 Stability conditions for a homogeneous substance

As a simple example of the preceding discussion, consider a one-component system in a single phase (e.g., a simple liquid or solid). Focus attention on

Fig. 8·4·1 A small portion *A* of a homogeneous substance is singled out for consideration to examine the conditions for stable equilibrium.



some small, but macroscopic, part *A* of this system where *A* consists of some fixed number of particles. The rest of the system is then relatively very large and acts like a reservoir at some *constant* temperature T_0 and *constant* pressure p_0 . By (8·3·5) the condition for stable equilibrium applied to *A* is that for this system the function

$$G_0 \equiv \bar{E} - T_0 S + p_0 V = \text{minimum} \quad (8\cdot4\cdot1)$$

Stability against temperature variations Let T and V be the two independent parameters specifying the macrostate of *A*. Consider first a situation where V is considered fixed but where the temperature parameter T is allowed to vary. Suppose that the minimum of G_0 occurs for $T = \tilde{T}$ when $G_0 = G_{\min}$. Expanding G about its minimum and writing $\Delta T \equiv T - \tilde{T}$, one obtains

$$\Delta_m G_0 = G_0 - G_{\min} = \left(\frac{\partial G_0}{\partial T} \right)_V \Delta T + \frac{1}{2} \left(\frac{\partial^2 G_0}{\partial T^2} \right)_V (\Delta T)^2 + \quad (8\cdot4\cdot2)$$

Here all derivatives are evaluated at $T = \tilde{T}$. Since G is a minimum, its stationary character at this point implies that $\Delta_m G = 0$ in a first approximation; i.e., the first-order terms proportional to ΔT must vanish so that

$$\left(\frac{\partial G_0}{\partial T} \right)_V = 0 \quad \text{for } T = \tilde{T} \quad (8\cdot4\cdot3)$$

The fact that G_0 is not only stationary but a *minimum* at $T = \tilde{T}$ requires that in next approximation, when terms in $(\Delta T)^2$ become important, one has

$$\begin{aligned} \Delta_m G_0 &\geq 0 \\ \text{or} \quad \left(\frac{\partial^2 G_0}{\partial T^2} \right)_V &\geq 0 \quad \text{for } T = \tilde{T} \end{aligned} \quad (8\cdot4\cdot4)$$

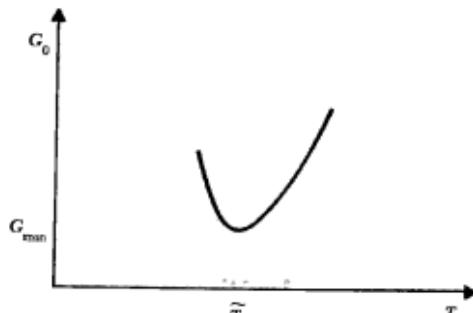


Fig. 8·4·2 Schematic dependence of $G_0(T, V)$ on the temperature T at a fixed volume V .

Using (8·4·1), the condition (8·4·3) that G_0 be stationary becomes, when V is kept constant,

$$\left(\frac{\partial G_0}{\partial T}\right)_V = \left(\frac{\partial \bar{E}}{\partial T}\right)_V - T_0 \left(\frac{\partial S}{\partial T}\right)_V = 0 \quad (8\cdot4\cdot5)$$

But by the fundamental thermodynamic relation

$$T dS = d\bar{E} + \bar{p} dV \quad (8\cdot4\cdot6)$$

it follows that for V constant, $dV = 0$, and

$$T \left(\frac{\partial S}{\partial T}\right)_V = \left(\frac{\partial \bar{E}}{\partial T}\right)_V$$

Thus (8·4·5) becomes

$$\left(\frac{\partial G}{\partial T}\right)_V = \left(1 - \frac{T_0}{T}\right) \left(\frac{\partial \bar{E}}{\partial T}\right)_V \quad (8\cdot4\cdot7)$$

Setting this equal to zero where $T = \tilde{T}$, one gets simply

$$\tilde{T} = T_0 \quad (8\cdot4\cdot8)$$

Hence we arrive at the obvious conclusion that a necessary condition for equilibrium is that the temperature of the subsystem A is the same as that of the surrounding medium.

Now we proceed to the second-order terms to satisfy the condition (8·4·4), which guarantees that G_0 is actually a minimum. By (8·4·7),

$$\left(\frac{\partial^2 G_0}{\partial T^2}\right)_V = \frac{T_0}{T^2} \left(\frac{\partial \bar{E}}{\partial T}\right)_V + \left(1 - \frac{T_0}{T}\right) \left(\frac{\partial^2 \bar{E}}{\partial T^2}\right)_V \geq 0$$

When this is evaluated at the minimum of G , where $T = T_0$ by virtue of (8·4·8), the second term vanishes and one obtains simply

$$\left(\frac{\partial \bar{E}}{\partial T}\right)_V \geq 0 \quad (8\cdot4\cdot9)$$

But this derivative is just the heat capacity C_V at constant volume. Thus

$$C_V \geq 0 \quad (8\cdot4\cdot10)$$

The condition (8·4·9) or (8·4·10) was already derived previously in (3·7·16) and in (6·5·8). It is a fundamental condition required to guarantee the intrinsic stability of any phase.

This condition is physically very reasonable. Indeed, the following statement, known as "Le Châtelier's principle," must be true quite generally:

- If a system is in *stable* equilibrium, then any spontaneous change of its parameters must bring about processes which tend to restore the system to equilibrium.

If this statement were not true, any slight fluctuation leading to a deviation from equilibrium would result in an increase of this deviation so that the system would clearly be unstable. To illustrate the principle in the present

example, suppose that the temperature T of the subsystem A has increased above that of the surroundings A' as a result of a spontaneous fluctuation. Then the process brought into play is heat transfer from the system A at the higher temperature to the surroundings A' , and a resulting *decrease* in the energy \bar{E} of A (i.e., $\Delta\bar{E} < 0$). But the stability condition expressed by Le Châtelier's principle requires that this process, induced by the original temperature increase, is such that the temperature is again *decreased* (i.e., $\Delta T < 0$). Hence it follows that $\Delta\bar{E}$ and ΔT must have the same sign, i.e., that $\partial\bar{E}/\partial T > 0$ in agreement with (8·4·9).

Stability against volume fluctuations Suppose that the temperature of the subsystem A is considered fixed at $T = T_0$, but that its volume V is allowed to vary. Then one can write

$$\Delta_m G_0 \equiv G_0 - G_{m\text{in}} = \left(\frac{\partial G_0}{\partial V}\right)_T \Delta V + \frac{1}{2} \left(\frac{\partial^2 G_0}{\partial V^2}\right)_T (\Delta V)^2 + \dots \quad (8\cdot4\cdot11)$$

where $\Delta V \equiv V - \tilde{V}$, and the expansion is about the volume $V = \tilde{V}$ where G_0 is minimum. The condition that G_0 is stationary demands that

$$\left(\frac{\partial G_0}{\partial V}\right)_T = 0 \quad (8\cdot4\cdot12)$$

Using the definition (8·4·1)

$$\left(\frac{\partial G_0}{\partial V}\right)_T = \left(\frac{\partial \bar{E}}{\partial V}\right)_T - T_0 \left(\frac{\partial S}{\partial V}\right)_T + p_0$$

But, by virtue of (8·4·6),

$$T \left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial \bar{E}}{\partial V}\right)_T + \bar{p}$$

Hence $\left(\frac{\partial G_0}{\partial V}\right)_T = T \left(\frac{\partial S}{\partial V}\right)_T - \bar{p} - T_0 \left(\frac{\partial S}{\partial V}\right)_T + p_0$

or $\left(\frac{\partial G_0}{\partial V}\right)_T = -\bar{p} + p_0 \quad (8\cdot4\cdot13)$

since $T = T_0$. The condition (8·4·12) then implies that at equilibrium, where G is minimum and $V = \tilde{V}$, the corresponding value of the pressure \bar{p} is such that

$$\bar{p} = p_0 \quad (8\cdot4\cdot14)$$

Again this is a rather obvious result which asserts merely that in equilibrium the pressure of the subsystem A must be equal to that of the surrounding medium.

The condition that G is actually a minimum is that $\Delta_m G_0 \geq 0$; or by (8·4·11), that the second derivative of G_0 is positive. By (8·4·13) this yields the condition

$$\left(\frac{\partial^2 G_0}{\partial V^2}\right)_T = -\left(\frac{\partial \bar{p}}{\partial V}\right)_T \geq 0 \quad (8\cdot4\cdot15)$$

In terms of the isothermal compressibility defined by

$$\kappa = -\frac{1}{V} \left(\frac{\partial V}{\partial \bar{p}} \right)_T \quad (8 \cdot 4 \cdot 16)$$

the condition (8·4·15) is equivalent to

$$\kappa \geq 0 \quad (8 \cdot 4 \cdot 17)$$

The stability condition (8·4·15) is again a physically quite reasonable result consistent with Le Châtelier's principle. Suppose that the volume of the subsystem A has *increased* by an amount ΔV as a result of a fluctuation. The pressure \bar{p} of A must then *decrease* below that of its surroundings (i.e., $\Delta \bar{p} < 0$) to guarantee that the net force exerted on A by its surroundings is of a direction such that it tends to reduce its volume to its former value.

Density fluctuations The preceding considerations permit one also to calculate the fluctuation in the volume V of the small subsystem A . The most probable situation is that where $V = \tilde{V}$ is such that G_0 is a minimum, $G_0(\tilde{V}) = G_{\min}$. Let $\mathcal{P}(V) dV$ denote the probability that the volume of A lies between V and $V + dV$. Then one has by (8·3·9)

$$\mathcal{P}(V) dV \propto e^{-G_0(V)/kT} dV \quad (8 \cdot 4 \cdot 18)$$

But when $\Delta V \equiv V - \tilde{V}$ is small, the expansion (8·4·11) is applicable. By virtue of (8·4·12) and (8·4·15) it yields

$$G_0(V) = G_{\min} - \frac{1}{2} \left(\frac{\partial \bar{p}}{\partial V} \right)_T (\Delta V)^2 = G_{\min} + \frac{(\Delta V)^2}{2 \tilde{V}_\kappa}$$

where we have used the definition (8·4·16) in the last step. Thus (8·4·18) becomes

$$\mathcal{P}(V) dV = B \exp \left[-\frac{(V - \tilde{V})^2}{2kT_0 \tilde{V}_\kappa} \right] dV \quad (8 \cdot 4 \cdot 19)$$

where we have absorbed G_{\min} into the proportionality constant B . This constant can, of course, be determined by the normalization requirement that the integral of (8·4·19) over all possible values of the volume V is equal to unity.*

The probability (8·4·19) is simply a Gaussian distribution with a maximum at the volume $V = \tilde{V}$. Thus \tilde{V} is also equal to the mean volume \bar{V} and the general result (1·6·9) implies that (8·4·19) yields a dispersion of the volume equal to

$$\overline{(\Delta V)^2} \equiv \overline{(V - \tilde{V})^2} = kT_0 \tilde{V}_\kappa \quad (8 \cdot 4 \cdot 20)$$

The presence of such volume fluctuations in a small amount of material containing a fixed number N of molecules implies, of course, corresponding fluctuations in the number $n = N/V$ of molecules per unit volume (and thus

* This integral can be extended from $V = -\infty$ to $V = +\infty$, since the integrand (8·4·19) becomes negligible when V differs appreciably from the value \tilde{V} where \mathcal{P} is maximum.

also in the mass density of the substance). The fluctuations in n are centered about the value $\bar{n} = N/\tilde{V}$, and for relatively small values of $\Delta n = n - \bar{n}$ one has $\Delta n = -(N/\tilde{V}^2) \Delta V = -(\bar{n}/\tilde{V}) \Delta V$. Hence (8.4.20) implies for the dispersion in the number density n the result

$$\overline{(\Delta n)^2} = \left(\frac{\bar{n}}{\tilde{V}} \right)^2 \overline{(\Delta V)^2} = \bar{n}^2 \left(\frac{kT_0}{\tilde{V}} \kappa \right) \quad (8.4.21)$$

Note that this depends on the size of the volume \tilde{V} under consideration.

An interesting case arises when

$$\left(\frac{\partial \bar{p}}{\partial V} \right)_T \rightarrow 0 \quad (8.4.22)$$

Then $\kappa \rightarrow \infty$ and the density fluctuations become very large.* The conditions of temperature and pressure which are such that $(\partial \bar{p} / \partial V)_T = 0$ define the so-called "critical point" of the substance. The very large density fluctuations at this point lead to a very large scattering of light. As a result a substance, which is ordinarily transparent, will assume a milky white appearance at its critical point (e.g., liquid CO₂ when it approaches its critical point at a temperature of 304°K and pressure of 73 atmospheres). This impressive phenomenon is known as "critical point opalescence."

Remark The result (8.4.19), proved under conditions of constant temperature, can be shown to remain valid even if both V and T are allowed to vary simultaneously (see Problem 8.1). Hence this discussion of density fluctuations is applicable to the actual case of experimental interest.

EQUILIBRIUM BETWEEN PHASES

8.5 Equilibrium conditions and the Clausius-Clapeyron equation

Consider a single component system which consists of two phases which we shall denote by 1 and 2. For example, these might be solid and liquid, or liquid and gas. We suppose that the system is in equilibrium with a reservoir at the constant temperature T and constant pressure p so that the system itself has always a temperature T and a mean pressure p . But the system can exist in either of its two possible phases or in a mixture of the two. Let us begin by finding the conditions which must be satisfied so that the two phases can coexist in equilibrium with each other.

In accordance with the discussion of Sec. 8.3, the equilibrium condition is that the Gibbs free energy G of the system is a minimum,

$$G = E - TS + pV = \text{minimum} \quad (8.5.1)$$

* They do not become infinite, since the approximations which allowed us to neglect in (8.4.11) terms beyond $(\Delta V)^2$ are then no longer justified.

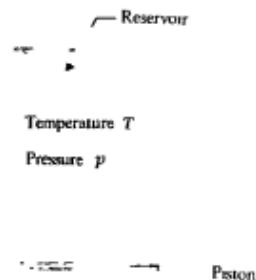


Fig. 8·5·1 A system consisting of two phases maintained at constant temperature and pressure.

Let

- ν_i = the number of moles of phase i present in the system
- $g_i(T, p)$ = the Gibbs free energy per mole of phase i at this temperature T and pressure p

Then G can be written

$$G = \nu_1 g_1 + \nu_2 g_2 \quad (8\cdot5\cdot2)$$

The conservation of matter implies that the total number ν of moles of the substance remain constant, i.e.,

$$\nu_1 + \nu_2 = \nu = \text{constant} \quad (8\cdot5\cdot3)$$

Thus one can take ν_1 to be the one independent parameter which is free to vary. In equilibrium (8·5·1) requires that G be stationary for changes in ν_1 ; thus

$$dG = g_1 d\nu_1 + g_2 d\nu_2 = 0$$

$$\text{or} \quad (g_1 - g_2) d\nu_1 = 0$$

since $d\nu_2 = -d\nu_1$ by (8·5·3). Hence we obtain as a necessary condition for equilibrium that

$$\blacktriangleright \quad g_1 = g_2 \quad (8\cdot5\cdot4)$$

Clearly, when this condition is satisfied, the transfer of a mole of substance from one phase to the other does not change G in (8·5·2); hence G is then stationary as required.

Remark One could go further and investigate the conditions that G is actually a *minimum*, but this would not yield anything very interesting other than the conditions that the heat capacity and compressibility of each phase must be positive to guarantee the stability of each phase (as shown in Sec. 8·4).

Let us look at the situation more closely. For given T and p , $g_1(T,p)$ is a well-defined function characteristic of the properties of phase 1; similarly, $g_2(T,p)$ is a well-defined function characteristic of phase 2.

If T and p are such that $g_1 < g_2$, then the minimum value of G in (8.5.2) is achieved if all the ν moles of substance transform into phase 1 so that $G = \nu g_1$. Phase 1 is then the stable one.

If T and p are such that $g_1 > g_2$, then the minimum value of G is achieved if all the substance transforms into phase 2 so that $G = \nu g_2$. Phase 2 is then the stable one.

If T and p are such that $g_1 = g_2$, then the condition (8.5.4) is satisfied and *any* amount ν_1 of phase 1 can coexist in equilibrium with the remaining amount $\nu_2 = \nu - \nu_1$ of phase 2. The value G remains unchanged when ν_1 is varied. The locus of points where T and p are such that the condition (8.5.4) is fulfilled then represents the "phase-equilibrium line" along which the two phases can coexist in equilibrium. This line, along which $g_1 = g_2$, divides the pT plane into two regions: one where $g_1 < g_2$, so that phase 1 is the stable one, and the other where $g_1 > g_2$, so that phase 2 is the stable one.

It is possible to characterize the phase-equilibrium line by a differential equation. In Fig. 8.5.2 consider any point, such as A , which lies *on* the phase-equilibrium line and corresponds to temperature T and pressure p . Then the condition (8.5.4) implies that

$$g_1(T,p) = g_2(T,p) \quad (8.5.5)$$

Consider now a neighboring point, such as B , which also lies on the phase-equilibrium line and corresponds to temperature $T + dT$ and pressure $p + dp$. Then the condition (8.5.4) implies that

$$g_1(T + dT, p + dp) = g_2(T + dT, p + dp) \quad (8.5.6)$$

Subtracting (8.5.5) from (8.5.6) yields the condition

$$dg_1 = dg_2 \quad (8.5.7)$$

where

$$dg_i = \left(\frac{\partial g_i}{\partial T} \right)_p dT + \left(\frac{\partial g_i}{\partial p} \right)_T dp$$

p

Phase 1

$g_1 = g_2$

$g_1 < g_2$

A

Phase 2

T

Fig. 8.5.2 Pressure-temperature plot showing the domains of relative stability of two phases and the phase-equilibrium line.

is the change in the molar Gibbs free energy for phase i in going from point A to point B .

But the change $d\epsilon$ for each phase can also be obtained by using the fundamental thermodynamic relation

$$d\epsilon = T \, ds - p \, dv$$

expressing the change in mean molar energy ϵ of this phase. Thus

$$dg \equiv d(\epsilon - Ts + pv) = -s \, dT + v \, dp \quad (8 \cdot 5 \cdot 8)$$

Hence (8·5·7) implies that

$$\begin{aligned} -s_1 \, dT + v_1 \, dp &= -s_2 \, dT + v_2 \, dp \\ (s_2 - s_1) \, dT &= (v_2 - v_1) \, dp \\ \frac{dp}{dT} &= \frac{\Delta s}{\Delta v} \end{aligned} \quad (8 \cdot 5 \cdot 9)$$

where $\Delta s \equiv s_2 - s_1$ and $\Delta v \equiv v_2 - v_1$. This is called the “Clausius-Clapeyron equation.” Consider any point on the phase-equilibrium line at a temperature T and corresponding pressure p . Equation (8·5·9) then relates the slope of the phase-equilibrium line at this point to the entropy change Δs and volume change Δv of the substance in “crossing the line” at this point, i.e., in undergoing a change of phase at this temperature and pressure. (Note that the quantities on the right side of (8·5·9) do not need to be referred to one mole of the substance; both numerator and denominator can be multiplied by the same number of moles, and dp/dT must obviously be left unchanged.)

Since there is an entropy change associated with the phase transformation, heat must also be absorbed. The “latent heat of transformation” L_{12} is defined as the heat absorbed when a given amount of phase 1 is transformed to phase 2. Since the process takes place at the constant temperature T , the corresponding entropy change is simply

$$\Delta S = S_2 - S_1 = \frac{L_{12}}{T} \quad (8 \cdot 5 \cdot 10)$$

where L_{12} is the latent heat at this temperature. Thus the Clausius-Clapeyron equation (8·5·9) can be written

$$\blacktriangleright \quad \frac{dp}{dT} = \frac{\Delta S}{\Delta V} = \frac{L_{12}}{T \Delta V} \quad (8 \cdot 5 \cdot 11)$$

Clearly, if V refers to the molar volume, then L_{12} is the latent heat per mole; if V refers to the volume per gram, then L_{12} is the latent heat per gram.

Let us discuss a few important illustrations.

Phase transformations of a simple substance Simple substances are capable of existing in phases of three types: solid, liquid, and gas. (There may also be several solid phases with different crystal structures.) The phase-equilibrium lines separating these phases appear typically as shown in Fig. 8·5·3. These lines separate solid from liquid, liquid from gas, and solid from

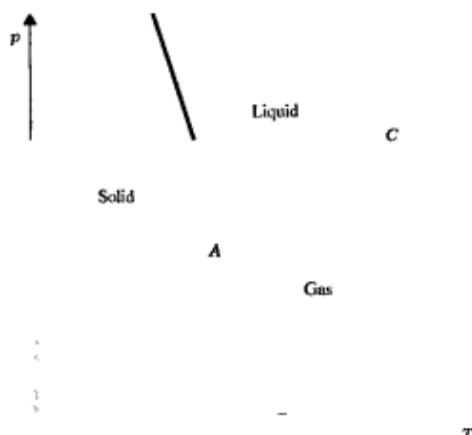


Fig. 8·5·3 Phase diagram for a simple substance. Point A is the triple point, point C the critical point.

gas.* The three lines meet at one common point *A*, called the “triple point”; at this unique temperature and pressure arbitrary amounts of *all three* phases can therefore coexist in equilibrium with each other. (This is the property which makes the triple point of water so suitable as a readily reproducible temperature standard.) At point *C*, the so-called “critical point,” the liquid-gas equilibrium line ends. The volume change ΔV between liquid and gas has then approached zero; beyond *C* there is no further phase transformation, since there exists only one “fluid” phase (the very dense gas has become indistinguishable from the liquid).

In going from solid (*s*) to liquid (*l*) the entropy of the substance (or degree of disorder) almost always increases.† Thus the corresponding latent heat L_{st} is positive and heat gets absorbed in the transformation. In most cases the solid expands upon melting, so that $\Delta V > 0$. In that case the Clausius-Clapeyron equation (8·5·11) asserts that the slope of the solid-liquid equilibrium line (i.e., of the melting curve) is positive. There are some substances, like water, which contract upon melting so that $\Delta V < 0$. For these the slope of the melting curve must thus be negative (as drawn in Fig. 8·5·3).

Approximate calculation of the vapor pressure The Clausius-Clapeyron equation can be used to derive an approximate expression for the pressure of a vapor in equilibrium with the liquid (or solid) at a temperature *T*. This pressure is called the “vapor pressure” of the liquid (or solid) at this temperature. By (8·5·11)

$$\frac{dp}{dT} = \frac{l}{T \Delta v} \quad (8\cdot5\cdot12)$$

* The gas phase is sometimes also called the “vapor phase.” The transformation from solid to liquid is called “melting,” that from liquid to gas is called “vaporization,” and that from solid to gas is called “sublimation.”

† An exceptional case occurs for solid He* in a certain temperature range where the nuclear spins in the solid are randomly oriented while those in the liquid are aligned anti-parallel to each other so as to satisfy the quantum-mechanical Fermi-Dirac statistics.

where $l \equiv l_{12}$ is the latent heat per mole and v is the molar volume. Let 1 refer to the liquid (or solid) phase and 2 to the vapor. Then

$$\Delta v = v_2 - v_1 \approx v_2$$

since the vapor is much less dense than the liquid, so that $v_2 \gg v_1$. Let us also assume that the vapor can be adequately treated as an ideal gas, so that its equation of state is simply

$$pv_2 = RT$$

Then $\Delta v = RT/p$ and (8·5·12) becomes

$$\frac{1}{p} \frac{dp}{dT} = \frac{l}{RT^2} \quad (8 \cdot 5 \cdot 13)$$

Assume that l is approximately temperature independent. Then (8·5·13) can be immediately integrated to give

$$\ln p = -\frac{l}{RT} + \text{constant}$$

$$\text{or} \quad p = p_0 e^{-l/RT} \quad (8 \cdot 5 \cdot 14)$$

where p_0 is some constant. This shows that the vapor pressure p is a very rapidly increasing function of T , the temperature dependence being determined by the magnitude of the latent heat of vaporization.

8 · 6 Phase transformations and the equation of state

Consider a single-component system. Suppose that the equation of state for a mole of this substance

$$p = p(v, T) \quad (8 \cdot 6 \cdot 1)$$

is assumed known (by theoretical considerations or from empirical information) for the range of variables where the substance is a gas or a liquid. For example, the equation of state might be the van der Waals equation mentioned in (5·8·13), i.e.,

$$\left(p + \frac{a}{v^2} \right) (v - b) = RT$$

The equation of state (8·6·1) can be represented in a two-dimensional diagram by drawing a set of curves of mean pressure p versus the molar volume v for various values of the temperature T . Such a diagram is illustrated schematically in Fig. 4·6·1.*

Consider the system described by (8·6·1) to be in contact with a reservoir at given temperature T and pressure p . The intensive parameters T and p can then be regarded as the independent variables of the problem. Focus attention

* Upon multiplication by v^2 , the van der Waals equation is seen to be cubic in v . Thus it also gives rise to S-shaped curves of the type illustrated in Fig. 4·6·1.

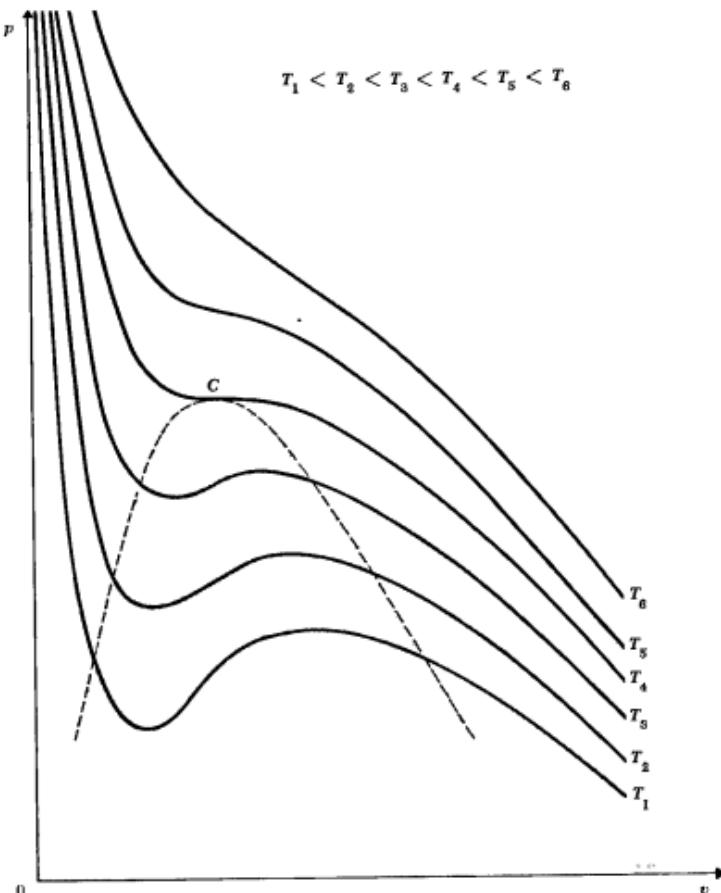


Fig. 8·6·1 Schematic diagram showing curves of constant temperature T for an equation of state (8·6·1) describing the fluid states of a substance. The point C is the critical point. In the shaded region mixtures of two phases can coexist along the horizontal lines.

on a particular curve of constant temperature T (or "isotherm") of the equation of state (8·6·1). A curve of this type is illustrated in Fig. 8·6·2 and contains a wealth of information. If at the given temperature T the pressure is sufficiently low so that $p < p_1$, the curve yields, correspondingly, a unique value of v . There exists then a well-defined single phase. Here the slope of the curve $\partial p / \partial v \leq 0$ as is necessary by the stability condition (8·4·15). Also, $|\partial p / \partial v|$ is relatively small, so that the compressibility of this phase is relatively large, as would be the case for a gaseous phase.

If at the given temperature T the pressure is sufficiently high that $p > p_2$, then there exists again a single phase with a unique value of v . The stability condition $\partial p / \partial v \leq 0$ is again satisfied, but $|\partial p / \partial v|$ is relatively large. Hence

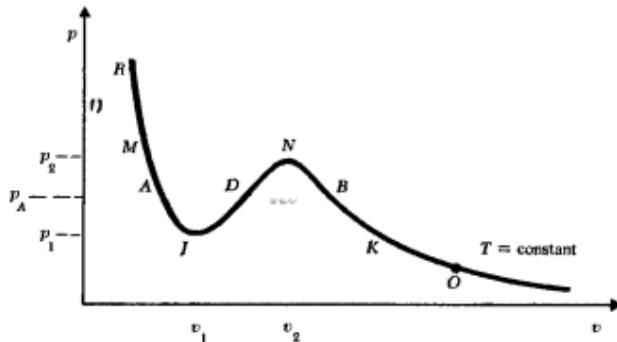


Fig. 8·6·2 Plot of the equation of state (8·6·1) for a particular temperature T .

the compressibility of this phase is relatively small as would be the case for a liquid phase.

Now consider the intermediate pressure range $p_1 < p < p_2$. At the given temperature T there are now, for each pressure p , three possible values of the volume v . The question is which value of v corresponds to the most stable situation. We see immediately that the stability condition $\partial p / \partial v \leq 0$ is violated in the region $v_1 < v < v_2$ where the curve has positive slope. Thus values of v in this range are certainly excluded since they would lead to an intrinsically unstable situation. But this still leaves two possible values of v between which one has to decide on the basis of relative stability. On the diagram, where we have labeled points on the curve by letters, one has to ask which is the more stable situation—where $v = v_A$ or where $v = v_B$. This question then reduces, by virtue of the discussion of Sec. 8·3, to an investigation of the relative magnitude of the molar-free energies $g_A(T, p)$ and $g_B(T, p)$.

Changes of the function $g \equiv \epsilon - Ts + pv$ can readily be computed along the constant temperature curve of Fig. 8·6·2. From the general thermodynamic relation

$$T \, ds = d\epsilon + p \, dv$$

it follows immediately that for a pressure change where T is kept constant,

$$dg = d(\epsilon - Ts + pv) = v \, dp \quad (8 \cdot 6 \cdot 2)$$

Differences between g on any point of the curve of Fig. 8·6·2 and some standard point O are then given by

$$g - g_O = \int_{p_0}^p v \, dp \quad (8 \cdot 6 \cdot 3)$$

The right side represents geometrically the area between the curve and the p axis in the range between p_0 and p . Note that, starting from the point O on the curve and performing the integral (8·6·3) along it, the value of this integral first increases until one reaches the point N , then decreases until one reaches the point J , and then increases again as one continues toward the point

M. Hence a curve of $g(T,p)$ versus p along the constant-temperature curve has the appearance shown in Fig. 8·6·3. (The points on this curve are labeled to correspond to those of Fig. 8·6·2.)

From this diagram one can readily see what happens for various values of the pressure. At O only the high-compressibility phase (in our example, the gas) exists. When the pressure is increased to the range $p_1 < p < p_2$, there are three possible values of g . The values of g along the curve $OKXN$ correspond to large values of $v > v_2$ in the region of high compressibility; this corresponds to the gas phase. The values of g along the curve $JXMR$ correspond to small values of $v < v_1$ in the region of low compressibility; this corresponds to the liquid phase. The values of g along the curve JDN correspond to the intrinsically unstable range $v_1 < v < v_2$. If p is only slightly larger than p_1 , then Fig. 8·6·3 shows that the gas phase, with volume near v_K , has the lower value of g and is thus the more stable one. This situation prevails until p is increased to the extent that it reaches the value $p = p_X$ corresponding to the point X where the curves KXN and JXM in Fig. 8·6·3 intersect. At this point the free energies g of both gas and liquid become equal. This then is the pressure at which arbitrary proportions of both these phases can coexist in equilibrium with each other. If the pressure is increased beyond p_X , the curve JXM corresponding to the liquid phase yields the lower free energy g so that this phase is the more stable one. At the point X the system shifts, therefore, from the curve $OKXN$ (corresponding to the gas phase) over to the curve $JXMR$ (corresponding to the liquid phase). Thus p_X corresponds to the pressure where the phase transformation from gas to liquid occurs.

Let us look at the phase transformation in greater detail. Assume that in Fig. 8·6·2 the pressure of the phase transformation is $p_A = p_X$. Then A

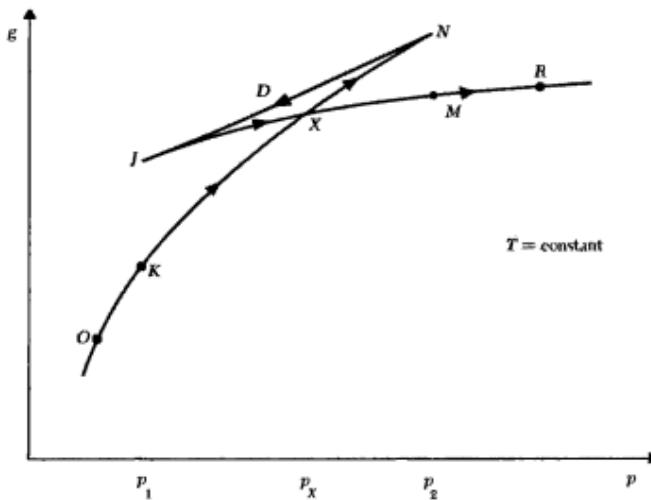


Fig. 8·6·3 Schematic illustration showing, as a function of pressure p , the behavior of $g(T,p)$ implied by the curve of Fig. 8·6·2.

and B both correspond to the point X in Fig. 8·6·3. Furthermore v_B is the molar volume of the gas and v_A that of the liquid at the pressure and temperature of the phase transformation. If under these circumstances a fraction ξ of the mole of substance is in the gaseous phase, then the total molar volume v_{tot} is given by

$$v_{\text{tot}} = \xi v_B + (1 - \xi) v_A \quad (8 \cdot 6 \cdot 4)$$

In the course of the phase transformation the total molar volume changes continuously from the value v_B for the gas to the value v_A for the liquid as the fraction ξ changes continuously from 1 to 0. In this process there will, of course, be a change of entropy and an associated latent heat. In Fig. 8·6·2 the horizontal line BDA along which the phase transformation occurs is characterized by the fact that

$$g_B = g_A \quad (8 \cdot 6 \cdot 5)$$

or, by virtue of (8·6·3), by the fact that the integral

$$\int_{BNDJA} v \, dp = 0 \quad (8 \cdot 6 \cdot 6)$$

when evaluated along the curve $BNDJA$ in Fig. 8·6·2. This integral can be broken up into several parts giving contributions of different signs:

$$\begin{aligned} \int_B^N v \, dp + \int_N^D v \, dp + \int_D^J v \, dp + \int_J^A v \, dp &= 0 \\ \left(\int_B^N v \, dp - \int_D^N v \, dp \right) + \left(- \int_J^D v \, dp + \int_J^A v \, dp \right) &= 0 \end{aligned}$$

or $\text{area } (DNB) - \text{area } (AJD) = 0$

where $\text{area } (DNB)$ denotes the area enclosed by the straight line DB and the curve DNB , and similarly, where $\text{area } (AJD)$ denotes the area enclosed by the straight line AD and the curve AJD . Thus we have the result that the location of the phase-transformation line ADB in Fig. 8·6·2 is determined by the condition that

► $\text{area } (AJD) = \text{area } (DNB) \quad (8 \cdot 6 \cdot 7)$

Remark The molar entropy change Δs and associated latent heat l of the transformation can also be determined from the equation of state. Since T is constant in the transformation so that $dT = 0$, one has simply

$$ds = \left(\frac{\partial s}{\partial v} \right)_T dv = \left(\frac{\partial p}{\partial T} \right)_v dv \quad (8 \cdot 6 \cdot 8)$$

where we have used a Maxwell relation in the last step. In going from A to B in Fig. 8·6·2, the corresponding entropy change can then be computed by evaluating the integral

$$\Delta s = s_B - s_A = \int_{AJDNB} \left(\frac{\partial p}{\partial T} \right)_v dv \quad (8 \cdot 6 \cdot 9)$$

Consider the given curve for temperature T and a neighboring one for tempera-

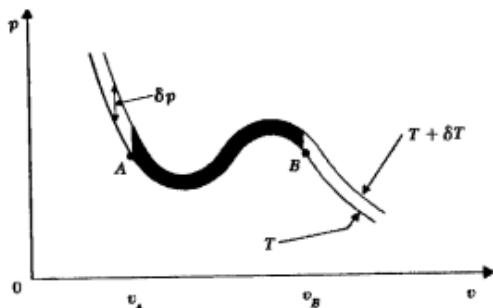


Fig. 8·6·4 Dependence of the pressure p on the molar volume v for two slightly different temperatures.

ture $T + \delta T$. At a given volume v , let δp denote the pressure difference between these two curves. Then (8·6·9) can be written

$$\Delta s = \frac{1}{\delta T} \int_{AJDNB} \delta p \, dv$$

$$\text{or } \Delta s = \frac{1}{\delta T} [\text{area between isotherms in the interval } v_A < v < v_B] \quad (8\cdot6\cdot10)$$

This area is shown shaded in Fig. 8·6·4.

This area is positive so that $\Delta s > 0$ or $s_B > s_A$, i.e., the entropy of the gas is greater than that of the liquid. The latent heat per mole absorbed in the transformation from liquid to gas is simply given by

$$l = T \Delta s \quad (8\cdot6\cdot11)$$

The change in the molar internal energy of the two phases is given by the fundamental relation $T ds = d\epsilon + p dv$ as

$$\Delta\epsilon = l - p \Delta v \quad (8\cdot6\cdot12)$$

since T and p are constant. Here $\Delta v = v_B - v_A$ and $\Delta\epsilon = \epsilon_B - \epsilon_A$.

As one goes to higher temperatures the two extremum points where $(\partial p / \partial v)_T = 0$ (indicated in Fig. 8·6·2 by $v = v_1$ and $v = v_2$) move closer together. This also implies that the volume change Δv in the phase transformation decreases. As the temperature is increased further one reaches the situation where these two extremum points v_1 and v_2 just coincide, so that there is no longer a change of sign of $(\partial p / \partial v)$, i.e., the derivative of $(\partial p / \partial v)$ vanishes also. At this point one has then $(\partial^2 p / \partial v^2)_T = 0$ as well as $(\partial p / \partial v)_T = 0$; the point is thus a point of inflection on the pv curve. This unique point is the "critical point" (point C in Fig. 8·6·1) and corresponds to values of T , p , and v which are called the critical temperature, pressure, and volume. There the phase transformation has barely disappeared, the volume change Δv having approached zero.* At still higher temperatures $(\partial p / \partial v) < 0$ everywhere so

* Since $(\partial p / \partial v)_T = 0$ at the critical point, it follows from our previous discussion following (8·4·22) that density fluctuations at this point become very large; i.e., the substance "cannot make up its mind" whether to be a liquid or a gas.

that there is no phase transformation. One then deals always with a single fluid phase with no sharp distinction between gas and liquid. As the pressure is raised one then goes *continuously* from the situation of large molar volume v and high compressibility to the situation of small v and low compressibility.

SYSTEMS WITH SEVERAL COMPONENTS; CHEMICAL EQUILIBRIUM

8 · 7 General relations for a system with several components

Consider a homogeneous system, of energy E and of volume V , which consists of m different kinds of molecules. Let N_i be the number of molecules of type i . Then the entropy of the system is a function of the following variables:

$$S = S(E, V, N_1, N_2, \dots, N_m) \quad (8 \cdot 7 \cdot 1)$$

These variables can, of course, all change in a general process. For example, the numbers of molecules of each species may change as a result of chemical reactions. In a completely general infinitesimal quasi-static process the entropy change is then given by*

$$dS = \left(\frac{\partial S}{\partial E}\right)_{V,N} dE + \left(\frac{\partial S}{\partial V}\right)_{E,N} dV + \sum_{i=1}^m \left(\frac{\partial S}{\partial N_i}\right)_{E,V,N} dN_i \quad (8 \cdot 7 \cdot 2)$$

Here the subscript N denotes the fact that all the numbers $\{N_1, \dots, N_m\}$ are kept constant in taking the partial derivative. In the case of a derivative such as $(\partial S / \partial N_i)$, the subscript N denotes the fact that all the numbers $\{N_1, \dots, N_{i-1}, N_{i+1}, \dots, N_m\}$, except the one number N_i with respect to which the derivative is taken, are kept constant in taking the partial derivative.

Equation (8 · 7 · 2) is a purely mathematical statement. But in the simple case when all the numbers N_i are kept fixed, the fundamental thermodynamic relation asserts that

$$dS = \frac{dQ}{T} = \frac{dE + p dV}{T} \quad (8 \cdot 7 \cdot 3)$$

Under these circumstances $dN_i = 0$ for all i in Eq. (8 · 7 · 2); comparison of the coefficients of dE and dV in (8 · 7 · 2) and (8 · 7 · 3) then yields

$$\begin{aligned} \left(\frac{\partial S}{\partial E}\right)_{V,N} &= \frac{1}{T} \\ \left(\frac{\partial S}{\partial V}\right)_{E,N} &= \frac{p}{T} \end{aligned} \quad (8 \cdot 7 \cdot 4)$$

* The system is, in general, interacting with some other systems. The variables in (8 · 7 · 2) and subsequent relations are to be evaluated for the equilibrium (or most probable) situation when the values of these variables are essentially equal to their mean values, \bar{E} , \bar{V} , \bar{N}_i , \bar{p} , and so on. We shall, however, omit writing the averaging bars over these symbols.

Let us introduce the abbreviation

$$\mu_j \equiv -T \left(\frac{\partial S}{\partial N_j} \right)_{E,V,N} \quad (8 \cdot 7 \cdot 5)$$

The quantity μ_j is called the “chemical potential per molecule” of the j th chemical species and has been defined so that it has the dimensions of energy. Then (8·7·2) can be written in the form

$$dS = \frac{1}{T} dE + \frac{p}{T} dV - \sum_i \frac{\mu_i}{T} dN_i \quad (8 \cdot 7 \cdot 6)$$

or

$$\blacktriangleright \quad dE = T dS - p dV + \sum_{i=1}^m \mu_i dN_i \quad (8 \cdot 7 \cdot 7)$$

This is just a generalization of the fundamental relation $dE = T dS - p dV$ to the case where the numbers of particles are allowed to vary.

Note that the chemical potential μ_j can be written in many forms equivalent to (8·7·5). For example, suppose that all the independent variables other than N_j are kept constant in (8·7·7). Then $dS = dV = 0$, $dN_i = 0$ for $i \neq j$, and (8·7·7) yields the relation

$$\mu_j = \left(\frac{\partial E}{\partial N_j} \right)_{S,V,N} \quad (8 \cdot 7 \cdot 8)$$

Alternatively, one can write (8·7·7) in the form

$$d(E - TS) = dF = -S dT - p dV + \sum_i \mu_i dN_i \quad (8 \cdot 7 \cdot 9)$$

If all independent variables other than N_j are kept constant, it follows immediately that

$$\mu_j = \left(\frac{\partial F}{\partial N_j} \right)_{T,V,N} \quad (8 \cdot 7 \cdot 10)$$

One can also write (8·7·7) in terms of the Gibbs free energy; thus

$$d(E - TS + pV) = dG = -S dT + V dp + \sum_i \mu_i dN_i \quad (8 \cdot 7 \cdot 11)$$

Hence one can also write

$$\mu_j = \left(\frac{\partial G}{\partial N_j} \right)_{T,p,N} \quad (8 \cdot 7 \cdot 12)$$

If there is *only one* chemical species present, say species j , then

$$G = G(T, p, N_j)$$

But G must be an extensive quantity. Thus, if all the independent extensive parameters are multiplied by a scale factor α , i.e., if N_j is multiplied by α , then G must be multiplied by the same factor α . Thus G must be proportional to N_j and can be written in the form

$$G(T, p, N_j) = N_j g'(T, p)$$

where $g'(T, p)$ does not depend on N_j . Then

$$\mu_j = \left(\frac{\partial G}{\partial N_j} \right)_{T, p} = g'(T, p) \quad (8 \cdot 7 \cdot 13)$$

i.e., the chemical potential per molecule is just equal to the Gibbs free energy $g' = G/N_j$ per molecule.

When several components are present, then $G = G(T, p, N_1, \dots, N_m)$ and in general

$$\mu_j = \left(\frac{\partial G}{\partial N_j} \right)_{T, p, N} \neq \frac{G}{N_j}$$

Remark The requirement that extensive quantities scale properly leads, however, to the following general conclusions. Consider, for example, the total energy

$$E = E(S, V, N_1, N_2, \dots, N_m) \quad (8 \cdot 7 \cdot 14)$$

If one increases all extensive variables by the same scale factor, Eq. (8·7·14) must remain valid. That is, if

$$S \rightarrow \alpha S, \quad V \rightarrow \alpha V, \quad N_i \rightarrow \alpha N_i$$

then one must also have $E \rightarrow \alpha E$. Thus

$$E(\alpha S, \alpha V, \alpha N_1, \dots, \alpha N_m) = \alpha E(S, V, N_1, \dots, N_m) \quad (8 \cdot 7 \cdot 15)$$

In particular let

$$\alpha = 1 + \gamma$$

where $|\gamma| \ll 1$. Then the left side of (8·7·15) becomes $E(S + \gamma S, V + \gamma V, N_1 + \gamma N_1, \dots)$ which can be expanded about its value $E(S, V, N_1, \dots)$, where $\gamma = 0$. Thus (8·7·15) implies the requirement that

$$E + \left(\frac{\partial E}{\partial S} \right)_{V, N} \gamma S + \left(\frac{\partial E}{\partial V} \right)_{S, N} \gamma V + \sum_{i=1}^m \left(\frac{\partial E}{\partial N_i} \right)_{S, V, N} \gamma N_i = (1 + \gamma)E$$

or* $E = \left(\frac{\partial E}{\partial S} \right)_{V, N} S + \left(\frac{\partial E}{\partial V} \right)_{S, N} V + \left(\frac{\partial E}{\partial N_i} \right)_{S, V, N} N_i \quad (8 \cdot 7 \cdot 16)$

But the derivatives are precisely given by the respective coefficients of dS , dV , and dN_i in (8·7·7). Hence (8·7·16) is equivalent to the relation

$$E = TS - pV + \sum_i \mu_i N_i \quad (8 \cdot 7 \cdot 17)$$

or

$$\blacktriangleright G = E - TS + pV = \sum_i \mu_i N_i \quad (8 \cdot 7 \cdot 18)$$

* This purely mathematical consequence of (8·7·15) is commonly referred to as "Euler's theorem for homogeneous functions."

If only a single kind j of molecule is present, (8·7·18) reduces to the previous relation

$$\mu_j = \frac{G}{N_j}$$

Eq. (8·7·17) implies that

$$dE = T dS + S dT - p dV - V dp + \sum_i \mu_i dN_i + \sum_i N_i d\mu_i$$

But since (8·7·7) must also be valid, one obtains the general result

$$S dT - V dp + \sum_i N_i d\mu_i = 0 \quad (8\cdot7\cdot19)$$

(This is known as the "Gibbs-Duhem relation.")

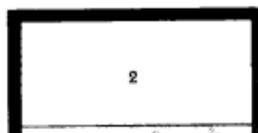
8·8 Alternative discussion of equilibrium between phases

In Sec. 8·5 we treated the problem of the equilibrium between two phases under conditions where the system was considered to be in equilibrium with a reservoir at constant temperature and pressure. It is instructive to treat this problem from a somewhat more general point of view, considering the total system to be isolated. Our discussion will be a straightforward extension of that used at the end of Sec. 3·9.

Consider N molecules forming a substance which consists of two phases denoted by 1 and 2. The whole system is isolated so that its total energy E and its total volume V are both fixed. Let there be N_i molecules of the substance in phase i , and denote the energy of this phase by E_i and its volume by V_i . Then we have the conservation conditions

$$\left. \begin{aligned} E_1 + E_2 &= E = \text{constant} \\ V_1 + V_2 &= V = \text{constant} \\ N_1 + N_2 &= N = \text{constant} \end{aligned} \right\} \quad (8\cdot8\cdot1)$$

The entropy of the whole system (or the total number of states accessible to the whole system) is a function of these parameters. The equilibrium condi-



1

Fig. 8·8·1 Equilibrium between two phases forming an isolated system of constant total energy and volume.

tion corresponding to the most probable situation is that the entropy is a maximum, i.e., that

$$S = S(E_1, V_1, N_1; E_2, V_2, N_2) = \text{maximum} \quad (8 \cdot 8 \cdot 2)$$

But

$$S = S_1(E_1, V_1, N_1) + S_2(E_2, V_2, N_2)$$

where S_i is the entropy of phase i . Thus the maximum condition (8·8·2) yields

$$dS = dS_1 + dS_2 = 0 \quad (8 \cdot 8 \cdot 3)$$

subject to the conditions (8·8·1), which become, in differential form,

$$\left. \begin{aligned} dE_1 + dE_2 &= 0 \\ dV_1 + dV_2 &= 0 \\ dN_1 + dN_2 &= 0 \end{aligned} \right\} \quad (8 \cdot 8 \cdot 4)$$

By using the relation (8·7·6) for each phase, (8·8·3) gives

$$dS = \left(\frac{1}{T_1} dE_1 + \frac{p_1}{T_1} dV_1 - \frac{\mu_1}{T_1} dN_1 \right) + \left(\frac{1}{T_2} dE_2 + \frac{p_2}{T_2} dV_2 - \frac{\mu_2}{T_2} dN_2 \right) = 0$$

or

$$dS = \left(\frac{1}{T_1} - \frac{1}{T_2} \right) dE_1 + \left(\frac{p_1}{T_1} - \frac{p_2}{T_2} \right) dV_1 - \left(\frac{\mu_1}{T_1} - \frac{\mu_2}{T_2} \right) dN_1 = 0 \quad (8 \cdot 8 \cdot 5)$$

where we have used the conservation conditions (8·8·4). Since (8·8·5) is to be valid for arbitrary variations dE_1, dV_1, dN_1 , it follows that the coefficients of all these differentials must separately vanish. Thus one obtains

$$\begin{aligned} \frac{1}{T_1} - \frac{1}{T_2} &= 0 \\ \frac{p_1}{T_1} - \frac{p_2}{T_2} &= 0 \\ \frac{\mu_1}{T_1} - \frac{\mu_2}{T_2} &= 0 \end{aligned}$$

or

$$\begin{aligned} T_1 &= T_2 \\ p_1 &= p_2 \\ \mu_1 &= \mu_2 \end{aligned} \quad (8 \cdot 8 \cdot 6)$$

These are the necessary conditions for equilibrium between two phases; they reflect the respective conservation equations (8·8·1). The relations (8·8·6) assert that the temperatures and mean pressures of the phases must be equal, as one would expect. The condition that the chemical potentials must also be equal may seem less familiar. But since each phase consists of only a single component, it follows by (8·7·13) that $\mu_i = g'_i$ is the chemical potential per molecule of phase i . The last relation of (8·8·6) is then equivalent to

$$g'_1 = g'_2 \quad (8 \cdot 8 \cdot 7)$$

Thus we regain the condition (8·5·4).

Remark In Sec. 8·5 we used the Gibbs free energy g per mole; thus $g_i = N_a g'_i$ where N_a is Avogadro's number. In the same way, it is sometimes useful to define a chemical potential per mole. This is given by the relation $(\partial G/\partial \nu_i)$; since $\nu_i = N_i/N_a$, it is N_a times larger than the corresponding chemical potential per molecule $(\partial G/\partial N_i)$.

One could readily extend the arguments of this section to treat the equilibrium between phases containing several components, or to calculate the fluctuations in the number of particles present in each phase.

Finally it is worth noting the microscopic implications of our discussion. One can take the equilibrium condition (8·8·6) or (8·8·7) and write it in the form

$$\mu_1(T,p) = \mu_2(T,p) \quad (8\cdot8\cdot8)$$

where we have expressed the chemical potentials in terms of T and p . Now we know, at least in principle, how to use statistical mechanics to calculate thermodynamic functions such as the entropy S for each phase. We can then go on to calculate for each phase the chemical potential μ (say, by its definition (8·7·5)) and we can express it in terms of T and p . The result is an equation of the form (8·8·8) which can be solved to find p as a function of T . In this way one can, for example, start from first principles to compute the vapor pressure of a substance at a given temperature. The only difficulty in such a calculation is the evaluation of the partition function for each phase. We shall illustrate such a vapor-pressure calculation in the next chapter.

8·9 General conditions for chemical equilibrium

Consider a homogeneous system (consisting of a single phase) which contains m different kinds of molecules. Let us designate the chemical symbols of these molecules by B_1, B_2, \dots, B_m . Assume that there exists the possibility of one chemical reaction occurring between these molecules whereby molecules can get transformed into each other. This chemical transformation must be consistent with the conservation of the total number of atoms of each kind. A properly balanced chemical equation expresses precisely these conservation conditions.

Example Suppose that the system consists of H_2 , O_2 , and H_2O molecules in the gas phase. There exists the possibility of transformation of molecules into each other through the chemical reaction



This chemical equation is properly balanced so that the total number of H atoms is the same on the left and right sides, as is the total number of O atoms.

Let b_i denote the coefficient of B_i in the chemical equation; thus b_i is some small integer designating the number of B_i molecules involved in the chemical transformation. For convenience we adopt the convention that, if the chemical reaction is regarded arbitrarily as proceeding in a given direction, one considers b_i positive for any "product" molecule formed as a result of the chemical reaction, and one considers b_i negative for any "reactant" molecule disappearing as a result of the reaction. For example, the reaction



would be written in standard form as

$$-2\text{H}_2 - \text{O}_2 + 2\text{H}_2\text{O} = 0 \quad (8 \cdot 9 \cdot 1)$$

A general chemical equation can then be written in the form

► $\sum_{i=1}^m b_i B_i = 0 \quad (8 \cdot 9 \cdot 2)$

Let N_i denote the number of B_i molecules in the system. The numbers N_i can change as a result of the chemical reaction between molecules. But they cannot change in arbitrary ways because the conservation of atoms requires that the chemical equation (8·9·2) be satisfied. The changes in the numbers N_i must therefore be proportional to the numbers of molecules appearing in the balanced chemical equation (8·9·2); i.e.,

$$dN_i = \lambda b_i \quad \text{for all } i \quad (8 \cdot 9 \cdot 3)$$

where λ is a constant of proportionality. Here $dN_i > 0$ for molecules formed as a result of the reaction, and $dN_i < 0$ for molecules disappearing as a result of the reaction.

Example In the reaction (8·9·1) the number of molecules $N_{\text{H}_2\text{O}}$, N_{H_2} , N_{O_2} can only vary in such a way that changes in these numbers are in the proportion

$$dN_{\text{H}_2\text{O}} : dN_{\text{H}_2} : dN_{\text{O}_2} = 2 : -2 : -1$$

Consider now an *equilibrium* situation where molecules confined within an isolated enclosure of volume V can react chemically with each other in accordance with the reaction (8·9·2). Let E denote the total energy of the system. The equilibrium condition then is that

$$S = S(E, V, N_1, \dots, N_m) = \text{maximum} \quad (8 \cdot 9 \cdot 4)$$

or $dS = 0 \quad (8 \cdot 9 \cdot 5)$

Under the assumed conditions of constant V and E , this condition becomes, by virtue of (8·7·6),

$$\sum_{i=1}^m \mu_i dN_i = 0 \quad (8 \cdot 9 \cdot 6)$$

Using the fact that the variations dN_i must satisfy the chemical equation (8·9·2), i.e., the restriction (8·9·3), the condition (8·9·6) then becomes simply

$$\blacktriangleright \quad \sum_{i=1}^m b_i \mu_i = 0 \quad (8\cdot9\cdot7)$$

This is the general condition for chemical equilibrium.

Remark If we had assumed that the reaction takes place under conditions of constant temperature T and volume V , then the condition that F is minimum or $dF = 0$ under these circumstances would, by virtue of (8·7·9), again lead to (8·9·6) and (8·9·7). Similarly, for a reaction at constant temperature T and pressure p , the condition that G is minimum or $dG = 0$ would, by virtue of (8·7·11), also lead to (8·9·6) and (8·9·7).

The chemical potentials μ_i are functions of the variables describing the system. For example, $\mu_i = \mu_i(E, V, N_1, \dots, N_m)$, if E and V are chosen as independent variables, or $\mu_i = \mu_i(T, V, N_1, \dots, N_m)$ if T and V are chosen as independent variables. Hence the condition (8·9·7) implies in an equilibrium situation a definite connection between the mean numbers N_i of molecules of each kind. Since statistical thermodynamics allows one to calculate thermodynamic functions like the entropy S , it also makes it possible to calculate the chemical potentials μ_i and thus to deduce explicitly the connection between the numbers N_1, \dots, N_m implied by the condition (8·9·7). We shall illustrate such a calculation in the next section.

8·10 Chemical equilibrium between ideal gases

Consider that the chemical reaction (8·9·2) can occur between m different types of molecules. Suppose that these form gases which can be considered ideal and that they are confined in a container of volume V at an absolute temperature T . What is the relationship between the mean numbers of the reacting molecules in the equilibrium situation?

The question is readily answered by exploiting the condition (8·9·7) explicitly. In other words, suppose that one knows the free energy

$$F = F(T, V, N_1, \dots, N_m)$$

of this mixture of gases. If at the constant specified temperature T and volume V one imagines a transformation in which, in accordance with the chemical equation (8·9·2), $|b_i|$ of each of the reactant molecules are transformed into b_i of each of the product molecules, then the corresponding small free-energy change ΔF in the reaction is given by

$$\Delta F = \sum_i \left(\frac{\partial F}{\partial N_i} \right)_{T, V, N} b_i = \sum_i \mu_i b_i \quad (8\cdot10\cdot1)$$

Here

$$\mu_i \equiv \left(\frac{\partial F}{\partial N_i} \right)_{T, V, N} \quad (8 \cdot 10 \cdot 2)$$

is, in accordance with the relation (8·7·10), just the chemical potential of a molecule of type i and is, like F itself, a function of the variables T, V, N_1, \dots, N_m . In equilibrium the free energy F is a minimum, so that (8·10·1) yields the familiar condition (8·9·7):

$$\Delta F = \sum_i b_i \mu_i = 0 \quad (8 \cdot 10 \cdot 3)$$

The remaining task is then merely that of calculating F , and hence the chemical potentials μ_i , of this mixture of gases.

Calculation of the chemical potential We again consider the gases to be at sufficiently high temperature and sufficiently low density that their translational motion can be treated classically.

Let the possible states of the k th molecule in the gas be labeled by s_k and let $\epsilon_k(s_k)$ denote the energy of the molecule in this state. Since there is negligible interaction between the molecules, the total energy of the gas in one of its possible states can always be written as a sum

$$E = \epsilon_1(s_1) + \epsilon_2(s_2) + \epsilon_3(s_3) + \dots$$

with a number of terms equal to the total number of molecules. Hence the partition function becomes (treating all molecules as distinguishable)

$$Z' = \sum_{s_1, s_2, s_3} e^{-\beta[\epsilon_1(s_1) + \epsilon_2(s_2) + \dots]}$$

where the summation is over *all* the states of *each* molecule. As usual this factors into the form

$$Z' = \left(\sum_{s_1} e^{-\beta \epsilon_1(s_1)} \right) \left(\sum_{s_2} e^{-\beta \epsilon_2(s_2)} \right) \dots \quad (8 \cdot 10 \cdot 4)$$

In this product there will be N_i equal factors for all molecules of type i , each of these factors being equal to

$$\xi_i \equiv \sum_s e^{-\beta \epsilon(s)} \quad (8 \cdot 10 \cdot 5)$$

where the sum is over all the states s and corresponding energies of one molecule of type i . Thus (8·10·4) becomes simply

$$Z' = \xi_1^{N_1} \xi_2^{N_2} \dots \xi_m^{N_m} \quad (8 \cdot 10 \cdot 6)$$

Here we have counted as distinct states of the gas all those which differ only by permutations of like molecules. As we saw in Sec. 7·3, it would be erroneous (Gibbs paradox) and inconsistent with the essential indistinguishability of the molecules in quantum mechanics to count these gas states as distinct. To get the correct partition function Z , the expression (8·10·6) must therefore be divided by the $(N_1! N_2! \dots N_m!)$ possible permutations of like molecules among themselves. Thus we get

$$Z = \frac{\xi_1^{N_1} \xi_2^{N_2} \dots \xi_m^{N_m}}{N_1! N_2! \dots N_m!} \quad (8 \cdot 10 \cdot 7)$$

This can also be written

$$\blacktriangleright \quad Z = Z_1 Z_2 \cdots Z_m \quad (8 \cdot 10 \cdot 8)$$

where

$$\blacktriangleright \quad Z_i = \frac{\xi_i^{N_i}}{N_i!} \quad (8 \cdot 10 \cdot 9)$$

is the partition function of a gas of N_i molecules occupying the given volume V by itself in the absence of all other gases.

A variety of important results follow from (8·10·8), i.e., from the relation

$$\ln Z = \sum_i \ln Z_i \quad (8 \cdot 10 \cdot 10)$$

These results all reflect the fact that the molecules are weakly interacting so that the thermodynamic functions are simply additive. For example, since the mean energy of a system is given quite generally by $\bar{E} = (-\partial \ln Z / \partial \beta)$, it follows from (8·10·10) that

$$\bar{E}(T, V) = \sum_i \bar{E}_i(T, V) \quad (8 \cdot 10 \cdot 11)$$

where \bar{E}_i is the mean energy of the i th gas occupying the given volume by itself. Also, since the mean pressure of a system is given quite generally by $\bar{p} = \beta^{-1}(\partial \ln Z / \partial V)$, it follows from (8·10·10) that

$$\bar{p} = \sum_i \bar{p}_i \quad (8 \cdot 10 \cdot 12)$$

where \bar{p}_i is the mean pressure that would be exerted by the i th gas if it occupied the given volume V by itself. This quantity \bar{p}_i is called the "partial pressure of the i th gas."

Now we have already derived the equation of state for a single gas; i.e., for gas i occupying the volume V by itself

$$\bar{p}_i = n_i k T, \quad n_i \equiv \frac{N_i}{V} \quad (8 \cdot 10 \cdot 13)$$

Hence (8·10·12) gives us immediately the equation of state for the gas mixture

$$\bar{p} = n k T, \quad \text{where } n \equiv \sum_{i=1}^m n_i \quad (8 \cdot 10 \cdot 14)$$

Since the Helmholtz free energy is given quite generally by the relation $F = -kT \ln Z$, it follows from (8·10·10) that

$$F(T, V) = \sum_i F_i(T, V) \quad (8 \cdot 10 \cdot 15)$$

where F_i is the free energy of the i th gas by itself. Since $F = \bar{E} - TS$, (8·10·11) and (8·10·14) also establish the additivity of the entropies

$$S(T, V) = \sum_i S_i(T, V) \quad (8 \cdot 10 \cdot 16)$$

where S_i is the entropy of the i th gas occupying the given volume by itself.

Let us now proceed to calculate the chemical potential. By (8·10·9),

$$\ln Z_i = N_i \ln \xi_i - \ln N_i!$$

where $\xi_i = \xi_i(T, V)$ is the partition function (8·10·5) for a single molecule and thus does not involve N_i . Using (6·6·9) and (8·10·10), one then obtains

$$F = -kT \ln Z = -kT \sum_i (N_i \ln \xi_i - \ln N_i!) \quad (8\cdot10\cdot17)$$

or

$$F = -kT \sum_i N_i (\ln \xi_i - \ln N_i + 1)$$

where we have used Stirling's formula $\ln N! = N \ln N - N$. Since

$$\frac{\partial \ln (N!)}{\partial N} = \ln N$$

(a result already encountered in (1·5·9)), it then follows by (8·10·2) that the chemical potential of the j th kind of molecule is simply given by

$$\mu_j = \left(\frac{\partial F}{\partial N_j} \right)_{T, V, N} = -kT(\ln \xi_j - \ln N_j)$$

or

$$\blacktriangleright \quad \mu_j = -kT \ln \frac{\xi_j}{N_j} \quad (8\cdot10\cdot18)$$

Law of mass action By (8·10·1), the free-energy change in the reaction is then equal to

$$\Delta F = -kT \sum_i b_i (\ln \xi_i - \ln N_i) = \Delta F_0 + kT \sum_i b_i \ln N_i \quad (8\cdot10\cdot19)$$

where

$$\Delta F_0 \equiv -kT \sum_i b_i \ln \xi_i \quad (8\cdot10\cdot20)$$

is a quantity (the so-called "standard free-energy change of the reaction") which depends only on T and V , but not on the numbers N_i of molecules present. The equilibrium condition (8·10·3) then becomes

$$\begin{aligned} \Delta F &= \Delta F_0 + kT \sum_i b_i \ln N_i = 0 \\ \sum_i \ln N_i^{b_i} &= \ln (N_1^{b_1} N_2^{b_2} \cdots N_m^{b_m}) = -\frac{\Delta F_0}{kT} \end{aligned}$$

or

$$\blacktriangleright \quad N_1^{b_1} N_2^{b_2} \cdots N_m^{b_m} = K_N(T, V) \quad (8\cdot10\cdot21)$$

where

$$\blacktriangleright \quad K_N(T, V) \equiv e^{-\Delta F_0/kT} = \xi_1^{b_1} \xi_2^{b_2} \cdots \xi_m^{b_m} \quad (8\cdot10\cdot22)$$

The quantity K_N is independent of the numbers of molecules present and is called the "equilibrium constant"; it is a function only of T and V through the dependence of the molecular partition functions ξ_i on these quantities.

Equation (8·10·21) is the desired explicit relation between the mean numbers of molecules present in equilibrium; it is called the "law of mass action" and is likely to be familiar from elementary chemistry.

Example Consider the reaction (8·9·1) in the gas phase



The law of mass action (8·10·21) then becomes

$$N_{\text{H}_2}^{-2} N_{\text{O}_2}^{-1} N_{\text{H}_2\text{O}}^2 = K_N$$

or

$$\frac{N_{\text{H}_2\text{O}}^2}{N_{\text{H}_2}^{-2} N_{\text{O}_2}^{-1}} = K_N(T, V)$$

Note that (8·10·22) gives an explicit expression for the equilibrium constant K_N in terms of the partition functions ξ_j for each type of molecule. Hence K_N can be calculated from first principles if the molecules are sufficiently simple so that ξ_j in (8·10·5) can be evaluated from a knowledge of the quantum states of a single molecule. Even when the molecules are more complex, it is still possible to use spectroscopic data to deduce their energy levels and thus to calculate ξ_j and the equilibrium constant K_N .

Remark Note that the factor $N!l!$, present in (8·10·9) to take proper account of the indistinguishability of the molecules, is absolutely essential to the whole theory. If this factor were not present, one would simply have $\ln Z = \Sigma N_j \ln \xi_j$, so that the chemical potential would become $\mu_j = -kT \ln \xi_j$, independent of N_j . Hence (8·10·3) would give us no relationship between numbers at all! Thus the classical difficulties exemplified by the Gibbs paradox lead to nonsense all down the line.

It is worth noting the simplifying property that $\xi_j(V, T)$ is, as in (7·2·6), simply proportional to V . Indeed (8·10·5) can be written

$$\xi_j \propto \int d^3r \int d^3p e^{-\beta p^2/2m} \sum_i e^{-\beta \epsilon_i^{(j=0)}}$$

where we have calculated the translational part of the partition function classically. (The remaining sum is over the states of *internal* motion, both vibration and rotation, if the molecule is not monatomic.) Here the center of mass position r occurs only in the integral $\int d^3r$ which yields the volume V . Thus one can write

$$\xi_j(V, T) = V \xi'_j(T) \quad (8·10·23)$$

where ξ'_j depends only on T . Hence the chemical potential (8·10·18) becomes

$$\mu_j = -kT \ln \frac{\xi'_j}{n_j} \quad (8·10·24)$$

where $n_j = N_j/V$ is the number of molecules of type j per unit volume.

The fundamental equilibrium condition (8·10·3) can then be written more simply

$$\sum_i b_i \ln n_i = \sum_i b_i \ln \xi'_i$$

or

► $n_1^{b_1} n_2^{b_2} \cdots n_m^{b_m} = K_n(T)$ (8·10·25)

where

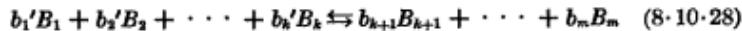
► $K_n(T) = \xi'_1^{b_1} \xi'_2^{b_2} \cdots \xi'_m^{b_m}$ (8·10·26)

and the equilibrium constant $K_n(T)$ depends *only* on the temperature.

By (8·10·22) and (8·10·26) one has the relation

$$K_N(T, V) = V^b K_n(T), \quad \text{where } b \equiv \sum_{i=1}^m b_i \quad (8·10·27)$$

Remark The law of mass action (8·10·25) can be given a detailed kinetic interpretation. Write the reaction (8·9·2) in the form



where $i = 1, \dots, k$ refers to the k kinds of reactant molecules (with $b_i' = -b_i > 0$) and where $i = k+1, \dots, m$ refers to the remaining $(m-k)$ kinds of product molecules. The probability P_+ per unit time that the reaction (8·10·28) occurs from left to right should be proportional to the probability of simultaneous encounter in a given element of volume of $b_1' B_1$ -molecules, $b_2' B_2$ -molecules, \dots , $b_k' B_k$ -molecules. Since molecules in an ideal gas are statistically independent, the probability of molecule i being in the volume element is simply proportional to n_i , and the probability P_+ becomes

$$P_+ = K_+(T) n_1^{b_1'} n_2^{b_2'} \cdots n_k^{b_k'}$$

where $K_+(T)$ is a constant of proportionality which can depend on T . Similarly, the probability P_- per unit time that the reaction (8·10·28) occurs from right to left should be proportional to the probability of simultaneous encounter of $b_{k+1}' B_{k+1}$ -molecules, \dots , $b_m' B_m$ -molecules, i.e.,

$$P_- = K_-(T) n_{k+1}^{b_{k+1}'} n_{k+2}^{b_{k+2}'} \cdots n_m^{b_m'}$$

In equilibrium one must have

$$P_+ = P_-$$

Hence it follows that

$$\frac{n_{k+1}^{b_{k+1}'} n_{k+2}^{b_{k+2}'} \cdots n_m^{b_m'}}{n_1^{b_1'} n_2^{b_2'} \cdots n_k^{b_k'}} = \frac{K_+(T)}{K_-(T)}$$

which is identical in form with (8·10·25).

Temperature dependence of the equilibrium constant The relation (8·10·22) gives explicitly

$$\ln K_N(T, V) = -\frac{\Delta F_0}{kT} \quad (8·10·29)$$

Hence $\left(\frac{\partial \ln K_N}{\partial T}\right)_V = -\left(\frac{\partial}{\partial T}\right)_V \left(\frac{\Delta F_0}{kT}\right) = -\left(\frac{\partial}{\partial T}\right)_{V,N} \frac{\Delta F}{kT}$ (8·10·30)

The last expression follows by (8·10·19), since ΔF and ΔF_0 differ only by an expression involving the numbers N_i ; and since these numbers are supposed to be held constant, as indicated, in the differentiation of $\Delta F/kT$. Hence

$$\left(\frac{\partial \ln K_N}{\partial T}\right)_V = \frac{1}{kT^2} \Delta F - \frac{1}{kT} \left(\frac{\partial \Delta F}{\partial T}\right)_{V,N} \quad (8·10·31)$$

But

$$-\left(\frac{\partial}{\partial T}\right) \Delta F = -\sum_i \left(\frac{\partial}{\partial T}\right) \left(\frac{\partial F}{\partial N_i}\right) b_i = -\sum_i \left(\frac{\partial}{\partial N_i}\right) \left(\frac{\partial F}{\partial T}\right) b_i = \sum \frac{\partial S}{\partial N_i} b_i \equiv \Delta S$$

Here we have used the general relation, implied by (8·7·9), that

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

and have denoted by ΔS the entropy change of the reaction when $|b_i|$ of each of the reactant molecules are transformed into b_i of each of the product molecules. Hence (8·10·31) becomes

$$\left(\frac{\partial \ln K_N}{\partial T}\right)_V = \frac{1}{kT^2} (\Delta F + T \Delta S) = \frac{\Delta E}{kT^2} \quad (8·10·32)$$

since $F \equiv E - TS$; thus $\Delta E = \Delta(F + TS)$ is simply the mean energy increase in the reaction. Since the reaction is carried out at constant volume, ΔE is also the heat absorbed in the reaction when $|b_i|$ of each of the reactant molecules are transformed into b_i of each of the product molecules.

Since K_n differs from K_N only by a factor involving the volume V , (8·10·32) implies equivalently that

$$\frac{d \ln K_n}{dT} = \frac{\Delta E}{kT^2} \quad (8·10·33)$$

If $\Delta E > 0$, (8·10·32) asserts that K_N increases as T is increased. This result is again in accord with what would be expected from Le Chatelier's principle. When $E > 0$, heat is absorbed as a result of the reaction. If the temperature T increases, more molecules must then be produced in order to absorb heat and thus to restore the original temperature. Thus K_N must increase.

SUGGESTIONS FOR SUPPLEMENTARY READING

Stability conditions

H. B. Callen: "Thermodynamics," chap. 8, John Wiley & Sons, Inc., New York, 1960.

Fluctuations of thermodynamic quantities

L. D. Landau and E. M. Lifshitz: "Statistical Physics," secs. 109–111, Addison-Wesley Publishing Company, Reading, Mass., 1959.

Phase transformations

- M. W. Zemansky: "Heat and Thermodynamics," 4th ed., chap. 15, McGraw-Hill Book Company, New York, 1957.
- W. P. Allis and M. A. Herlin: "Thermodynamics and Statistical Mechanics," secs. 33-36, McGraw-Hill Book Company, New York, 1952.
- H. B. Callen: "Thermodynamics," chap. 9, John Wiley & Sons, Inc., New York, 1960.
- A. B. Pippard: "The Elements of Classical Thermodynamics," chaps. 8 and 9, Cambridge University Press, Cambridge, 1957. (Includes a good discussion of "higher-order" phase transformations.)

Chemical equilibrium

- M. W. Zemansky: "Heat and Thermodynamics," 4th ed., chaps. 17 and 18, McGraw-Hill Book Company, New York, 1957.
- H. B. Callen: "Thermodynamics," chap. 12, John Wiley & Sons, Inc., New York, 1960.
- T. L. Hill: "An Introduction to Statistical Thermodynamics," chap. 10, Addison-Wesley Publishing Company, Reading, Mass., 1960.
- J. F. Lee, F. W. Sears, and D. L. Turcotte: "Statistical Thermodynamics," chap. 13, Addison-Wesley Publishing Company, Reading, Mass., 1963.

PROBLEMS

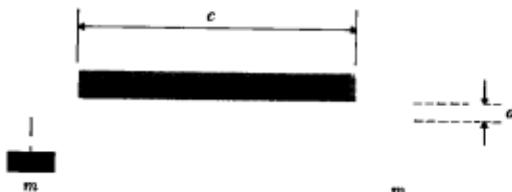
- 8.1** In some homogeneous substance at absolute temperature T_0 (e.g., a liquid or gas) focus attention on some small portion of mass M . This small portion is in equilibrium with the rest of the substance; it is large enough to be macroscopic and can be characterized by a volume V and temperature T . Calculate the probability $\mathcal{P}(V, T) dV dT$ that the volume of this portion lies between V and $V + dV$ and that its temperature lies between T and $T + dT$. Express your answer in terms of the compressibility κ of the substance, its density ρ_0 , and its specific heat per gram c_V at constant volume.
- 8.2** The vapor pressure p (in millimeters of mercury) of solid ammonia is given by $\ln p = 23.03 - 3754/T$ and that of liquid ammonia by $\ln p = 19.49 - 3063/T$.
- What is the temperature of the triple point?
 - What are the latent heats of sublimation and vaporization at the triple point?
 - What is the latent heat of melting at the triple point?
- 8.3** A simple substance of molecular weight μ has its triple point at the absolute temperature T_0 and pressure p_0 . At this point the densities of the solid and liquid are ρ_s and ρ_l , respectively, while the vapor can be approximated by a dilute ideal gas. If at this triple point the slope of the melting curve is $(dp/dT)_m$ and that of the liquid vaporization curve is $(dp/dT)_v$, what is the slope $(dp/dT)_s$ of the sublimation curve of the solid?
- 8.4** Helium remains a liquid down to absolute zero at atmospheric pressure, but becomes a solid at sufficiently high pressures. The density of the solid is, as usual, greater than that of the liquid. Consider the phase-equilibrium line between the solid and liquid. In the limit as $T \rightarrow 0$, is the slope dp/dT of this line positive, zero, or negative?
- 8.5** Liquid helium boils at a temperature T_0 (4.2°K) when its vapor pressure is equal to $p_0 = 1$ atmosphere. The latent heat of vaporization per mole of the liquid is equal to L and approximately independent of temperature. The liquid is

contained within a dewar which serves to insulate it thermally from the room temperature surroundings. Since the insulation is not perfect, an amount of heat Q per second flows into the liquid and evaporates some of it. (This heat influx Q is essentially constant, independent of whether the temperature of the liquid is T_0 or less.) In order to reach low temperatures one can reduce the pressure of the He vapor over the liquid by pumping it away with a pump at room temperature T_r . (By the time it reaches the pump, the He vapor has warmed up to room temperature.) The pump has a maximum pumping speed such that it can remove a constant volume V of gas per second, irrespective of the pressure of the gas. (This is a characteristic feature of ordinary mechanical pumps which simply sweep out a fixed volume of gas per revolution.)

(a) Calculate the minimum vapor pressure p_m which this pump can maintain over the surface of the liquid if the heat influx is Q .

(b) If the liquid is thus maintained in equilibrium with its vapor at this pressure p_m , calculate its approximate temperature T_m .

- 8.6 An atomic beam of sodium (Na) atoms is produced by maintaining liquid sodium in an enclosure at some elevated temperature T . Atoms of Na from the vapor above the liquid escape by effusion through a narrow slit in the enclosure and thus give rise to an atomic beam of intensity I . (The intensity I is defined as the number of atoms in the beam which cross unit area per unit time.) The latent heat of vaporization per mole of liquid Na into a vapor of Na atoms is L . To estimate how sensitive the beam intensity is to fluctuations in the temperature of the enclosure, calculate the relative intensity change $I^{-1}(dI/dT)$ in terms of L and the absolute temperature T of the enclosure.
- 8.7 The molar latent heat of transformation in going from phase 1 to phase 2 at the temperature T and pressure p is l . What is the latent heat of the phase transformation at a slightly different temperature (and corresponding pressure), i.e., what is (dl/dT) ? Express your answer in terms of l and the molar specific heat c_p , coefficient of expansion α , and molar volume v of each phase at the original temperature T and pressure p .
- 8.8 A steel bar of rectangular cross section (height a and width b) is placed on a block of ice with its ends extending a trifle as shown in the figure. A weight of mass m is hung from each end of the bar. The entire system is at 0°C . As a result of the pressure exerted by the bar, the ice melts beneath the bar and refreezes above the bar. Heat is therefore liberated above the bar, conducted through the metal, and then absorbed by the ice beneath the bar. (We assume that this is the most important way in which heat reaches the ice immediately beneath the bar in order to melt it.) Find an approximate expression for the speed with which the bar thus sinks through the ice. The answer should be in terms of the latent heat of fusion l per gram of ice, the densities ρ_i and ρ_w of ice and water respectively, the thermal conductivity κ of steel, the temperature T (0°C) of the ice, the acceleration due to gravity g , the mass m , and the dimensions a , b , and c , where c is the width of the block of ice.



- 8.9** Careful measurements were made of the vapor pressure of liquid pentane as a function of temperature. The temperature was measured very precisely in terms of the emf of a thermocouple whose reference junction was maintained at the triple point of water. Thus one determined the curve of vapor pressure p versus measured thermocouple emf ϕ . Also measured as a function of ϕ along the vapor pressure curve were the latent heat of vaporization L per gram of liquid pentane, and the volume change ΔV per gram of pentane in going from liquid to vapor. Show that this information is sufficient to calibrate the thermocouple; i.e., write an explicit expression (in terms of an integral) for the absolute temperature T when the measured thermocouple emf is ϕ .
- 8.10** Consider any substance in equilibrium in the presence of externally applied forces due to gravitational or electromagnetic fields. Focus attention on any two volume elements of this substance, both of fixed size and infinitesimally small on a macroscopic scale. By using the fact that the total entropy of the substance must remain stationary if a small amount of energy or a small number of particles, or both, are transferred from one of these volume elements to the other, show that the temperature T and the chemical potential μ must each have a constant value throughout the substance.
- 8.11** Consider a classical ideal gas in thermal equilibrium at temperature T in a container of volume V in the presence of a uniform gravitational field. The acceleration due to gravity is g and directed along the $-z$ direction.
- (a) Calculate the chemical potential μ of an element of volume of such a gas as a function of the pressure p , the temperature T , and the height z .
 - (b) Show that the requirement that μ is constant implies immediately the law of atmospheres which gives the dependence of p on T and z .
- 8.12** At a fixed temperature $T = 1200^\circ\text{K}$, the gases



are in chemical equilibrium in a vessel of volume V . If the volume of this vessel is increased, its temperature being maintained constant, does the relative concentration of CO_2 increase, decrease, or remain the same?

- 8.13** An experiment on iodine (I) atoms is carried out in a molecular beam apparatus. The beam is obtained by effusion of molecules from a small slit in an oven containing, as a result of thermal dissociation, a mixture of I_2 molecules and I atoms. If the temperature of the oven is kept the same but the total gas pressure within it is doubled, by what factor is the intensity of I atoms in the beam changed?
- 8.14** Consider the following chemical reaction between ideal gases:

$$\sum_{i=1}^m b_i B_i = 0$$

Let the temperature be T , the total pressure be p . Denote the partial pressure of the i th species by p_i . Show that the law of mass action can be put into the form

$$p_1^{b_1} p_2^{b_2} \cdots p_m^{b_m} = K_p(T)$$

where the constant $K_p(T)$ depends only on T .

- 8.15** Show that if the chemical reaction of the preceding problem is carried out under conditions of constant total pressure, the heat of reaction per mole (i.e., the heat which must be supplied to transform $|b_i|$ moles of each of the reactants to $|b_i|$ moles of each of the reaction products) is given by the enthalpy change

$$\Delta H = \Sigma b_i h_i$$

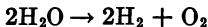
where h_i is the enthalpy per mole of the i th gas at the given temperature and pressure.

- 8.16** Show that

$$\frac{d \ln K_p}{dT} = \frac{\Delta H}{RT^2}$$

where R is the gas constant per mole.

- 8.17** Suppose that v_0 moles of H_2O gas are introduced into a container of fixed volume V at a temperature low enough so that virtually all the gas remains undisassociated in the form of H_2O vapor. At higher temperatures dissociation can take place according to the reaction



Let ξ denote the fraction of H_2O molecules which are dissociated at any temperature T corresponding to a total gas pressure p . Write an equation relating ξ to p and $K_p(T)$.

- 8.18** In the preceding problem the degree of dissociation ξ at atmospheric pressure and at various temperatures T is experimentally found to have the following values:

T ($^{\circ}\text{K}$)	ξ
1500	1.97×10^{-4}
1705	1.2×10^{-3}
2155	1.2×10^{-2}

What is the heat required to dissociate one mole of water vapor at 1 atmosphere into O_2 and H_2 at a temperature of 1700°K ?

- 8.19** The partition function for an ideal gas of molecules in a volume V can be written in the form

$$Z = \frac{1}{N!} (V\xi')^N$$

where $V\xi'$ is the partition function for a single molecule (involving its kinetic energy, plus internal energy if it is not monatomic) and ξ' depends only on the absolute temperature T .

When these molecules are condensed so as to form a liquid, the crudest approximation consists of treating the liquid as if the molecules still formed a gas of molecules moving independently, provided that (1) each molecule is assumed to have a *constant* potential energy $-\eta$ due to its average interaction with the rest of the molecules; and (2) each molecule is assumed free to move throughout a total volume Nv_0 , where v_0 is the (constant) volume available per molecule in the liquid phase.

(a) With these assumptions, write down the partition function for a liquid consisting of N_l molecules.

(b) Write down the chemical potential μ_v for N_v molecules of the vapor in a volume V_v at the temperature T . Treat it as an ideal gas.

(c) Write down the chemical potential μ_l for N_l molecules of liquid at the temperature T .

(d) By equating chemical potentials, find an expression relating the vapor pressure to the temperature T where the gas is in equilibrium with the liquid.

- (e) Calculate the molar entropy difference between gas and liquid in equilibrium at the same temperature and pressure. From this calculate the molar heat of evaporation L . Show that $L = N_A \eta$ if $\eta \gg kT$.
- (f) The boiling point T_b is that temperature where the vapor pressure is 1 atmosphere. Express the ratio L/RT_b in terms of v_0 and the volume v_e per molecule in the vapor phase at one atmosphere at the temperature T_b .
- (g) Estimate the order of magnitude of L/RT_b and show that it is a number of the order of 10 for all ordinary liquids. (This result is called "Trouton's rule.")
- (h) Compare this simple theory with experiment by looking up the densities and molecular weights of some liquids, computing L/T_b , and comparing with the experimental ratio of L/T_b . Data can be found in the "Handbook of Physics and Chemistry" (Chemical Rubber Publishing Company, Cleveland, Ohio). Try nitrogen and benzene, for example.

Quantum statistics of ideal gases

9

THIS CHAPTER will be devoted to a discussion of systems consisting of particles with negligible mutual interaction, i.e., of "ideal gases." But we shall now treat these systems from a completely quantum-mechanical point of view. This will allow us to discuss problems involving gases at low temperatures or high densities and to avoid the problems encountered in Sec. 7·3 in connection with the indistinguishability of the particles. It will also permit us to calculate unique values of entropies, to make absolute calculations of vapor pressures or chemical-equilibrium constants, and to treat distinctly nonclassical gases such as photons or conduction electrons in metals.

MAXWELL - BOLTZMANN, BOSE - EINSTEIN, AND FERMI - DIRAC STATISTICS

9 · 1 *Identical particles and symmetry requirements*

Consider a gas consisting of N identical structureless particles enclosed within a container of volume V . Let Q_i denote collectively all the coordinates of the i th particle (e.g., its three cartesian position coordinates and its spin coordinate, if any). Let s_i be an index labeling the possible quantum states of this single particle (e.g., each possible value of s_i corresponds to a specification of the three momentum components of the particle and of its direction of spin orientation; we postpone more detailed discussion to Sec. 9·9). The state of the whole gas is then described by the set of quantum numbers

$$\{s_1, s_2, \dots, s_N\} \quad (9 \cdot 1 \cdot 1)$$

which characterize the wave function Ψ of the gas in this state.

$$\Psi = \Psi_{\{s_1, \dots, s_N\}}(Q_1, Q_2, \dots, Q_N) \quad (9 \cdot 1 \cdot 2)$$

Let us now discuss the various cases of interest.

"Classical" case (Maxwell-Boltzmann statistics) In this case the particles are considered to be distinguishable, and any number of particles can be in the same single-particle state s . This "classical" description imposes no symmetry requirements on the wave function when two particles are interchanged. The particles are then said to obey "Maxwell-Boltzmann statistics" (abbreviated "MB statistics"). This description is *not* correct quantum mechanically, but is interesting for purposes of comparison.

Quantum mechanics The quantum-mechanical description is, of course, the one which is actually applicable. But when quantum mechanics is applied to a system of identical particles, it imposes definite symmetry requirements on the wave function (9·1·2) under interchange of any two identical particles. The net result is that one does not obtain a new state of the whole gas by simply interchanging two such particles. When counting the distinct possible states accessible to the whole gas, the particles must thus be considered as intrinsically indistinguishable. In enumerating these possible states of the gas, it does then not matter *which* particle is in *which* particle state, but only *how many* particles there are in each single-particle state s .

The symmetry requirements can be regarded as fundamental quantum-mechanical postulates* and are intimately connected with the spin of the particles. There are two possible cases which may arise: either (a) the particles have integral spin or (b) the particles have half-integral spin.

a. Particles with integral spin (Bose-Einstein statistics):

This is the case where each particle has a total spin angular momentum (measured in units of \hbar) which is integral, i.e., 0, 1, 2, . . . (examples might be He^4 atoms or photons). Then the fundamental quantum-mechanical symmetry requirement is that the total wave function Ψ be *symmetric* (i.e., that it remain unchanged) under interchange of any two particles (i.e., interchange of *both* their spatial and spin coordinates). In symbols,

$$\Psi(\dots Q_j \dots Q_i \dots) = \Psi(\dots Q_i \dots Q_j \dots) \quad (9 \cdot 1 \cdot 3)$$

(Here we have omitted the subscript $\{s_1, \dots, s_N\}$ in (9·1·2) for the sake of brevity.) Thus interchange of two particles does *not* lead to a new state of the whole gas. The particles must, therefore, be considered as genuinely indistinguishable in enumerating the distinct states of the gas. Note that there is no restriction on how many particles can be in any one single-particle state s . Particles satisfying the symmetry requirement (9·1·3) are said to obey "Bose-Einstein statistics" (abbreviated "BE statistics") and are sometimes called "bosons."

b. Particles with half-integral spin (Fermi-Dirac statistics):

This is applicable when each particle has a total spin angular momentum (measured in units of \hbar) which is half-integral, i.e., $\frac{1}{2}, \frac{3}{2}, \dots$ (examples might

* These postulates can, however, be derived (as was first done by Pauli) from a much more profound point of view which examines the requirements for a consistent description of the particles in terms of quantum field theory.

be electrons or He^3 atoms). Then the fundamental quantum-mechanical symmetry requirement is that the total wave function Ψ be *antisymmetric* (i.e., that it change sign) under interchange of any two particles. In symbols

$$\Psi(\dots Q_j \dots Q_i \dots) = -\Psi(\dots Q_i \dots Q_j \dots) \quad (9 \cdot 1 \cdot 4)$$

Once again, interchange of two particles does not lead to a new state of the gas. Hence the particles must again be considered as genuinely indistinguishable in enumerating the distinct states of the gas. But the change of sign in (9·1·4) does imply one additional consequence: Suppose that two particles i and j , both in the *same* single-particle state s , are interchanged. In that case one obviously has

$$\Psi(\dots Q_j \dots Q_i \dots) = \Psi(\dots Q_i \dots Q_j \dots) \quad (9 \cdot 1 \cdot 5)$$

But since the fundamental symmetry requirement (9·1·4) must also be valid, (9·1·4) and (9·1·5) together imply that

$$\Psi = 0 \quad \text{when particles } i \text{ and } j \text{ are in the same state } s \quad (9 \cdot 1 \cdot 6)$$

Thus in the Fermi-Dirac case there exists *no* state of the whole gas for which two or more particles are in the same single-particle state. This is the so-called "Pauli exclusion principle."* In enumerating the distinct states of the gas one must then keep in mind the restriction that there can never be more than one particle in any given single-particle state. Particles satisfying the antisymmetry requirement (9·1·4) are said to obey Fermi-Dirac statistics (abbreviated "FD statistics") and are sometimes called "fermions."

Illustration A very simple example should help to make these general ideas much clearer. Consider a "gas" of only two particles; call them A and B . Assume that each particle can be in one of three possible quantum states, $s = 1, 2, 3$. Let us enumerate the possible states of the whole gas. This is the same as asking in how many distinct ways one can put two particles (A and B) into three single-particle states (labeled 1, 2, 3).

Maxwell-Boltzmann statistics: The particles are to be considered distinguishable. Any number of particles can be in any one state.

1	2	3
AB
...	AB	...
...	...	AB
A	B	...
B	A	...
A	...	B
B	...	A
...	A	B
...	B	A

* This principle should be familiar since it applies to the important case of electrons (which have spin $\frac{1}{2}$) and accounts for the periodic table of the elements.

Each of the two particles can be placed in any one of the three states. Hence there exist a total of $3^2 = 9$ possible states for the whole gas.

Bose-Einstein statistics: The particles are to be considered *indistinguishable*. Any number of particles can be in any one state. The indistinguishability implies that $B = A$, so that the three states in the MB case which differed only in interchange of A and B are now no longer to be counted as distinct. The enumeration is then as follows:

1	2	3
AA
...	AA	...
...	...	AA
A	A	...
A	...	A
...	A	A

There are now three distinct ways of placing the particles in the same state. There are three distinct ways of placing the particles in different states. Hence there exist a total of $3 + 3 = 6$ possible states for the whole gas.

Fermi-Dirac statistics: The particles are to be considered as *indistinguishable*. No more than one particle can be in any one state. The three states in the BE case where two particles were in the same state must be eliminated in this case. One is thus left with the following enumeration:

1	2	3
A	A	...
A	...	A
...	A	A

There exist now only a total of 3 possible states for the whole gas.

This example shows one further qualitative feature of interest. Let

$$\xi \equiv \frac{\text{probability that the two particles are found in the same state}}{\text{probability that the two particles are found in different states}}$$

Then we have for the three cases

$$\begin{aligned}\xi_{\text{MB}} &= \frac{3}{6} = \frac{1}{2} \\ \xi_{\text{BE}} &= \frac{3}{3} = 1 \\ \xi_{\text{FD}} &= \frac{0}{3} = 0\end{aligned}$$

Thus in the BE case there is a greater relative tendency for particles to bunch together in the same state than in classical statistics. On the other hand, in the FD case there is a greater relative tendency for particles to remain apart in different states than there is in classical statistics.

Discussion in terms of wave functions The same simple example can be discussed equivalently in terms of the possible wave functions for the whole gas. Let

$$\psi_s(Q) = \text{the one-particle wave function for a single particle (with coordinate } Q \text{) in state } s$$

As before, let Ψ be the wave function for the whole gas. Since the particles are noninteracting, Ψ can be written as a simple product of single-particle wave functions, or of proper linear combinations thereof. Let us again discuss the cases in turn.

Maxwell-Boltzmann statistics: There is no particular symmetry requirement on Ψ under particle interchange. Apart from normalization, a complete set of wave functions Ψ for the gas are then the $3 \times 3 = 9$ functions of the form

$$\psi_i(Q_A)\psi_j(Q_B)$$

where $i = 1, 2, 3$ and $j = 1, 2, 3$.

Bose-Einstein statistics: Here Ψ must be symmetric under interchange of the two particles. From the nine wave functions listed above one can construct only six symmetric ones. A complete (unnormalized) set of distinct wave functions are then the three functions of the form

$$\psi_i(Q_A)\psi_i(Q_B)$$

where $i = 1, 2, 3$ and the three functions of the form

$$\psi_i(Q_A)\psi_j(Q_B) + \psi_j(Q_B)\psi_i(Q_A)$$

where $j > i$; again $i = 1, 2, 3$ and $j = 1, 2, 3$.

Fermi-Dirac statistics: Here Ψ must be antisymmetric under interchange of the two particles. From the nine wave functions listed under the MB case one can construct only three antisymmetric ones. A complete (unnormalized) set of distinct wave functions are then the three functions of the form

$$\psi_i(Q_A)\psi_j(Q_B) - \psi_j(Q_B)\psi_i(Q_A)$$

where $j > i$; again $i = 1, 2, 3$ and $j = 1, 2, 3$.

9 · 2 Formulation of the statistical problem

We consider a gas of identical particles in a volume V in equilibrium at the temperature T . We shall use the following notation:

Label the possible quantum states of a single particle by r (or s).

Denote the energy of a particle in state r by ϵ_r .

Denote the number of particles in state r by n_r .

Label the possible quantum states of the whole gas by R .

The assumption of negligibly small interaction between the particles allows us to write for the total energy of the gas, when it is in some state R where there are n_1 particles in state $r = 1$, n_2 particles in state $r = 2$, etc., the additive expression

$$E_R = n_1 \epsilon_1 + n_2 \epsilon_2 + n_3 \epsilon_3 + \dots = \sum_r n_r \epsilon_r \quad (9 \cdot 2 \cdot 1)$$

where the sum extends over all the possible states r of a particle. Furthermore, if the total number of particles in the gas is known to be N , one must have

$$\sum_r n_r = N \quad (9 \cdot 2 \cdot 2)$$

In order to calculate the thermodynamic functions of the gas (e.g., its entropy), it is necessary to calculate its partition function

$$Z = \sum_R e^{-\beta E_R} = \sum_R e^{-\beta(n_1 \epsilon_1 + n_2 \epsilon_2 + \dots)} \quad (9 \cdot 2 \cdot 3)$$

Here the sum is over all the possible states R of the whole gas, i.e., essentially over all the various possible values of the numbers n_1, n_2, n_3, \dots .

Since $\exp[-\beta(n_1 \epsilon_1 + n_2 \epsilon_2 + \dots)]$ is the relative probability of finding the gas in a particular state where there are n_1 particles in state 1, n_2 particles in state 2, etc., one can write for the mean number of particles in a state s

$$\bar{n}_s = \frac{\sum_R n_s e^{-\beta(n_1 \epsilon_1 + n_2 \epsilon_2 + \dots)}}{\sum_R e^{-\beta(n_1 \epsilon_1 + n_2 \epsilon_2 + \dots)}} \quad (9 \cdot 2 \cdot 4)$$

Hence $\bar{n}_s = \frac{1}{Z} \sum_R \left(-\frac{1}{\beta} \frac{\partial}{\partial \epsilon_s} \right) e^{-\beta(n_1 \epsilon_1 + n_2 \epsilon_2 + \dots)} = \frac{-1}{\beta Z} \frac{\partial Z}{\partial \epsilon_s}$

or

$$\blacktriangleright \quad \bar{n}_s = -\frac{1}{\beta} \frac{\partial \ln Z}{\partial \epsilon_s} \quad (9 \cdot 2 \cdot 5)$$

Thus the mean number of particles in a given single-particle state s can also be expressed in terms of the partition function Z .

Calculation of the dispersion One can similarly write down an expression for the dispersion of the number of particles in state s . One can use the general relation

$$\overline{(\Delta n_s)^2} = \overline{(n_s - \bar{n}_s)^2} = \overline{n_s^2} - \bar{n}_s^2 \quad (9 \cdot 2 \cdot 6)$$

But for $\overline{n_s^2}$ one can write, by definition,

$$\overline{n_s^2} = \frac{\sum_R n_s^2 e^{-\beta(n_1 \epsilon_1 + n_2 \epsilon_2 + \dots)}}{\sum_R e^{-\beta(n_1 \epsilon_1 + n_2 \epsilon_2 + \dots)}} \quad (9 \cdot 2 \cdot 7)$$

Hence $\overline{n_s^2} = \frac{1}{Z} \sum_R \left(-\frac{1}{\beta} \frac{\partial}{\partial \epsilon_s} \right) \left(-\frac{1}{\beta} \frac{\partial}{\partial \epsilon_s} \right) e^{-\beta(n_1 \epsilon_1 + n_2 \epsilon_2 + \dots)} = \frac{1}{Z} \left(-\frac{1}{\beta} \frac{\partial}{\partial \epsilon_s} \right)^2 Z$

or

$$\overline{n_s^2} = \frac{1}{\beta^2 Z} \frac{\partial^2 Z}{\partial \epsilon_s^2} \quad (9 \cdot 2 \cdot 8)$$

This can be put into a more convenient form involving \bar{n}_s in (9·2·5). Thus

$$\overline{n_s^2} = \frac{1}{\beta^2} \left[\frac{\partial}{\partial \epsilon_s} \left(\frac{1}{Z} \frac{\partial Z}{\partial \epsilon_s} \right) + \frac{1}{Z^2} \left(\frac{\partial Z}{\partial \epsilon_s} \right)^2 \right] = \frac{1}{\beta^2} \left[\frac{\partial}{\partial \epsilon_s} \left(\frac{\partial \ln Z}{\partial \epsilon_s} \right) + \beta^2 \bar{n}_s^2 \right]$$

Thus (9·2·6) becomes

$$\overline{(\Delta n_s)^2} = \frac{1}{\beta^2} \frac{\partial^2 \ln Z}{\partial \epsilon_s^2} \quad (9 \cdot 2 \cdot 9)$$

or by (9·2·5),

$$\overline{(\Delta n_s)^2} = - \frac{1}{\beta} \frac{\partial \bar{n}_s}{\partial \epsilon_s} \quad (9 \cdot 2 \cdot 10)$$

Calculation of all physical quantities of interest thus requires simply the evaluation of the partition function (9·2·3). Let us now be very specific about what we mean by the sum over all possible states R of the gas. In accordance with the discussion of Sec 9·1 we mean the following:

Maxwell-Boltzmann statistics: Here one must sum over all possible numbers of particles in each state, i.e., over all values

$$n_r = 0, 1, 2, 3, \dots \quad \text{for each } r \quad (9 \cdot 2 \cdot 11)$$

subject to the restriction (9·2·2) of a fixed total number of particles

$$\sum_r n_r = N \quad (9 \cdot 2 \cdot 12)$$

But the particles have also to be considered as *distinguishable*. Thus any permutation of two particles in different states must be counted as a distinct state of the whole gas even though the numbers $\{n_1, n_2, n_3, \dots\}$ are left unchanged. This is so because it is not enough to specify how many particles are in each single-particle state, but it is necessary to specify *which* particular particle is in *which* state.

Bose-Einstein and photon statistics: Here the particles are to be considered as indistinguishable, so that mere specification of the numbers $\{n_1, n_2, n_3, \dots\}$ is enough to specify the state of the gas. Thus it is necessary only to sum over all possible numbers of particles in each single-particle state, i.e., over all possible values

$$n_r = 0, 1, 2, 3, \dots \quad \text{for each } r \quad (9 \cdot 2 \cdot 13)$$

If the total number of particles is fixed, these numbers must satisfy only the restriction (9·2·2)

$$\sum_r n_r = N \quad (9 \cdot 2 \cdot 14)$$

A simpler special case is that where there is *no* requirement fixing the total number of particles. This is the case, for example, when one considers the particles to be photons enclosed in a container of volume V , since the photons can be readily absorbed and emitted by the walls. There is then no equation of constraint (9·2·14) to be satisfied, and one obtains the special case of "photon statistics."

Fermi-Dirac statistics: Here the particles are again to be considered as indistinguishable, so that mere specification of the numbers $\{n_1, n_2, n_3, \dots\}$ is enough to specify the state of the gas. Thus it is necessary only to sum over all possible numbers of particles in each single-particle state, remembering that there can be no more than one particle in any one such state; i.e., one has to sum over the two possible values

$$n_r = 0, 1 \quad \text{for each } r \quad (9 \cdot 2 \cdot 15)$$

If the total number of particles is fixed, these numbers must satisfy only the restriction (9·2·2)

$$\sum_r n_r = N \quad (9 \cdot 2 \cdot 16)$$

9 · 3 *The quantum distribution functions*

Before turning to a systematic calculation of the partition functions in the various cases of interest, we shall devote this section to a simple discussion of the essential features of the quantum theory of ideal gases. We begin by noting that there is a profound difference between gases obeying BE statistics and those obeying FD statistics. This difference becomes most striking in the limit as $T \rightarrow 0$, when the gas as a whole is in its state of lowest energy.

Consider a gas consisting of a fixed number N of particles, and suppose that the state of lowest energy of a single particle has an energy ϵ_1 . (This corresponds to a state where the particle has essentially zero momentum.) In the case of BE statistics, where there is no restriction on how many particles can be placed in any one single-particle state, the lowest energy of the whole gas is then obtained if *all* the N particles of the gas are put into their lowest-lying state of energy ϵ_1 (e.g., all particles are in their state of zero momentum). This then describes the situation at $T = 0$.

But in the case of FD statistics one cannot put more than one particle into any one single-particle state. If one is interested in obtaining the lowest energy of the whole gas, one is then forced to populate single-particle states of increasing energy; i.e., one can start from the state of lowest energy ϵ_1 and must then fill up single-particle states of successively higher energies, one at a time, until all the N particles have been accommodated. The net result is that, even when $T = 0$ and the gas as a whole is in its state of lowest possible energy, there are particles in the gas which have a very high energy compared

to ϵ_1 ; similarly, the gas as a whole has an energy considerably greater than the energy $N\epsilon_1$ which it would have if the particles obeyed BE statistics. The Pauli exclusion principle thus has very pronounced consequences.

Let us now consider the case of arbitrary temperature T and calculate, for the several cases of interest, the mean number of particles \bar{n}_s in a particular state s . We can proceed directly from the expression (9.2.4) for this mean value, i.e.,

$$\bar{n}_s = \frac{\sum_{n_1, n_2, \dots} n_s e^{-\beta(n_1\epsilon_1 + n_2\epsilon_2 + \dots + n_s\epsilon_s + \dots)}}{\sum_{n_1, n_2, \dots} e^{-\beta(n_1\epsilon_1 + n_2\epsilon_2 + \dots + n_s\epsilon_s + \dots)}} \quad (9.3.1)$$

By summing first over all possible values of n_s , using the multiplicative property of the exponential function, and rearranging the order of summation, (9.3.1) can also be written in the form

$$\bar{n}_s = \frac{\sum_{n_s} n_s e^{-\beta n_s \epsilon_s} \sum_{n_1, n_2, \dots}^{(s)} e^{-\beta(n_1\epsilon_1 + n_2\epsilon_2 + \dots)}}{\sum_{n_s} e^{-\beta n_s \epsilon_s} \sum_{n_1, n_2, \dots}^{(s)} e^{-\beta(n_1\epsilon_1 + n_2\epsilon_2 + \dots)}} \quad (9.3.2)$$

Here the last sums in the numerator and denominator omit from consideration the particular state s (this is indicated by the superscript s on the summation symbol).

Photon statistics This is the case of BE statistics with an unspecified total number of particles. In accordance with the discussion of the last section, the numbers n_1, n_2, \dots assume here all values $n_r = 0, 1, 2, 3, \dots$ for each r , without any further restriction. The sums $\Sigma^{(s)}$ in the numerator and denominator of (9.3.2) are then identical and cancel. Thus one is left simply with

$$\bar{n}_s = \frac{\sum_{n_s} n_s e^{-\beta n_s \epsilon_s}}{\sum_{n_s} e^{-\beta n_s \epsilon_s}} \quad (9.3.3)$$

The rest of the calculation is straightforward. The result (9.3.3) becomes

$$\bar{n}_s = \frac{(-1/\beta)(\partial/\partial\epsilon_s)\Sigma e^{-\beta n_s \epsilon_s}}{\Sigma e^{-\beta n_s \epsilon_s}} = -\frac{1}{\beta} \frac{\partial}{\partial\epsilon_s} \ln (\Sigma e^{-\beta n_s \epsilon_s}) \quad (9.3.4)$$

But the last sum is just an infinite geometric series which can be summed. Thus

$$\sum_{n_s=0}^{\infty} e^{-\beta n_s \epsilon_s} = 1 + e^{-\beta \epsilon_s} + e^{-2\beta \epsilon_s} + \dots = \frac{1}{1 - e^{-\beta \epsilon_s}}$$

Hence (9.3.4) gives

$$\bar{n}_s = \frac{1}{\beta} \frac{\partial}{\partial\epsilon_s} \ln (1 - e^{-\beta \epsilon_s}) = \frac{e^{-\beta \epsilon_s}}{1 - e^{-\beta \epsilon_s}}$$

or

$$\blacktriangleright \quad \bar{n}_s = \frac{1}{e^{\beta \epsilon_s} - 1} \quad (9 \cdot 3 \cdot 5)$$

This is called the "Planck distribution."

Fermi-Dirac statistics Let us now turn to cases where the total number N of particles is fixed. This restriction makes the calculation slightly more complicated. We discuss first the case of FD statistics, since it is somewhat simpler. In accordance with the discussion of Sec. 9·2, the sums in (9·3·2) range here over all values of the numbers n_1, n_2, \dots such that $n_r = 0$ and 1 for each r ; but these numbers must always satisfy the restriction

$$\sum_r n_r = N \quad (9 \cdot 3 \cdot 6)$$

This restriction implies, for example, that if one particle is in state s , the sum $\Sigma^{(s)}$ in (9·3·2) extends only over the remaining $(N - 1)$ particles which can be put into the states other than s . Let us then introduce for the sum $\Sigma^{(s)}$, extended over all states except s , the convenient abbreviation

$$Z_s(N) \equiv \sum_{n_1, n_2, \dots}^{(s)} e^{-\beta(n_1 \epsilon_1 + n_2 \epsilon_2 + \dots)} \quad (9 \cdot 3 \cdot 7)$$

if N particles are to be distributed over these remaining states, i.e., if

$$\sum_r^{(s)} n_r = N \quad (\text{state } s \text{ omitted from this sum})$$

By performing explicitly the sum over $n_s = 0$ and 1, the expression (9·3·2) becomes then

$$\bar{n}_s = \frac{0 + e^{-\beta \epsilon_s} Z_s(N - 1)}{Z_s(N) + e^{-\beta \epsilon_s} Z_s(N - 1)} \quad (9 \cdot 3 \cdot 8)$$

or
$$\bar{n}_s = \frac{1}{[Z_s(N)/Z_s(N - 1)] e^{\beta \epsilon_s} + 1} \quad (9 \cdot 3 \cdot 9)$$

This can be simplified by relating $Z_s(N - 1)$ to $Z_s(N)$. Thus one can write quite generally, if $\Delta N \ll N$,

$$\ln Z_s(N - \Delta N) = \ln Z_s(N) - \frac{\partial \ln Z_s}{\partial N} \Delta N = \ln Z_s(N) - \alpha_s \Delta N$$

or
$$Z_s(N - \Delta N) = Z_s(N) e^{-\alpha_s \Delta N} \quad (9 \cdot 3 \cdot 10)$$

where
$$\alpha_s \equiv \frac{\partial \ln Z_s}{\partial N} \quad (9 \cdot 3 \cdot 11)$$

But since $Z_s(N)$ is a sum over very many states, one expects that the variation of its logarithm with the total number of particles N should be very insensitive as to which particular state s has been omitted from the sum (9·3·7). Let us then introduce the approximation (whose validity can be verified later) that α_s is independent of s , so that one can write simply

$$\alpha_s = \alpha \quad (9 \cdot 3 \cdot 12)$$

for all s . The derivative (9.3.11) can then also be expressed approximately in terms of the derivative of the full partition function $Z(N)$ (over all states) which occurs in the denominator of (9.3.1) or (9.3.8); i.e.,

$$\alpha = \frac{\partial \ln Z}{\partial N} \quad (9.3.13)$$

Using (9.3.10) with $\Delta N = 1$ and the approximation (9.3.12), the result (9.3.9) becomes then

► $\bar{n}_s = \frac{1}{e^{\alpha+\beta\epsilon_s} + 1} \quad (9.3.14)$

This is called the "Fermi-Dirac distribution."

The parameter α in (9.3.14) can be determined from the condition (9.3.6), which demands that the mean values must satisfy the relation

$$\sum_r \bar{n}_r = N \quad (9.3.15)$$

or $\sum_r \frac{1}{e^{\alpha+\beta\epsilon_r} + 1} = N \quad (9.3.16)$

Note that since the free energy $F = -kT \ln Z$, the relation (9.3.13) is equivalent to

$$\alpha = -\frac{1}{kT} \frac{\partial F}{\partial N} = -\frac{\mu}{kT} = -\beta\mu \quad (9.3.17)$$

where μ is the chemical potential per particle defined in (8.7.10). The result (9.3.14) can thus also be written in the form

$$\bar{n}_s = \frac{1}{e^{\beta(\epsilon_s-\mu)} + 1} \quad (9.3.18)$$

Note that $\bar{n}_s \rightarrow 0$ if ϵ_s becomes large enough. On the other hand, since the denominator in (9.3.14) can never become less than unity no matter how small ϵ_s becomes, it follows that $\bar{n}_s \leq 1$. Hence

$$0 \leq \bar{n}_s \leq 1$$

a relation which reflects properly the requirement imposed by the Pauli exclusion principle

Remark concerning the validity of the approximation The partition function given by the denominator of (9.3.1) or (9.3.8) is related to $Z_s(N)$ by

$$Z(N) = Z_s(N) + e^{-\beta\epsilon_s} Z_s(N-1) = Z_s(N)(1 + e^{-\alpha-\beta\epsilon_s})$$

or $\ln Z = \ln Z_s + \ln(1 + e^{-\alpha-\beta\epsilon_s})$

where we have used (9.3.10) and (9.3.12). Hence

$$\frac{\partial \ln Z}{\partial N} = \frac{\partial \ln Z_s}{\partial N} - \frac{e^{-\alpha-\beta\epsilon_s}}{1 + e^{-\alpha-\beta\epsilon_s}} \frac{\partial \alpha}{\partial N}$$

or $\alpha = \alpha_s - \bar{n}_s \frac{\partial \alpha}{\partial N}$

The assumption (9·3·12) is then satisfied if

$$\frac{\partial \alpha}{\partial N} \bar{n}_s \ll \alpha \quad (9\cdot3\cdot19)$$

or, for FD statistics where $\bar{n}_s < 1$, if $\partial\alpha/\partial N \ll \alpha$; i.e., if the number of particles N is large enough so that the chemical potential does not change appreciably upon addition of one more particle to the system.

Bose-Einstein statistics The discussion here is very similar to that just given in the case of FD statistics. Here the sums in (9·3·2) range over all values of the numbers n_1, n_2, \dots such that $n_r = 0, 1, 2, 3, \dots$ for each r ; but the situation differs from the case of photons because these numbers must always satisfy the restriction (9·3·6) of a fixed total number N of particles. Performing explicitly the sum over n_s , Eq. (9·3·2) then becomes

$$\bar{n}_s = \frac{0 + e^{-\beta \epsilon_s} Z_s(N-1) + 2e^{-2\beta \epsilon_s} Z_s(N-2) + \dots}{Z_s(N) + e^{-\beta \epsilon_s} Z_s(N-1) + e^{-2\beta \epsilon_s} Z_s(N-2) + \dots} \quad (9\cdot3\cdot20)$$

where $Z_s(N)$ is defined as in (9·3·7). Using (9·3·10) and the approximation (9·3·12), the result (9·3·20) becomes

$$\begin{aligned} \bar{n}_s &= \frac{Z_s(N)[0 + e^{-\beta \epsilon_s} e^{-\alpha} + 2e^{-2\beta \epsilon_s} e^{-2\alpha} + \dots]}{Z_s(N)[1 + e^{-\beta \epsilon_s} e^{-\alpha} + e^{-2\beta \epsilon_s} e^{-2\alpha} + \dots]} \\ \text{or } \bar{n}_s &= \frac{\sum_s n_s e^{-n_s(\alpha+\beta \epsilon_s)}}{\sum_s e^{-n_s(\alpha+\beta \epsilon_s)}} \end{aligned} \quad (9\cdot3\cdot21)$$

But this simple expression is similar to (9·3·3), except that $\beta \epsilon_s$ in that expression is replaced by $(\alpha + \beta \epsilon_s)$. The remainder of the calculation is then identical to that leading to (9·3·4) and yields therefore

$$\bar{n}_s = \frac{1}{e^{\alpha+\beta \epsilon_s} - 1} \quad (9\cdot3\cdot22)$$

This is called the "Bose-Einstein distribution." Note that \bar{n}_s can become very large in this case. The parameter α can again be determined by the condition (9·3·15), i.e., by the relation

$$\sum_r \frac{1}{e^{\alpha+\beta \epsilon_r} - 1} = N \quad (9\cdot3\cdot23)$$

It is again related to the chemical potential μ by the relation $\alpha = -\beta \mu$ of (9·3·17), so that (9·3·22) can also be written in the form

$$\bar{n}_s = \frac{1}{e^{\beta(\epsilon_s-\mu)} - 1} \quad (9\cdot3\cdot24)$$

In the case of photons, the sums are to be performed without any restriction as to the total number N of particles, so that $Z(N)$ [or $Z_s(N)$] does not depend on N . Thus $\alpha = 0$ by (9·3·13), and the Bose-Einstein distribution (9·3·22) reduces properly to the special case of the Planck distribution (9·3·5).

Remark In the case of photons (or other particles whose total number is not fixed) ϵ_s denotes the unambiguously defined energy necessary to create one particle in state s (e.g., $\epsilon_s = \hbar\omega_s$ if ω_s is the angular frequency of the photon). Suppose that the energy scale is shifted by an arbitrary constant η , so that the ground state of the photon gas (the situation when $n_1 = n_2 = n_3 = \dots = 0$) has energy η instead of zero energy. Then the energy of the gas in a particular state becomes $E = \sum n_s \epsilon_s + \eta$. But the constant η cancels in (9.3.1), so that the Planck distribution (9.3.5) is properly unaffected.

In the case of ordinary gases with a fixed number N of particles, ϵ_s denotes the energy level of a particle in state s . Suppose that the energy scale is shifted by an arbitrary constant. Then all single-particle energy levels are shifted by the same constant η' and the energy of all states of the whole gas is shifted by the constant $\eta = N\eta'$. Once again this additive constant cancels in (9.3.1); thus the FD and BE distributions (9.3.18) and (9.3.24) are properly unaffected (the chemical potential μ being also shifted by η').

This completes the discussion of the essential features of the quantum statistics of ideal gases. It is, however, worth looking at the various cases in greater detail with the aim of calculating not only the distribution functions \tilde{n}_s , but also thermodynamic functions (e.g., the entropy) and the magnitude of fluctuations in the number of particles in a given state. We shall, therefore, devote the next few sections to calculate systematically the partition function Z for each case of interest, i.e., to find an explicit expression for Z in terms of the energy levels of a *single* particle. The remainder of the calculation will then involve only the simple problem of finding explicitly the energy levels of a single particle.

9.4 Maxwell-Boltzmann statistics

For purposes of comparison, it is instructive to deal first with the strictly classical case of Maxwell-Boltzmann statistics. Here the partition function is

$$Z = \sum_R e^{-\beta(n_1 \epsilon_1 + n_2 \epsilon_2 + \dots)} \quad (9.4.1)$$

where the sum is to be evaluated, as described at the end of Sec. 9.2, by summing over all states R of the gas, i.e., by summing over all possible values of the numbers n_s and taking into account the *distinguishability* of the particles. If there is a total of N molecules, there are, for given values of $\{n_1, n_2, \dots\}$,

$$\frac{N!}{n_1! n_2! \dots}$$

possible ways in which the particles can be put into the given single-particle states, so that there are n_1 particles in state 1, n_2 particles in state 2, etc. By virtue of the distinguishability of the particles, each of these possible arrange-

ments corresponds then to a *distinct* state for the whole gas. Hence (9·4·1) can be written explicitly as

$$Z = \sum_{n_1, n_2, \dots} \frac{N!}{n_1! n_2! \dots} e^{-(\epsilon_1 + \epsilon_2 + \dots)} \beta \quad (9\cdot4\cdot2)$$

where one sums over all values $n_r = 0, 1, 2, \dots$ for each r , subject to the restriction

$$\sum_r n_r = N \quad (9\cdot4\cdot3).$$

But (9·4·2) can be written

$$Z = \sum_{n_1, n_2, \dots} \frac{N!}{n_1! n_2! \dots} (e^{-\beta \epsilon_1})^{n_1} (e^{-\beta \epsilon_2})^{n_2} \dots$$

which, by virtue of (9·4·3), is just the result of expanding a polynomial. Thus

$$Z = (e^{-\beta \epsilon_1} + e^{-\beta \epsilon_2} + \dots)^N$$

or

$$\blacktriangleright \quad \ln Z = N \ln \left(\sum_r e^{-\beta \epsilon_r} \right) \quad (9\cdot4\cdot4)$$

where the argument of the logarithm is simply the partition function for a single particle.

Alternative method One may equally well write the partition function of the whole gas in the form

$$Z = \sum_{r_1, r_2, \dots} \exp [-\beta(\epsilon_{r_1} + \epsilon_{r_2} + \dots + \epsilon_{r_N})] \quad (9\cdot4\cdot5)$$

where the summation is now over all the possible *states* of *each* individual particle. Clearly, this way of summing considers particles as distinguishable and produces distinct terms in the sum when [particle 1 is in state r_1 and particle 2 is in state r_2] and when [particle 2 is in state r_1 and particle 1 is in state r_2]. Now (9·4·5) immediately factors to give

$$\begin{aligned} Z &= \sum_{r_1, r_2, \dots} \exp (-\beta \epsilon_{r_1}) \exp (-\beta \epsilon_{r_2}) \dots \\ &= \left[\sum_{r_1} \exp (-\beta \epsilon_{r_1}) \right] \left[\sum_{r_2} \exp (-\beta \epsilon_{r_2}) \right] \dots \\ \text{or} \quad Z &= \left[\sum_{r_1} \exp (-\beta \epsilon_{r_1}) \right]^N \end{aligned} \quad (9\cdot4\cdot6)$$

Thus one regains the result (9·4·4).

By applying (9·2·5) to the partition function (9·4·4), one obtains, by differentiating with respect to the one term involving ϵ_s ,

$$\bar{n}_s = -\frac{1}{\beta} \frac{\partial \ln Z}{\partial \epsilon_s} = -\frac{1}{\beta} N \frac{-\beta e^{-\beta \epsilon_s}}{\sum_r e^{-\beta \epsilon_r}}$$

or

$$\blacktriangleright \quad \bar{n}_s = N \frac{e^{-\beta \epsilon_s}}{\sum_r e^{-\beta \epsilon_r}} \quad (9 \cdot 4 \cdot 7)$$

This is called the "Maxwell-Boltzmann distribution." It is, of course, just the result we encountered previously in a classical approach where we applied the canonical distribution to a single particle.

Calculation of the dispersion By combining the general result (9·2·10) with (9·4·7) one obtains

$$\begin{aligned} \overline{(\Delta n_s)^2} &= -\frac{1}{\beta} \frac{\partial \bar{n}_s}{\partial \epsilon_s} = -\frac{N}{\beta} \left[\frac{-\beta e^{-\beta \epsilon_s}}{\sum e^{-\beta \epsilon_r}} - \frac{-\beta e^{-\beta \epsilon_s} e^{-\beta \epsilon_s}}{(\sum e^{-\beta \epsilon_r})^2} \right] \\ &= \bar{n}_s - \frac{\bar{n}_s^2}{N} \end{aligned}$$

or $\overline{(\Delta n_s)^2} = \bar{n}_s \left(1 - \frac{\bar{n}_s}{N} \right) \approx \bar{n}_s \quad (9 \cdot 4 \cdot 8)$

This last step follows since $\bar{n}_s \ll N$ unless the temperature T is exceedingly low. The relative dispersion is then

$$\blacktriangleright \quad \frac{\overline{(\Delta n_s)^2}}{\bar{n}_s^2} = \frac{1}{\bar{n}_s} \quad (9 \cdot 4 \cdot 9)$$

9 · 5 Photon statistics

The partition function is given by

$$Z = \sum_R e^{-\beta(n_1 \epsilon_1 + n_2 \epsilon_2 + \dots)} \quad (9 \cdot 5 \cdot 1)$$

where, in accordance with the discussion at the end of Sec. 9·2, the summation is simply over all values $n_r = 0, 1, 2, 3, \dots$ for each r , without further restriction. Thus (9·5·1) becomes explicitly

$$\begin{aligned} Z &= \sum_{n_1, n_2, \dots} e^{-\beta n_1 \epsilon_1} e^{-\beta n_2 \epsilon_2} e^{-\beta n_3 \epsilon_3} \dots \\ \text{or} \quad Z &= \left(\sum_{n_1=0}^{\infty} e^{-\beta n_1 \epsilon_1} \right) \left(\sum_{n_2=0}^{\infty} e^{-\beta n_2 \epsilon_2} \right) \left(\sum_{n_3=0}^{\infty} e^{-\beta n_3 \epsilon_3} \right) \dots \quad (9 \cdot 5 \cdot 2) \end{aligned}$$

But each sum is just an infinite geometric series whose first term is 1 and where the ratio between successive terms is $e^{-\beta \epsilon_r}$. It can thus be immediately summed. Hence (9·5·2) becomes

$$Z = \left(\frac{1}{1 - e^{-\beta \epsilon_1}} \right) \left(\frac{1}{1 - e^{-\beta \epsilon_2}} \right) \left(\frac{1}{1 - e^{-\beta \epsilon_3}} \right) \dots$$

or

$$\blacktriangleright \quad \ln Z = - \sum_r \ln (1 - e^{-\beta \epsilon_r}) \quad (9 \cdot 5 \cdot 3)$$

By applying (9·2·5) to (9·5·3), differentiation with respect to the one term involving ϵ_s yields

$$\bar{n}_s = -\frac{1}{\beta} \frac{\partial \ln Z}{\partial \epsilon_s} = \frac{e^{-\beta \epsilon_s}}{1 - e^{-\beta \epsilon_s}}$$

or

► $\bar{n}_s = \frac{1}{e^{\beta \epsilon_s} - 1}$ (9·5·4)

Thus we regain the Planck distribution previously derived in (9·3·5).

Calculation of the dispersion The dispersion in n_s can be calculated by applying (9·2·10) to (9·5·4). Thus

$$\overline{(\Delta n_s)^2} = -\frac{1}{\beta} \frac{\partial \bar{n}_s}{\partial \epsilon} = \frac{e^{\beta \epsilon_s}}{(e^{\beta \epsilon_s} - 1)^2}$$

One can use (9·5·4) to write this in terms of \bar{n}_s . Thus

$$\overline{(\Delta n_s)^2} = \frac{(e^{\beta \epsilon_s} - 1) + 1}{(e^{\beta \epsilon_s} - 1)^2} = \bar{n}_s + \bar{n}_s^2$$

Hence

$$\overline{(\Delta n_s)^2} = \bar{n}_s(1 + \bar{n}_s) \quad (9·5·5)$$

or

► $\frac{\overline{(\Delta n_s)^2}}{\bar{n}_s^2} = \frac{1}{\bar{n}_s} + 1 \quad (9·5·6)$

Note that this dispersion is greater than in the MB case of Eq. (9·4·8). In dealing with photons, therefore, the relative dispersion does *not* become arbitrarily small even if $\bar{n}_s \gg 1$.

9 · 6 Bose-Einstein statistics

The partition function is again given by

$$Z = \sum_R e^{-\beta(n_1 \epsilon_1 + n_2 \epsilon_2 + \dots)} \quad (9·6·1)$$

where, in accordance with the discussion of Sec. 9·2, the summation is over all values

$$n_r = 0, 1, 2, \dots \quad \text{for each } r \quad (9·6·2)$$

Unlike the photon case, however, these numbers must now satisfy the restrictive condition

$$\sum_r n_r = N \quad (9·6·3)$$

where N is the total number of particles in the gas. If it were not for the equation of constraint (9·6·3), the sum (9·6·1) could be easily evaluated just as in the last section. But the condition (9·6·3) introduces a complication.

There are various ways of handling the problem presented by the condition (9·6·3). Let us use an approximation method similar to that described in Sec. 6·8. As a result of (9·6·3), Z depends on the total number N of particles in the system. If the number of particles were N' instead of N , the partition function would have some other value $Z(N')$. Indeed, since there are so many terms in the sum (9·6·1), $Z(N')$ is a very rapidly increasing function of N' . But, by virtue of (9·6·3), we are interested only in the value of Z for $N' = N$. We can, however, exploit the rapidly increasing property of $Z(N')$ by noting that multiplication by the rapidly decreasing function $e^{-\alpha N'}$ produces a function $Z(N')e^{-\alpha N'}$ with a very sharp maximum which can be made to occur at the value $N' = N$ by proper choice of the positive parameter α . A sum of this function over all possible numbers N' thus selects only those terms of interest near $N' = N$, i.e.,

$$\sum_{N'} Z(N') e^{-\alpha N'} = Z(N) e^{-\alpha N} \Delta^* N' \quad (9 \cdot 6 \cdot 4)$$

where the right side is just the maximum value of the summand multiplied by the width $\Delta^* N'$ of its maximum (where $\Delta^* N' \ll N$).

Let us introduce the abbreviation

$$\blacktriangleright \quad Z \equiv \sum_{N'} Z(N') e^{-\alpha N'} \quad (9 \cdot 6 \cdot 5)$$

Taking the logarithm of (9·6·4) one then obtains to an excellent approximation

$$\blacktriangleright \quad \ln Z(N) = \alpha N + \ln Z \quad (9 \cdot 6 \cdot 6)$$

where we have neglected the term $\ln(\Delta^* N')$ which is utterly negligible compared to the other terms which are of order N . Here the sum (9·6·5) is easily performed, since it extends over all possible numbers without any restriction. (The quantity Z is called a "grand partition function.")

Let us evaluate Z . By (9·6·1) this becomes

$$\blacktriangleright \quad Z = \sum_R e^{-\beta(n_1\epsilon_1 + n_2\epsilon_2 + \dots)} e^{-\alpha(n_1 + n_2 + \dots)} \quad (9 \cdot 6 \cdot 7)$$

where the sum is over all possible numbers (9·6·2) without restriction. By regrouping terms one obtains

$$\begin{aligned} Z &= \sum_{n_1, n_2, \dots} e^{-(\alpha+\beta\epsilon_1)n_1 - (\alpha+\beta\epsilon_2)n_2 - \dots} \\ &= \left(\sum_{n_1=0}^{\infty} e^{-(\alpha+\beta\epsilon_1)n_1} \right) \left(\sum_{n_2=0}^{\infty} e^{-(\alpha+\beta\epsilon_2)n_2} \right) \dots \end{aligned}$$

This is just a product of simple geometric series. Hence

$$Z = \left(\frac{1}{1 - e^{-(\alpha+\beta\epsilon_1)}} \right) \left(\frac{1}{1 - e^{-(\alpha+\beta\epsilon_2)}} \right) \dots$$

or $\ln Z = - \sum_r \ln (1 - e^{-\alpha-\beta\epsilon_r}) \quad (9 \cdot 6 \cdot 8)$

Equation (9·6·6) then yields

$$\blacktriangleright \quad \ln Z = \alpha N - \sum_r \ln (1 - e^{-\alpha - \beta \epsilon_r}) \quad (9 \cdot 6 \cdot 9)$$

Our argument assumed that the parameter α is to be chosen so that the function $Z(N')e^{-\alpha N'}$ has its maximum for $N' = N$, i.e., so that

$$\frac{\partial}{\partial N'} [\ln Z(N') - \alpha N'] = \frac{\partial \ln Z(N)}{\partial N} - \alpha = 0 \quad (9 \cdot 6 \cdot 10)$$

Since this condition involves the particular value $N' = N$, α itself must be a function of N . By virtue of (9·6·6), the condition (9·6·10) is equivalent to

$$\begin{aligned} & \left[\alpha + \left(N + \frac{\partial \ln Z}{\partial \alpha} \right) \frac{\partial \alpha}{\partial N} \right] - \alpha = 0 \\ \text{or} \quad & N + \frac{\partial \ln Z}{\partial \alpha} = \frac{\partial \ln Z}{\partial \alpha} = 0 \end{aligned} \quad (9 \cdot 6 \cdot 11)$$

Using the expression (9·6·9), the relation (9·6·11) which determines α is then

$$N - \sum_r \frac{e^{-\alpha - \beta \epsilon_r}}{1 - e^{-\alpha - \beta \epsilon_r}} = 0$$

or

$$\blacktriangleright \quad \sum_r \frac{1}{e^{\alpha + \beta \epsilon_r} - 1} = N \quad (9 \cdot 6 \cdot 12)$$

By applying (9·2·5) to (9·6·9) one obtains then

$$\bar{n}_s = -\frac{1}{\beta} \frac{\partial \ln Z}{\partial \epsilon_s} = -\frac{1}{\beta} \left[-\frac{\beta e^{-\alpha - \beta \epsilon_s}}{1 - e^{-\alpha - \beta \epsilon_s}} + \frac{\partial \ln Z}{\partial \alpha} \frac{\partial \alpha}{\partial \epsilon_s} \right]$$

The last term takes into account the fact that α is a function of ϵ_s through the relation (9·6·12). But this term vanishes by virtue of (9·6·11). Hence one has simply

$$\blacktriangleright \quad \bar{n}_s = \frac{1}{e^{\alpha + \beta \epsilon_s} - 1} \quad (9 \cdot 6 \cdot 13)$$

Thus one regains the Bose-Einstein distribution already derived previously in (9·3·22). Note that the condition (9·6·12) which determines α is then equivalent to

$$\sum_r \bar{n}_r = N \quad (9 \cdot 6 \cdot 14)$$

the obvious requirement needed to satisfy the conservation of particles (9·6·3).

The chemical potential of the gas is given by

$$\mu = \frac{\partial F}{\partial N} = -kT \frac{\partial \ln Z}{\partial N} = -kT\alpha$$

where we have used (9·6·10). Thus the parameter

$$\alpha = -\beta\mu \quad (9\cdot6\cdot15)$$

is directly related to the chemical potential of the gas. In the case of photons, where there is no restriction on the total number of particles, Z is independent of N ; then $\alpha = 0$, and all our relations reduce to those of the preceding section.

Calculation of the dispersion By applying (9·2·10) to (9·6·13) one obtains

$$\overline{(\Delta n_s)^2} = -\frac{1}{\beta} \frac{\partial \bar{n}_s}{\partial \epsilon_s} = \frac{1}{\beta} \frac{e^{\alpha+\beta\epsilon_s}}{(e^{\alpha+\beta\epsilon_s} - 1)^2} \left(\frac{\partial \alpha}{\partial \epsilon_s} + \beta \right)$$

But $\frac{e^{\alpha+\beta\epsilon_s}}{(e^{\alpha+\beta\epsilon_s} - 1)^2} = \frac{(e^{\alpha+\beta\epsilon_s} - 1) + 1}{(e^{\alpha+\beta\epsilon_s} - 1)^2} = \bar{n}_s + \bar{n}_s^2$

Hence $\overline{(\Delta n_s)^2} = \bar{n}_s(1 + \bar{n}_s) \left(1 + \frac{1}{\beta} \frac{\partial \alpha}{\partial \epsilon_s} \right) \approx \bar{n}_s(1 + \bar{n}_s) \quad (9\cdot6\cdot16)$

and

$$\blacktriangleright \quad \overline{(\Delta n_s)^2} \approx \frac{1}{\bar{n}_s} + 1 \quad (9\cdot6\cdot17)$$

where we have neglected the term $\partial \alpha / \partial \epsilon_s$. This term is usually very small, since α is to be determined by (9·6·11) and (unless the temperature $T = (k\beta)^{-1}$ is so low that only a very few terms in the sum have appreciable magnitude) a small change of *one* energy ϵ_s leaves the sum (and hence α) essentially unchanged.

Note that the relation (9·6·17) is exactly the same as that of (9·5·6) for photons. The relative dispersion is again greater than in the MB case of (9·4·9). Thus the relative dispersion does *not* become arbitrarily small even when $\bar{n}_s \gg 1$.

The correction term in (9·6·16) can, of course, be evaluated explicitly by taking (9·6·12), which determines α , and differentiating it with respect to ϵ_s . Thus

$$-\frac{\beta e^{\alpha+\beta\epsilon_s}}{(e^{\alpha+\beta\epsilon_s} - 1)^2} - \sum_r \frac{e^{\alpha+\beta\epsilon_r}}{(e^{\alpha+\beta\epsilon_r} - 1)^2} \frac{\partial \alpha}{\partial \epsilon_s} = 0$$

or $-\beta(\bar{n}_s + \bar{n}_s^2) - \left[\sum_r (\bar{n}_r + \bar{n}_r^2) \right] \frac{\partial \alpha}{\partial \epsilon_s} = 0$

Hence $\frac{\partial \alpha}{\partial \epsilon_s} = -\beta \frac{\bar{n}_s(1 + \bar{n}_s)}{\sum_r \bar{n}_r(1 + \bar{n}_r)}$

and $\overline{(\Delta n_s)^2} = \bar{n}_s(1 + \bar{n}_s) \left[1 - \frac{\bar{n}_s(1 + \bar{n}_s)}{\sum_r \bar{n}_r(1 + \bar{n}_r)} \right] \quad (9\cdot6\cdot18)$

The dispersion is thus slightly smaller than if the last term in the square brackets were neglected. But if one goes to the limit where $T \rightarrow 0$, then all the particles tend to be in the one single-particle state $s = 1$ of lowest energy, so that $\bar{n}_1 \approx N$ while $\bar{n}_s \approx 0$ for all other states. The correction term in (9·6·18) is then important since it predicts properly that the fluctuation in the number of particles in the ground state $s = 1$ goes to zero.

9 · 7 Fermi-Dirac statistics

The discussion here is very similar to that for Bose-Einstein statistics. The problem is again to evaluate the partition function (9 · 6 · 1). But, in accord with the discussion of Sec. 9 · 2, the summation is only over the two values

$$n_r = 0 \text{ and } 1 \quad \text{for each } r \quad (9 \cdot 7 \cdot 1)$$

where these numbers must again satisfy the restrictive condition (9 · 6 · 3).

The problem can be handled in a manner identical to that used in the last section for BE statistics. The unrestricted sum Z of (9 · 6 · 5) becomes

$$\begin{aligned} Z &= \sum_{n_1, n_2, n_3} e^{-\beta(n_1\epsilon_1 + n_2\epsilon_2 + \dots) - \alpha(n_1 + n_2 + \dots)} \\ &= \left(\sum_{n_1=0}^1 e^{-(\alpha+\beta\epsilon_1)n_1} \right) \left(\sum_{n_2=0}^1 e^{-(\alpha+\beta\epsilon_2)n_2} \right) \dots \end{aligned} \quad (9 \cdot 7 \cdot 2)$$

Here each sum consists, by virtue of (9 · 7 · 1), of only two terms and is thus trivial. Hence

$$Z = (1 + e^{-\alpha-\beta\epsilon_1})(1 + e^{-\alpha-\beta\epsilon_2}) \dots$$

$$\text{or} \quad \ln Z = \sum_r \ln (1 + e^{-\alpha-\beta\epsilon_r}) \quad (9 \cdot 7 \cdot 3)$$

Hence (9 · 6 · 5) becomes

$$\blacktriangleright \quad \ln Z = \alpha N + \sum_r \ln (1 + e^{-\alpha-\beta\epsilon_r}) \quad (9 \cdot 7 \cdot 4)$$

Except for some important sign changes, this expression is of the same form as (9 · 6 · 9) for the BE case. The parameter α is again to be determined from the condition (9 · 6 · 11). Thus

$$\frac{\partial \ln Z}{\partial \alpha} = N - \sum_r \frac{e^{-\alpha-\beta\epsilon_r}}{1 + e^{-\alpha-\beta\epsilon_r}} = 0$$

or

$$\blacktriangleright \quad \sum_r \frac{1}{e^{\alpha+\beta\epsilon_r} + 1} = N \quad (9 \cdot 7 \cdot 5)$$

By applying (9 · 2 · 5) to (9 · 7 · 4), one obtains

$$\bar{n}_s = -\frac{1}{\beta} \frac{\partial \ln Z}{\partial \epsilon_s} = \frac{1}{\beta} \frac{\beta e^{-\alpha-\beta\epsilon_s}}{1 + e^{-\alpha-\beta\epsilon_s}}$$

or

$$\blacktriangleright \quad \bar{n}_s = \frac{1}{e^{\alpha+\beta\epsilon_s} + 1} \quad (9 \cdot 7 \cdot 6)$$

Thus one regains the Fermi-Dirac distribution derived previously in (9 · 3 · 14). The relation (9 · 7 · 5) which is used to determine α is again just the condition (9 · 6 · 14) and the parameter α is again related to the chemical potential μ by the relation (9 · 6 · 15).

Calculation of the dispersion By applying (9·2·10) to (9·7·6), one obtains

$$\overline{(\Delta n_s)^2} = -\frac{1}{\beta} \frac{\partial \bar{n}_s}{\partial \epsilon_s} = \frac{1}{\beta} \frac{e^{\alpha+\beta\epsilon_s}}{(e^{\alpha+\beta\epsilon_s} + 1)^2} \left(\frac{\partial \alpha}{\partial \epsilon_s} + \beta \right)$$

But $\frac{e^{\alpha+\beta\epsilon_s}}{(e^{\alpha+\beta\epsilon_s} + 1)^2} = \frac{(e^{\alpha+\beta\epsilon_s} + 1) - 1}{(e^{\alpha+\beta\epsilon_s} + 1)^2} = \bar{n}_s - \bar{n}_s^2$

Hence $\overline{(\Delta n_s)^2} = \bar{n}_s(1 - \bar{n}_s) \left(1 + \frac{1}{\beta} \frac{\partial \alpha}{\partial \epsilon_s} \right) \approx \bar{n}_s(1 - \bar{n}_s)$ (9·7·7)

and

► $\frac{\overline{(\Delta n_s)^2}}{\bar{n}_s^2} \approx \frac{1}{\bar{n}_s} - 1$ (9·7·8)

Note that the relative dispersion is smaller than in the MB case discussed in Eq. (9·4·9). For example, if $\bar{n}_s \rightarrow 1$, the maximum value it can attain in accordance with the exclusion principle, then the dispersion vanishes. There is no fluctuation in \bar{n}_s for states which are completely filled.

9·8 Quantum statistics in the classical limit

The preceding sections dealing with the quantum statistics of ideal gases can be summarized by the statements that

► $\bar{n}_r = \frac{1}{e^{\alpha+\beta\epsilon_r} \pm 1}$ (9·8·1)

where the upper sign refers to FD and the lower one to BE statistics. If the gas consists of a fixed number N of particles, the parameter α is to be determined by the condition

► $\sum_r \bar{n}_r = \sum_r \frac{1}{e^{\alpha+\beta\epsilon_r} \pm 1} = N$ (9·8·2)

The partition function Z of the gas is given by

► $\ln Z = \alpha N \pm \sum_r \ln (1 \pm e^{-\alpha-\beta\epsilon_r})$ (9·8·3)

Let us now investigate the magnitude of α in some limiting cases. Consider first the case of a gas at a given temperature when its concentration is made sufficiently low, i.e., when N is made sufficiently small. The relation (9·8·2) can then only be satisfied if each term in the sum over all states is sufficiently small, i.e., if $\bar{n}_r \ll 1$ or $\exp(\alpha + \beta\epsilon_r) \gg 1$ for all states r . Similarly, consider the case of a gas with some fixed number N of particles when its temperature is made sufficiently large, i.e., when β is made sufficiently small. In the sum of (9·8·2) the terms of appreciable magnitude are those for which $\beta\epsilon_r \ll \alpha$; hence it follows that as $\beta \rightarrow 0$, an increasing number of terms with large values of ϵ_r contribute substantially to this sum. To prevent this sum from exceeding N , the parameter α must become large enough so that each

term is sufficiently small; i.e., it is again necessary that $\exp(\alpha + \beta\epsilon_r) \gg 1$ or $\bar{n}_r \ll 1$ for all states r . Thus one arrives at the conclusion that, if the concentration is made sufficiently low or if the temperature is made sufficiently high, α must become so large that,

$$\text{for all } r, \quad e^{\alpha+\beta\epsilon_r} \gg 1 \quad (9 \cdot 8 \cdot 4)$$

Equivalently this means that the occupation numbers become then small enough so that,

$$\text{for all } r, \quad \bar{n}_r \ll 1 \quad (9 \cdot 8 \cdot 5)$$

We shall call the limit of sufficiently low concentration or sufficiently high temperature where (9·8·4) or (9·8·5) are satisfied the "classical limit."

In this limit it follows by (9·8·4) that for both FD and BE statistics (9·8·1) reduces to

$$\bar{n}_r = e^{-\alpha-\beta\epsilon_r} \quad (9 \cdot 8 \cdot 6)$$

By virtue of (9·8·2), the parameter α is then determined by the condition

$$\sum_r e^{-\alpha-\beta\epsilon_r} = e^{-\alpha} \sum_r e^{-\beta\epsilon_r} = N$$

$$\text{or} \quad e^{-\alpha} = N \left(\sum_r e^{-\beta\epsilon_r} \right)^{-1} \quad (9 \cdot 8 \cdot 7)$$

$$\text{Thus} \quad \bar{n}_r = N \frac{e^{-\beta\epsilon_r}}{\sum_r e^{-\beta\epsilon_r}} \quad (9 \cdot 8 \cdot 8)$$

Hence it follows that in the classical limit of sufficiently low density or sufficiently high temperature the quantum distribution laws, whether FD or BE, reduce to the MB distribution.

The present conclusion is in agreement with our discussion of Sec. 7·4, where we estimated more quantitatively just how low the concentration and how high the temperature must be for classical results to be applicable.

Let us now consider the partition function of (9·8·3). In the classical limit, where (9·8·4) is satisfied, one can expand the logarithm in (9·8·3) to get

$$\ln Z = \alpha N \pm \sum_r (\pm e^{-\alpha-\beta\epsilon_r}) = \alpha N + N$$

But by virtue of (9·8·7)

$$\alpha = -\ln N + \ln \left(\sum_r e^{-\beta\epsilon_r} \right)$$

$$\text{Hence} \quad \ln Z = -N \ln N + N + N \ln \left(\sum_r e^{-\beta\epsilon_r} \right) \quad (9 \cdot 8 \cdot 9)$$

Note that this does *not* equal the partition function Z_{MB} computed in Eq. (9·4·4) for MB statistics

$$\ln Z_{MB} = N \ln \left(\sum_r e^{-\beta\epsilon_r} \right) \quad (9 \cdot 8 \cdot 10)$$

Indeed

$$\ln Z = \ln Z_{MB} - (N \ln N - N)$$

Thus

$$\ln Z = \ln Z_{MB} - \ln N!$$

or

$$Z = \frac{Z_{MB}}{N!} \quad (9 \cdot 8 \cdot 11)$$

where we have used Stirling's formula since N is large. Here the factor $N!$ corresponds simply to the number of possible permutations of the particles, permutations which are physically meaningless when the particles are identical. It was precisely this factor which we had to introduce in an *ad hoc* fashion in Sec. 7·3 to save ourselves from the nonphysical consequences of the Gibbs paradox. What we have done in this section is to justify the whole discussion of Sec. 7·3 as being appropriate for a gas treated properly by quantum mechanics in the limit of sufficiently low concentration or high temperature. The partition function is automatically correctly evaluated by (9·8·9), there is no Gibbs paradox, and everything is consistent.

A gas in the classical limit where (9·8·6) is satisfied is said to be "non-degenerate." On the other hand, if the concentration and temperature are such that the actual FD or BE distribution (9·8·1) must be used, the gas is said to be "degenerate."

IDEAL GAS IN THE CLASSICAL LIMIT

9·9 *Quantum states of a single particle*

Wave function To complete the discussion of the statistical problem it is necessary to enumerate the possible quantum states s and corresponding energies ϵ_s of a single noninteracting particle. Consider this particle to be nonrelativistic and denote its mass by m , its position vector by \mathbf{r} , and its momentum by \mathbf{p} . Suppose that the particle is confined within a container of volume V within which the particle is subject to no forces. Neglecting for the time being the effect of the bounding walls, the wave function $\Psi(\mathbf{r},t)$ of the particle is then simply described by a plane wave of the form

$$\Psi = Ae^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} = \psi(\mathbf{r}) e^{-i\omega t} \quad (9\cdot9\cdot1)$$

which propagates in a direction specified by the "wave vector" \mathbf{k} and which has some constant amplitude A . Here the energy ϵ of the particle is related to the frequency ω by

$$\epsilon = \hbar\omega \quad (9\cdot9\cdot2)$$

while its momentum is related to its wave vector \mathbf{k} by the de Broglie relation

$$\mathbf{p} = \hbar\mathbf{k} \quad (9\cdot9\cdot3)$$

Thus one has

$$\epsilon = \frac{\mathbf{p}^2}{2m} = \frac{\hbar^2\mathbf{k}^2}{2m} \quad (9\cdot9\cdot4)$$

The basic justification for these statements is, of course, the fact that Ψ must satisfy the Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} = \mathcal{H}\Psi \quad (9\cdot9\cdot5)$$

Since one can choose the potential energy to be zero inside the container, the Hamiltonian \mathcal{H} reduces there to the kinetic energy alone; i.e.,

$$\mathcal{H} = \frac{1}{2m} \mathbf{p}^2 = \frac{1}{2m} \left(\frac{\hbar}{i} \nabla \right)^2 = -\frac{\hbar^2}{2m} \nabla^2$$

where

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

Putting

$$\Psi = \psi e^{-i\epsilon t} = \psi e^{-(i/\hbar)\epsilon t} \quad (9 \cdot 9 \cdot 6)$$

where ψ does not depend on time, (9 · 9 · 5) reduces to the time-independent Schrödinger equation

$$\mathcal{H}\psi = \epsilon\psi \quad (9 \cdot 9 \cdot 7)$$

or

$$\nabla^2\psi + \frac{2me}{\hbar^2}\psi = 0 \quad (9 \cdot 9 \cdot 8)$$

Equation (9 · 9 · 7) shows that ϵ corresponds to the possible values of \mathcal{H} and is thus the energy of the particle. The wave equation (9 · 9 · 8) has solutions of the general form

$$\psi = A e^{i(\kappa_x x + \kappa_y y + \kappa_z z)} = A e^{i\boldsymbol{\kappa} \cdot \mathbf{r}} \quad (9 \cdot 9 \cdot 9)$$

where $\boldsymbol{\kappa}$ is the constant "wave vector" with components κ_x , κ_y , κ_z . By substitution of (9 · 9 · 9) into (9 · 9 · 8) one finds that the latter equation is satisfied if

$$-(\kappa_x^2 + \kappa_y^2 + \kappa_z^2) + \frac{2me}{\hbar^2} = 0$$

Thus

$$\epsilon = \frac{\hbar^2 \kappa^2}{2m} \quad (9 \cdot 9 \cdot 10)$$

and ϵ is only a function of the magnitude $\kappa \equiv |\boldsymbol{\kappa}|$ of $\boldsymbol{\kappa}$. Since

$$p\psi = \frac{\hbar}{i} \nabla \psi = \hbar \boldsymbol{\kappa} \psi$$

one obtains then the relations (9 · 9 · 3) and (9 · 9 · 4).

Up to now we have considered only the translational degrees of freedom. If the particle also has an intrinsic spin angular momentum, the situation is scarcely more complicated; there is then simply a different function ψ for each possible orientation of the particle spin. For example, if the particle has spin $\frac{1}{2}$ (e.g., if it is an electron), then there are two possible wave functions ψ_{\pm} corresponding to the two possible values $m = \pm \frac{1}{2}$ of the quantum number specifying the orientation of the particle's spin angular momentum.

Boundary conditions and enumeration of states The wave function ψ must satisfy certain boundary conditions. Accordingly, not all possible values of $\boldsymbol{\kappa}$ (or \mathbf{p}) are allowed, but only certain discrete values. The corresponding energies of the particle are then also quantized by virtue of (9 · 9 · 4).

The boundary conditions can be treated in a very general and simple way in the usual situation where the container enclosing the gas of particles is

large enough that its smallest linear dimension L is much greater than the de Broglie wavelength $\lambda = 2\pi/|\kappa|$ of the particle under consideration.* It is then physically clear that the detailed properties of the bounding walls of the container (e.g., their shape or the nature of the material of which they are made) must become of negligible significance in describing the behavior of a particle located well within the container.† To make the argument more precise, let us consider any macroscopic volume element which is large compared to λ and which lies well within the container so that it is everywhere removed from the container walls by distances large compared to λ . The actual wave function anywhere within the container can always be written as a superposition of plane waves (9·9·1) with all possible wave vectors κ . Hence one can regard the volume element under consideration as being traversed by waves of the form (9·9·1) traveling in all possible directions specified by κ , and with all possible wavelengths related to the magnitude of κ . Since the container walls are far away (compared to λ), it does not really matter just how each such wave is ultimately reflected from these walls, or which wave gets reflected how many times before it passes again through the volume element under consideration. The number of waves of each kind traversing this volume element should be quite insensitive to any such details which describe what happens near the container walls and should be substantially unaffected if the shape or properties of these walls are modified. Indeed, it is simplest if one imagines these walls moved out to infinity, i.e., if one effectively eliminates the walls altogether. One can then avoid the necessity of treating the problem of reflections at the walls, a problem which is really immaterial in describing the situation in the volume element under consideration. It does not matter whether a given wave enters this volume element after having been reflected somewhere far away, or after coming in from infinity without ever having been reflected at all.

The foregoing comments show that, for purposes of discussing the properties of a gas anywhere but in the immediate vicinity of the container walls, the exact nature of the boundary conditions imposed on each particle should be unimportant. One can therefore formulate the problem in a way which makes these boundary conditions as simple as possible. Let us therefore choose the basic volume V of gas under consideration to be in the shape of a rectangular parallelepiped with edges parallel to the x , y , z axes and with respective edge lengths equal to L_x , L_y , L_z . Thus $V = L_x L_y L_z$. The simplest boundary conditions to impose are such that a traveling wave of the form (9·9·1) is indeed an exact solution of the problem. This requires that the wave (9·9·1) be able to propagate indefinitely without suffering any reflections. In order to make the boundary conditions consistent with this simple situation, one can neglect

* This condition is ordinarily very well satisfied for essentially all molecules of a gas since a typical order of magnitude, already estimated in Sec. 7·4, is $\lambda \approx 1 \text{ \AA}$ for an atom of thermal energy at room temperature.

† Note that the fraction of particles near the surface of the container, i.e., within a distance λ of its walls, is of the order of $\lambda L^2/L^3 = \lambda/L$ and is thus ordinarily utterly negligible for a macroscopic container.

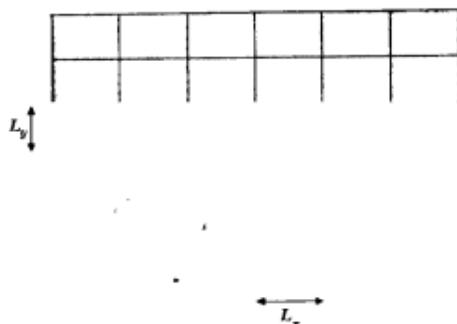


Fig. 9·9·1 The volume under consideration (indicated in darker gray) is here considered embedded in an array of similar volumes extending throughout all space. Wall effects are thus effectively eliminated.

completely the presence of any container walls and can imagine that the volume of gas under consideration is embedded in an infinite set of similar volumes in each of which the physical situation is exactly the same (see Fig. 9·9·1). The wave function must then satisfy the conditions

$$\left. \begin{aligned} \psi(x + L_x, y, z) &= \psi(x, y, z) \\ \psi(x, y + L_y, z) &= \psi(x, y, z) \\ \psi(x, y, z + L_z) &= \psi(x, y, z) \end{aligned} \right\} \quad (9 \cdot 9 \cdot 11)$$

The requirement that the wave function be the same in any of the parallel-epipeds should not affect the physics of interest in the one volume under consideration if its dimensions are large compared to the de Broglie wavelength λ of the particle.

Remark Suppose that the problem were one-dimensional so that a particle moves in the x direction in a container of length L_x . Then one can eliminate the effects of reflections by imagining the container to be bent around in the form of a circle as shown in Fig. 9·9·2. If L_x is very large, the curvature is quite negligible so that the situation inside the container is substantially the same as before. But the advantage is that there are now no container walls to worry about. Hence traveling waves described by (9·9·1) and going around without reflection are perfectly good solutions of the problem. It is only necessary to note that the points x and $x + L_x$ are now coincident; the requirement that the wave function be single-valued implies the condition

$$\psi(x + L_x) = \psi(x) \quad (9 \cdot 9 \cdot 12)$$



Fig. 9·9·2 A one-dimensional container of length L_x bent into a circle by joining its ends.

This is precisely the analog of (9·9·11) in one dimension. Indeed, one could regard the condition (9·9·11) as resulting from the attempt to eliminate reflections in three dimensions by imagining the original parallelepiped to be bent into a doughnut in four dimensions. (This is, admittedly, difficult to visualize.)

—

This point of view, which describes the situation in terms of simple traveling waves satisfying the periodic boundary conditions (9·9·11), is very convenient and mathematically exceedingly easy. By virtue of (9·9·1) or (9·9·9)

$$\psi = e^{i\kappa \cdot r} = e^{i(\kappa_x x + \kappa_y y + \kappa_z z)}$$

To satisfy (9·9·11) one must require that

$$\kappa_x(x + L_x) = \kappa_x x + 2\pi n_x \quad (n_x \text{ integral})$$

$$\left. \begin{aligned} \text{or} \quad \kappa_x &= \frac{2\pi}{L_x} n_x \\ \text{Similarly,} \quad \kappa_y &= \frac{2\pi}{L_y} n_y \\ \text{and} \quad \kappa_z &= \frac{2\pi}{L_z} n_z \end{aligned} \right\} \quad (9 \cdot 9 \cdot 13)$$

Here the numbers n_x, n_y, n_z are *any* set of integers—positive, negative, or zero.

The components of $\kappa = p/\hbar$ are thus quantized in discrete units. Accordingly (9·9·4) yields the possible quantized particle energies

$$\epsilon = \frac{\hbar^2}{2m} (\kappa_x^2 + \kappa_y^2 + \kappa_z^2) = \frac{2\pi^2 \hbar^2}{m} \left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2} \right) \quad (9 \cdot 9 \cdot 14)$$

Note that for any kind of macroscopic volume where L_x, L_y, L_z are large, the possible values of the wave-vector components given by (9·9·13) are very closely spaced. There are thus very many states of the particle (i.e., very many possible integers n_x) corresponding to any small range $d\kappa_x$ of a wave-vector component. It is easy to do some counting. For given values of κ_y and κ_z , it follows by (9·9·13) that the number Δn_x of possible integers n_x for which κ_x lies in the range between κ_x and $\kappa_x + d\kappa_x$ is equal to

$$\Delta n_x = \frac{L_x}{2\pi} d\kappa_x \quad (9 \cdot 9 \cdot 15)$$

The number of translational states $\rho(\kappa) d^3\kappa$ for which κ is such that it lies in the range between κ and $\kappa + d\kappa$ (i.e., in the range such that its x component is between κ_x and $\kappa_x + d\kappa_x$, its y component between κ_y and $\kappa_y + d\kappa_y$, and its z component between κ_z and $\kappa_z + d\kappa_z$) is then given by the product of the numbers of possible integers in the three component ranges. Thus

$$\rho d^3\kappa = \Delta n_x \Delta n_y \Delta n_z = \left(\frac{L_x}{2\pi} d\kappa_x \right) \left(\frac{L_y}{2\pi} d\kappa_y \right) \left(\frac{L_z}{2\pi} d\kappa_z \right) = \frac{L_x L_y L_z}{(2\pi)^3} d\kappa_x d\kappa_y d\kappa_z$$

or

$$\blacktriangleright \quad \rho d^3\kappa = \frac{V}{(2\pi)^3} d^3\kappa \quad (9 \cdot 9 \cdot 16)$$

where $d^3\kappa = dk_x dk_y dk_z$ is the element of volume in " κ space." Note that the density of states ρ is independent of κ and proportional to the volume V under consideration; i.e., the number of states *per unit volume*, with a wave number κ (or momentum $p = \hbar\kappa$) lying in some given range, is a constant independent of the magnitude or shape of the volume.

Remark Note that (9 · 9 · 3) yields for the number of translational states $\rho_p d^3p$ in the momentum range between p and $p + dp$ the expression

$$\rho_p d^3p = \rho d^3\kappa = \frac{V}{(2\pi)^3} \frac{d^3p}{\hbar^3} = V \frac{d^3p}{\hbar^3} \quad (9 \cdot 9 \cdot 17)$$

where $\hbar = 2\pi\hbar$ is the ordinary Planck's constant. Now $V d^3p$ is the volume of the classical six-dimensional phase space occupied by a particle in a box of volume V and with momentum between p and $p + dp$. Thus (9 · 9 · 17) shows that subdivision of this phase space into cells of size \hbar^3 yields the correct number of quantum states for the particle.

Various other relations can be deduced from the result (9 · 9 · 16). For example, let us find the number of translational states $\rho_\kappa d\kappa$ for which κ is such that its magnitude $|\kappa|$ lies in the range between κ and $\kappa + dk$. This is obtained by summing (9 · 9 · 16) over all values of κ in this range, i.e., over the volume in κ space of the portion of spherical shell lying between radii κ and $\kappa + dk$. Thus

$$\rho_\kappa d\kappa = \frac{V}{(2\pi)^3} (4\pi\kappa^2 d\kappa) = \frac{V}{2\pi^2} \kappa^2 d\kappa \quad (9 \cdot 9 \cdot 18)$$

Remark Since ϵ depends only on $\kappa = |\kappa|$, (9 · 9 · 18) gives immediately, corresponding to this range of κ , the corresponding number of translational states $\rho_\epsilon d\epsilon$ for which the energy of the particle lies between ϵ and $\epsilon + de$. From the equality of states one has

$$|\rho_\epsilon d\epsilon| = |\rho_\kappa d\kappa| = \rho_\kappa \left| \frac{d\kappa}{d\epsilon} \right| d\epsilon = \rho_\kappa \left| \frac{d\epsilon}{d\kappa} \right|^{-1} d\epsilon$$

By (9 · 9 · 4) one then obtains

$$\rho_\epsilon d\epsilon = \frac{V}{2\pi^2} \kappa^2 \left| \frac{d\kappa}{d\epsilon} \right| d\epsilon = \frac{V}{4\pi^2} \frac{(2m)^{\frac{1}{2}}}{\hbar^3} \epsilon^{\frac{1}{2}} d\epsilon \quad (9 \cdot 9 \cdot 19)$$

—

Alternative discussion It is, of course, possible to adopt a slightly more complicated point of view which does take into account explicitly reflections occurring at the walls of the container. Since the exact boundary conditions

are immaterial let us, for simplicity, assume that the container is in the shape of a rectangular parallelepiped with walls located at $x = 0$ and $x = L_x$, $y = 0$ and $y = L_y$, and $z = 0$ and $z = L_z$. Let us further assume that these walls are perfectly reflecting, i.e., that the potential energy U of the particle equals $U = 0$ inside the box and $U = \infty$ outside the box. Then the wave function ψ must satisfy the requirement that

$$\psi = 0 \quad \begin{cases} \text{whenever } x = 0 \text{ or } L_x \\ \quad y = 0 \text{ or } L_y \\ \quad z = 0 \text{ or } L_z \end{cases} \quad (9 \cdot 9 \cdot 20)$$

The particular solution $\psi = e^{ik_x x}$ of (9·9·9) represents a traveling wave and does *not* satisfy the boundary conditions (9·9·20). But one can construct suitable linear combinations of (9·9·9) (all of which automatically also satisfy the Schrödinger equation (9·9·8)) which do satisfy the boundary conditions (9·9·20). What this means physically is that in this box with perfectly reflecting parallel walls standing waves are set up which result from the superposition of traveling waves propagating back and forth.* Mathematically, since $e^{ik_x x}$ is a solution of (9·9·8), so is $e^{-ik_x x}$. The combination

$$(e^{ik_x x} - e^{-ik_x x}) \propto \sin \kappa_x x \quad (9 \cdot 9 \cdot 21)$$

vanishes properly when $x = 0$. It can also be made to vanish for $x = L_x$, provided one chooses κ_x so that

$$\kappa_x L_x = \pi n_x$$

where n_x is any integer. Here the possible values n_x should be restricted to the positive set

$$n_x = 1, 2, 3, \dots$$

since a sign reversal of n_x (or κ_x) just turns the function (9·9·20) into

$$\sin(-\kappa_x)x = -\sin \kappa_x x$$

which is not a distinct new wave function. Thus a standing wave solution is specified completely by $|\kappa_x|$.

Forming standing waves analogous to (9·9·21) also for the y and z directions, one obtains the product wave function

$$\psi = A(\sin \kappa_x x)(\sin \kappa_y y)(\sin \kappa_z z) \quad (9 \cdot 9 \cdot 22)$$

where A is some constant. This satisfies the Schrödinger equation (9·9·8) and also the boundary conditions (9·9·20) provided that

$$\kappa_x = \frac{\pi}{L_x} n_x, \quad \kappa_y = \frac{\pi}{L_y} n_y, \quad \kappa_z = \frac{\pi}{L_z} n_z \quad (9 \cdot 9 \cdot 23)$$

* Simple standing waves of the form (9·9·21) would not be set up if the walls of the container were not exactly parallel. Hence our previous discussion in terms of traveling waves criss-crossing the volume in all directions, in a manner insensitive to the precise boundary conditions, affords a more convenient and general point of view.

where n_x, n_y, n_z are any *positive* integers. The possible energies of the particle are then given by

$$\epsilon = \frac{\hbar^2}{2m} \kappa^2 = \frac{\pi^2 \hbar^2}{2m} \left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2} \right)$$

For given values of κ_y and κ_z , the number of translational states with κ_x in the range between κ_x and $\kappa_x + d\kappa_x$ is now equal to

$$\Delta n_x = \frac{L_x}{\pi} d\kappa_x \quad (9 \cdot 9 \cdot 24)$$

The number of translational states with κ in the range between κ and $\kappa + d\kappa$ is then given by

$$\rho d^3\kappa = \Delta n_x \Delta n_y \Delta n_z = \left(\frac{L_x}{\pi} d\kappa_x \right) \left(\frac{L_y}{\pi} d\kappa_y \right) \left(\frac{L_z}{\pi} d\kappa_z \right)$$

or

$$\rho d^3\kappa = \frac{V}{\pi^3} d^3\kappa \quad (9 \cdot 9 \cdot 25)$$

The number of translational states $\rho_\kappa d\kappa$ for which κ is such that its *magnitude* lies in the range between κ and $\kappa + d\kappa$ is obtained by summing (9·9·25) over all values of κ in this range, i.e., over the volume in κ space of the portion of spherical shell lying between radii κ and $\kappa + d\kappa$ and located in the first octant where $\kappa_x, \kappa_y, \kappa_z > 0$ so as to satisfy (9·9·23). Thus (9·9·25) yields

$$\rho_\kappa d\kappa = \frac{V}{\pi^3} \left(\frac{4\pi\kappa^2 d\kappa}{8} \right) = \frac{V}{2\pi^2} \kappa^2 d\kappa \quad (9 \cdot 9 \cdot 26)$$

This is the *same* result as was obtained in (9·9·18). The reason is simple. By (9·9·24) there are, compared to (9·9·15), twice as many states lying in a given interval $d\kappa_x$, but since only positive values of κ_x are now to be counted, the number of such intervals is decreased by a compensating factor of 2.

By (9·9·26) it also follows that $\rho_\kappa d\kappa$ is the same as in (9·9·19). This just illustrates the result (which can also be established by rather elaborate general mathematical arguments)* that this density of states should be the same irrespective of the shape of the container or of the exact boundary conditions imposed on its surface, so long as the de Broglie wavelength of the particle is small compared to the dimensions of the container.

9 · 10 Evaluation of the partition function

We are now ready to calculate the partition function Z of a monatomic ideal gas in the classical limit of sufficiently low density or sufficiently high temperature. By (9·8·9) one has

$$\ln Z = N(\ln \zeta - \ln N + 1) \quad (9 \cdot 10 \cdot 1)$$

where

$$\zeta \equiv \sum_r e^{-\beta \epsilon_r} \quad (9 \cdot 10 \cdot 2)$$

* See, for example, R. Courant and D. Hilbert, "Methods of Mathematical Physics," vol. I, pp. 429–445, Interscience Publishers, New York, 1953.

is the sum over all states of a single particle. The expression (9·10·1) is identical to the result (7·3·3); i.e.,

$$Z = \frac{\zeta^N}{N!} \quad (9 \cdot 10 \cdot 3)$$

Since we have just enumerated the possible states of a single particle, the sum (9·10·2) is readily evaluated. By (9·9·14)

$$\zeta = \sum_{\kappa_x, \kappa_y, \kappa_z} \exp \left[-\frac{\beta \hbar^2}{2m} (\kappa_x^2 + \kappa_y^2 + \kappa_z^2) \right] \quad (9 \cdot 10 \cdot 4)$$

where the sum is over all possible values of κ_x , κ_y , κ_z given by (9·9·13). Since the exponential function factors, ζ becomes the product of three similar sums

$$\zeta = \left(\sum_{\kappa_x} e^{-(\beta \hbar^2 / 2m) \kappa_x^2} \right) \left(\sum_{\kappa_y} e^{-(\beta \hbar^2 / 2m) \kappa_y^2} \right) \left(\sum_{\kappa_z} e^{-(\beta \hbar^2 / 2m) \kappa_z^2} \right) \quad (9 \cdot 10 \cdot 5)$$

Successive terms in a sum like that over $\kappa_x = (2\pi/L_x)n_x$ correspond to a very small increment $\Delta \kappa_x = 2\pi/L_x$ in κ_x and differ, therefore, very little from each other; i.e.,

$$\left| \frac{\partial}{\partial \kappa_x} [e^{-(\beta \hbar^2 / 2m) \kappa_x^2}] \left(\frac{2\pi}{L_x} \right) \right| \ll e^{-(\beta \hbar^2 / 2m) \kappa_x^2} \quad (9 \cdot 10 \cdot 6)$$

Provided that this condition is satisfied, it is an excellent approximation to replace the sums in (9·10·5) by integrals. A small range between κ_x and $\kappa_x + d\kappa_x$ contains then, by (9·9·15), $\Delta n_x = (L_x/2\pi) d\kappa_x$ terms which have nearly the same magnitude and can be grouped together. Summing over all possible ranges of κ_x completes the sum. Thus

$$\begin{aligned} \sum_{\kappa_x=-\infty}^{\infty} e^{-(\beta \hbar^2 / 2m) \kappa_x^2} &\approx \int_{-\infty}^{\infty} e^{-(\beta \hbar^2 / 2m) \kappa_x^2} \left(\frac{L_x}{2\pi} d\kappa_x \right) \\ &= \frac{L_x}{2\pi} \left(\frac{2\pi m}{\beta \hbar^2} \right)^{\frac{1}{2}} = \frac{L_x}{2\pi \hbar} \left(\frac{2\pi m}{\beta} \right)^{\frac{1}{2}} \quad \text{by (A·4·2)} \end{aligned}$$

Hence (9·10·5) becomes

$$\zeta = \frac{V}{(2\pi \hbar)^3} \left(\frac{2\pi m}{\beta} \right)^{\frac{3}{2}} = \frac{V}{\hbar^3} (2\pi m k T)^{\frac{3}{2}} \quad (9 \cdot 10 \cdot 7)$$

Note that this is the same result as that obtained by the classical calculation in (7·2·6), provided that we set the arbitrary parameter h_0 (which measures the size of a cell in classical phase space) equal to Planck's constant \hbar .

It then follows, by (9·10·1), that

$$\ln Z = N \left(\ln \frac{V}{N} - \frac{3}{2} \ln \beta + \frac{3}{2} \ln \frac{2\pi m}{\hbar^2} + 1 \right) \quad (9 \cdot 10 \cdot 8)$$

$$\text{Hence } \bar{E} = - \frac{\partial \ln Z}{\partial \beta} = \frac{3}{2} \frac{N}{\beta} = \frac{3}{2} N k T \quad (9 \cdot 10 \cdot 9)$$

and

$$\blacktriangleright \quad S = k(\ln Z + \beta\bar{E}) = Nk \left(\ln \frac{V}{N} + \frac{3}{2} \ln T + \sigma_0 \right) \quad (9 \cdot 10 \cdot 10)$$

where

$$\blacktriangleright \quad \sigma_0 \equiv \frac{3}{2} \ln \frac{2\pi mk}{h^2} + \frac{5}{2} \quad (9 \cdot 10 \cdot 11)$$

These results are exactly the same as those we obtained in (7 · 3 · 5), with one important difference. Since we have now treated the problem by quantum mechanics, the constant σ_0 has a definite value in terms of Planck's constant h (unlike the classical case where h_0 was an arbitrary parameter). The fact that the entropy does not involve any arbitrary constants has important physical consequences which we shall discuss in Sec. 9 · 11. All quantities such as \bar{E} or the mean pressure \bar{p} , which depend only on derivatives of S are, of course, the same as those calculated in Sec. 7 · 2.

Let us verify that the condition (9 · 10 · 6) justifying the replacement of the sum over states by an integral is indeed satisfied. This condition requires that

$$\left| \frac{\beta h^2}{m} \kappa_x \frac{2\pi}{L_x} \right| \ll 1 \quad (9 \cdot 10 \cdot 12)$$

But the mean value of κ_x can be estimated from (9 · 10 · 9) or the equipartition theorem. Thus

$$\frac{\hbar^2 \kappa_x^2}{2m} = \frac{1}{3} \frac{\hbar^2 \kappa^2}{2m} = \frac{1}{2} kT$$

$$\hbar \bar{\kappa}_x \approx \sqrt{mkT}$$

or

Hence (9 · 10 · 12) becomes

$$\frac{\hbar}{mkT} \sqrt{mkT} \frac{2\pi}{L_x} = \frac{\hbar}{\sqrt{mkT}} \frac{1}{L_x} \ll 1$$

or approximately

$$\bar{\lambda} \ll L_x \quad (9 \cdot 10 \cdot 13)$$

where $\bar{\lambda} = \hbar/\bar{\kappa}$ is the mean de Broglie wavelength of the particle.

Thus (9 · 10 · 12) demands only that $\bar{\lambda}$ is smaller than the smallest dimension L of the container. On the other hand, we saw in Sec. 7 · 4 that the requirement for the very applicability of the classical approximation is that $\bar{\lambda}$ be smaller than the mean interparticle separation, i.e.,

$$\bar{\lambda} \ll \frac{L}{N^{\frac{1}{3}}} \quad (9 \cdot 10 \cdot 14)$$

which is a much more stringent condition than (9 · 10 · 13).

Finally, let us point out what happens if each particle has also an intrinsic spin angular momentum J . The possible orientations of this spin are specified by its projection $m_J = -J, -J+1, \dots, J-1, J$. There are then $(2J+1)$ possible states of the same energy associated with each possible translational state of a particle. The net result is that the sum over states ζ simply is multiplied by $(2J+1)$, so that the entropy is increased by the constant $Nk \ln (2J+1)$.

9 · 11 *Physical implications of the quantum-mechanical enumeration of states*

Although the results of the quantum-mechanical calculation of Z are virtually the same as those of the semiclassical calculation in Sec. 7 · 3, there are two significant differences:

- The correct dependence (9 · 10 · 1) of $\ln Z$ on N (i.e., the factor $N!$ in (9 · 10 · 3)) is an automatic consequence of the theory. Thus the Gibbs paradox does not arise, and $\ln Z$ in (9 · 10 · 8) behaves properly like an extensive quantity under simultaneous change of scale of N and V .
- There are no arbitrary constants occurring in Z or the entropy S derived therefrom; instead Z is a well-defined number involving Planck's constant h .

These differences reflect the fact that we have now unambiguously counted the number of quantum states available to the gas. We should expect this enumeration to be particularly important in cases involving transfer of particles from one phase to another (or from one component to another), since in these cases a calculation of the equilibrium situation must compare the actual number of states available in one phase with that in another (or for one type of molecule with that for another). Mathematically, this is manifested by the properties of the chemical potential

$$\mu = \left(\frac{\partial F}{\partial N} \right)_{V,T} = -kT \left(\frac{\partial \ln Z}{\partial N} \right)_{V,T} \quad (9 \cdot 11 \cdot 1)$$

In the last chapter we saw that the chemical potential is the important parameter determining the equilibrium conditions between phases or chemical components. On the other hand, it is clear from (9 · 11 · 1) and (9 · 10 · 1) that

$$\mu = -kT \ln \frac{\zeta}{N} \quad (9 \cdot 11 \cdot 2)$$

does depend on N and the various constants, such as Planck's constant, involved in ζ . Thus the quantum-mechanical calculation of Z in terms of these constants allows one to make predictions completely outside the realm of any theory based on classical statistical mechanics. We shall give two representative illustrations.

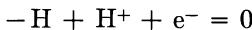
Thermal ionization of hydrogen atoms Suppose that H atoms are enclosed in a container of volume V at a high temperature T .* There then exists the possibility of ionization into a hydrogen ion H^+ and an electron e^- . This can be described in terms of the reaction



Let ϵ_0 denote the energy necessary to ionize the atom, i.e., its "ionization potential." This means that the ground state of the H atom has an energy $(-\epsilon_0)$ relative to the state where the proton H^+ and the electron e^- are at rest

* We assume that this temperature is high enough that the number of H_2 molecules is negligible, practically all of them being dissociated into H atoms.

separated by an infinite distance from each other.* Viewing (9·11·3) as a chemical equilibrium of the type discussed in Sec. 8·9, we can write it in the standard form



so that the law of mass action (8·10·21) becomes

$$\frac{N_+ N_-}{N_{\text{H}}} = K_N \quad (9\cdot11\cdot4)$$

where

$$K_N = \frac{\xi_+ \xi_-}{\xi_{\text{H}}} \quad (9\cdot11\cdot5)$$

Here N denotes the mean number of particles of each kind and the subscripts $+$, $-$, and H refer to the H^+ ion, the electron, and the H atom, respectively.

We are now in a position to calculate the quantities ξ from first principles. It is only necessary to be sure, for the sake of consistency, that all energies in the problem are measured from the same standard state. We shall choose this standard state to be the one where the electron and proton are at rest at infinite separation from each other. Furthermore, we shall assume that the H^+ and e^- concentrations are relatively small. The classical limit is then applicable at these high temperatures, and any coulomb attraction between the separated protons and electrons can be neglected.

Thus one can use (9·10·7) to write for the electron of mass m ,

$$\xi_- = 2 \frac{V}{h^3} (2\pi mkT)^{\frac{1}{2}} \quad (9\cdot11\cdot6)$$

Here the factor of 2 is introduced, since the *electron* has spin $\frac{1}{2}$ and, therefore, has two possible spin states for each translational state. Similarly, for the freely moving proton of mass M , one obtains

$$\xi_+ = 2 \frac{V}{h^3} (2\pi MkT)^{\frac{1}{2}} \quad (9\cdot11\cdot7)$$

Here the factor of 2 is introduced, because the *nuclear* spin of the proton is $\frac{1}{2}$, so that there are two possible nuclear spin orientations for each translational state of the proton.

The H atom has a mass $M + m \approx M$, since $m \ll M$. Its *internal* energy measured with respect to our chosen standard state is $(-\epsilon_0)$, since practically all H atoms are in their ground state at the temperature under consideration.† Hence one can write for the H atom

$$\xi_{\text{H}} = 4 \frac{V}{h^3} (2\pi MkT)^{\frac{1}{2}} e^{\epsilon_0/kT} \quad (9\cdot11\cdot8)$$

* From atomic physics we know that $\epsilon_0 = \frac{1}{2}(e^2/a_0)$, where $a_0 = \hbar^2/me^2$ is the Bohr radius. Numerically $\epsilon_0 = 13.6$ electron volts. (This is about three times larger than the energy necessary to dissociate a H_2 molecule.)

† The first excited state has an energy $-\frac{1}{4}\epsilon_0$, so that the relative probability of finding an atom in this state rather than in the ground state is

$$e^{\frac{1}{4}\beta\epsilon_0}/e^{\beta\epsilon_0} = e^{-\frac{3}{4}\beta\epsilon_0}$$

which is very small, even if $T = (k\beta)^{-1} = 10,000^\circ\text{K}$.

Here the factor of 4 is introduced, since there are four possible states of the atom for each translation state: two states of possible electron spin orientation, and for each of these, two states of possible nuclear spin orientation.

Combining these various expressions, one obtains, by (9·11·5),

$$K_N = \frac{V}{h^3} (2\pi mkT)^{\frac{3}{2}} e^{-\epsilon_0/kT} \quad (9\cdot11\cdot9)$$

which is the desired expression for the equilibrium constant. Note that all the statistical weighting factors due to the existence of spin have cancelled.

What (9·11·4) and (9·11·9) say physically is that the large ionization potential ϵ_0 tends to favor the existence of the H atom since this is the system of lowest energy. On the other hand, many more states become accessible to the system, i.e., its entropy tends to be greater, if one deals with two separate particles. The equilibrium situation represents the compromise between these two tendencies. More generally speaking, the most probable situation is that where the free energy $F = E - TS$ is minimum. At low temperatures where $F \approx E$ this favors the situation of low energy, i.e., the H atom. On the other hand, when T becomes large, F can become small if the entropy S is large, and this favors dissociation.

Suppose that a number N_0 of H atoms are present in the container at some temperature low enough that $N_- = N_+ \approx 0$, and that the temperature is then raised to the value T . Let ξ denote the fraction of atoms dissociated at this temperature, i.e.,

$$\xi \equiv \frac{N_+}{N_0} \quad (9\cdot11\cdot10)$$

But by virtue of (9·11·3)

$$N_+ = N_- = N_0\xi$$

and

$$N_H = N_0 - N_0\xi = N_0(1 - \xi) \approx N_0$$

since $\xi \ll 1$. Then the law of mass action (9·11·4) gives, by (9·11·9),

$$\xi^2 = \left(\frac{V}{N_0}\right) \left(\frac{2\pi mkT}{h^2}\right)^{\frac{3}{2}} e^{-\epsilon_0/kT} \quad (9\cdot11\cdot11)$$

so that the degree of dissociation can be readily calculated. Note that Planck's constant appears quite explicitly in this relation.

Vapor pressure of a solid Consider a solid consisting of monatomic molecules, e.g., solid argon. If it is in equilibrium with its vapor the equilibrium condition is, by (8·8·8),

$$\mu_1 = \mu_2 \quad (9\cdot11\cdot12)$$

where μ_1 is the chemical potential of the vapor and μ_2 that of the solid. Unless the temperature is exceedingly high, the vapor is not too dense and can be treated as an ideal gas. Then for the chemical potential of N_1 atoms of vapor in a volume V_1 , (9·11·2) and (9·10·7) give

$$\mu_1 = -kT \ln \left[\frac{V_1}{N_1} \left(\frac{2\pi mkT}{h^2} \right)^{\frac{3}{2}} \right] \quad (9\cdot11\cdot13)$$

Here we have assumed, for simplicity, that the atoms of mass m have no spin degrees of freedom.

Let us now turn to a discussion of the solid. If it consists of N_2 atoms and has a volume V_2 , its chemical potential is related to its partition function Z by

$$\mu_2 = \left(\frac{\partial F}{\partial N_2} \right)_{T, V_2} = -kT \left(\frac{\partial \ln Z}{\partial N_2} \right)_{T, V_2} \quad (9 \cdot 11 \cdot 14)$$

Although we could try to calculate Z by using a model such as the Einstein model of Sec. 7 · 7, let us keep the discussion more general and relate Z directly to specific heat information. The mean energy of the solid is related to Z by

$$\bar{E}(T) = - \left(\frac{\partial \ln Z}{\partial \beta} \right)_V = kT^2 \left(\frac{\partial \ln Z}{\partial T} \right)_V$$

This can be integrated immediately to give

$$\ln Z(T) - \ln Z(T_0) = \int_{T_0}^T \frac{\bar{E}(T')}{kT'^2} dT' \quad (9 \cdot 11 \cdot 15)$$

Here we shall choose $T_0 \rightarrow 0$.

Since the solid is almost incompressible, its volume V_2 is very nearly constant, and its thermodynamic functions are essentially only functions of T . Let us denote by $c(T)$ the specific heat per atom of the solid. (It matters little whether it is measured at constant volume or constant pressure, since the solid is nearly incompressible.) Since $(\partial \bar{E}/\partial T)_V = N_2 c$, we can express $\bar{E}(T)$ in terms of the specific heat. Thus

$$\bar{E}(T) = -N_2\eta + N_2 \int_0^T c(T'') dT'' \quad (9 \cdot 11 \cdot 16)$$

Here we have put $\bar{E}(0) \equiv -N_2\eta$. This is simply the ground-state energy of the solid measured from the same standard state as that of the vapor, i.e., from the state where all atoms are at rest at very large distances from each other. Thus η is the latent heat of sublimation per atom at $T = 0$.

Finally we note that as $T \rightarrow 0$ or $\beta \rightarrow \infty$

$$\begin{aligned} Z &= \Sigma e^{-\beta E_r} \rightarrow \Omega_0 e^{-\beta(-N_2\eta)} \\ \text{or} \quad \ln Z(T_0) &= \frac{N_2\eta}{kT_0} \quad \text{as } T_0 \rightarrow 0 \end{aligned} \quad (9 \cdot 11 \cdot 17)$$

since the number of states Ω_0 accessible to the solid in its ground state is of the order of unity.* (The atoms were assumed to have no spin degrees of freedom which might lead to many states at $T_0 = 0$.) Using (9 · 11 · 16) and (9 · 11 · 17) in (9 · 11 · 15), and putting $T_0 \rightarrow 0$, one obtains

$$\ln Z(T) = \frac{N_2\eta}{kT} + N_2 \int_0^T \frac{dT'}{kT'^2} \int_0^{T'} c(T'') dT'' \quad (9 \cdot 11 \cdot 18)$$

Hence (9 · 11 · 14) yields

$$\mu_2(T) = -\eta - T \int_0^T \frac{dT'}{kT'^2} \int_0^{T'} c(T'') dT'' \quad (9 \cdot 11 \cdot 19)$$

* That is, the entropy $S = k \ln \Omega$ of the solid vanishes as $T \rightarrow 0$, in accordance with the third law.

The equilibrium condition (9·11·12) becomes then

$$\ln \left[\frac{V_1}{N_1} \left(\frac{2\pi mkT}{h^2} \right)^{\frac{3}{2}} \right] = - \frac{\mu_2(T)}{kT} \quad (9 \cdot 11 \cdot 20)$$

To find the vapor pressure \bar{p} , one needs only to use the ideal gas equation of state $\bar{p}V_1 = N_1 kT$ for the vapor. Thus (9·11·20) becomes

$$\ln \left[\frac{kT}{\bar{p}} \left(\frac{2\pi mkT}{h^2} \right)^{\frac{3}{2}} \right] = - \frac{\mu_2}{kT}$$

$$\text{Hence } \ln \bar{p} = \ln \left[\frac{(2\pi m)^{\frac{3}{2}}}{h^3} (kT)^{\frac{3}{2}} \right] + \frac{\mu_2}{kT}$$

$$\text{and } \bar{p}(T) = \frac{(2\pi m)^{\frac{3}{2}}}{h^3} (kT)^{\frac{3}{2}} \exp \left[- \frac{\eta}{kT} - \frac{1}{k} \int_0^T \frac{dT'}{T'^2} \int_0^{T'} c(T'') dT'' \right] \quad (9 \cdot 11 \cdot 21)$$

This is the desired expression for the vapor pressure. Note that it again involves Planck's constant in an essential way.

The specific heat of the solid can be obtained either from microscopic calculation by some model (e.g., the Einstein model of Sec. 7·7) or from experimental measurements. The double integral in (9·11·21) is seen to be a positive increasing function of T ; it converges without difficulty, since $c \rightarrow 0$ sufficiently rapidly as $T \rightarrow 0$.

Note that if we had tried to compute the vapor pressure by means of the Clapeyron equation in a manner similar to that used at the end of Sec. 8·5, we could not have determined the constant of integration, i.e., all the constants on the right side of (9·11·21).

*9·12 Partition functions of polyatomic molecules

Let us sketch briefly how one goes about calculating the partition function for an ideal gas consisting of N polyatomic molecules. In the classical limit where the mean de Broglie wavelength $\bar{\lambda}$ associated with the momentum of the center-of-mass motion is small compared to the mean separation of the molecules one has again

$$Z = \frac{\zeta^N}{N!} \quad (9 \cdot 12 \cdot 1)$$

Here

$$\zeta = \sum_s e^{-\beta \epsilon(s)} \quad (9 \cdot 12 \cdot 2)$$

is the partition function for an individual molecule, the summation being over all the quantum states s of the molecule. To a good approximation one can write the Hamiltonian of a molecule in the additive form

$$\mathcal{H} = \mathcal{H}_t + \mathcal{H}_e + \mathcal{H}_r + \mathcal{H}_v \quad (9 \cdot 12 \cdot 3)$$

and correspondingly the energy levels of the molecule in the form

$$\epsilon(s) = \epsilon_t(s_t) + \epsilon_e(s_e) + \epsilon_r(s_r) + \epsilon_v(s_v) \quad (9 \cdot 12 \cdot 4)$$

\mathcal{H}_t denotes the Hamiltonian describing the translational motion of the center of mass of the molecule; $\epsilon_t(s_t)$ denotes the corresponding translational energy of the translational state labeled s_t .

\mathcal{H}_e denotes the Hamiltonian describing the motion of the electrons about the nuclei assumed in a fixed configuration; $\epsilon_e(s_e)$ denotes the corresponding electronic state labeled s_e .

\mathcal{H}_r denotes the Hamiltonian describing the rotation of the nuclei of the molecule about their center of mass; $\epsilon_r(s_r)$ denotes the corresponding rotational energy of the rotational state labeled s_r .

\mathcal{H}_v denotes the Hamiltonian describing the vibrational motion of the nuclei of the molecule relative to each other; $\epsilon_v(s_v)$ denotes the corresponding vibrational energy of the vibrational state labeled by s_v .

The additivity of (9·12·4) implies immediately that the partition function ζ factors into a product; i.e.,

$$\begin{aligned}\zeta &= \sum_{s_t, s_e, \dots} e^{-\beta[\epsilon_t(s_t) + \epsilon_e(s_e) + \epsilon_r(s_r) + \epsilon_v(s_v)]} \\ &= \left(\sum_{s_t} e^{-\beta\epsilon_t(s_t)}\right) \left(\sum_{s_e} e^{-\beta\epsilon_e(s_e)}\right) \left(\sum_{s_r} e^{-\beta\epsilon_r(s_r)}\right) \left(\sum_{s_v} e^{-\beta\epsilon_v(s_v)}\right)\end{aligned}$$

or (9·12·5)

$$\zeta = \zeta_t \zeta_e \zeta_r \zeta_v$$

where ζ_t is the partition function for the translational motion of the center of mass, ζ_e is the partition function for electronic motion, etc.

Let us discuss these partition functions specifically for a diatomic molecule with atoms of masses m_1 and m_2 .

Translational motion of the center of mass The center of mass moves like a particle of mass $m_1 + m_2$. Thus

$$\mathcal{H}^{(t)} = \frac{\mathbf{p}^2}{2(m_1 + m_2)}$$

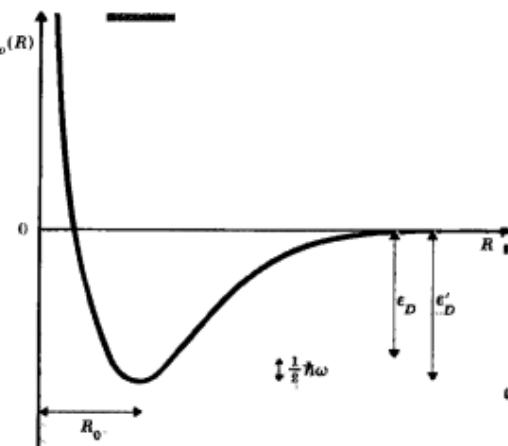
where \mathbf{p} denotes the linear momentum of the center of mass. Using the mass $m_1 + m_2$, the translational states are then the same as those discussed in connection with the monatomic gas. Hence the sum over translational states gives by comparison with (9·10·7)

$$\zeta_t = \frac{V}{h^3} [2\pi(m_1 + m_2)kT]^{\frac{3}{2}} \quad (9·12·6)$$

Electronic motion We turn next to the internal motion of the atoms relative to the center of mass. Consider first the possible electronic states of the molecule. For *fixed* nuclei the electronic ground-state energy ϵ_{e0} can be calculated as a function of the internuclear separation R and yields a curve of the type shown in Fig. 9·12·1.

The minimum of this curve determines, for the electronic ground state of the molecule, the equilibrium internuclear separation R_0 , where $\epsilon_{e0} = -\epsilon_D'$.

Fig. 9·12·1 Energy of the electronic ground state $\epsilon_{e0}(R)$ of a diatomic molecule as a function of the internuclear separation R . The dissociation energy is denoted by ϵ_D , the vibrational zero-point energy by $\frac{1}{2}\hbar\omega$.



This energy is negative when measured with respect to a standard state where the nuclei are at rest at infinite separation from each other. Since the first excited electronic state is, for almost all molecules, higher than the ground state by an energy of the order of a few electron volts, i.e., very large compared to kT , all terms in the electronic partition function other than the one of lowest energy are negligible. (That is, the molecule is with overwhelming probability in its electronic ground state.) Thus one has simply

$$\zeta_e = \Omega_0 e^{\beta \epsilon_D} \quad (9\cdot12\cdot7)$$

where Ω_0 is the degree of degeneracy (if any) of the electronic ground state.

Rotation Consider now the rotation of the molecule. This is in first approximation like the rotation of a rigid dumbbell consisting of two masses m_1 and m_2 separated by the atomic equilibrium distance R_0 in the molecule. The moment

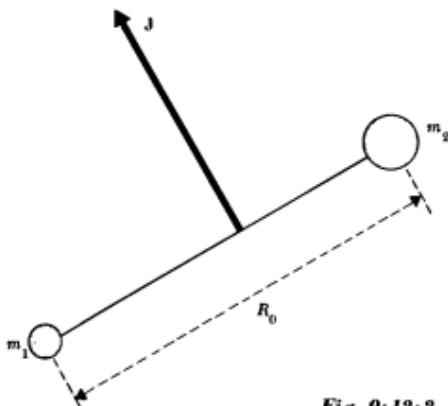


Fig. 9·12·2 Rotation of a rigid dumbbell molecule.

of inertia A of the molecule about an axis through the center of mass and perpendicular to the line joining the atomic nuclei is given by

$$A = \frac{1}{2}\mu^*R_0^2 \quad (9 \cdot 12 \cdot 8)$$

where μ^* is the reduced mass of the two atoms

$$\mu^* = \frac{m_1 m_2}{m_1 + m_2} \quad (9 \cdot 12 \cdot 9)$$

If $\hbar J$ denotes the rotational angular momentum of this dumbbell, its classical energy is given by $(\hbar J)^2/2A$. Quantum mechanically J^2 can assume the possible values $J(J+1)$, where the quantum number $J = 0, 1, 2, 3, \dots$. Hence the rotational energy levels are given by

$$\epsilon_r = \frac{\hbar^2}{2A} J(J+1) \quad (9 \cdot 12 \cdot 10)$$

(Note that a small moment of inertia implies a large spacing of the rotational energy levels.) The vector J can, of course, have several discrete spatial orientations labeled by m_J , its projection along some axis. The possible values of m_J are

$$m_J = -J, -J+1, \dots, (J-1), J$$

so that for each value of J , there are $2J+1$ possible quantum states of the same energy (9·12·10). The rotational partition function becomes then

$$\xi_r = \sum_{J=0}^{\infty} (2J+1) e^{-(\beta \hbar^2/2A)J(J+1)} \quad (9 \cdot 12 \cdot 11)$$

The significant parameter in (9·12·11) is the argument of the exponential, i.e., the ratio of rotational to thermal energy. For low temperature T or small moment of inertia, $\hbar^2/(2AkT) \gg 1$; then practically all molecules are in the very lowest rotational states, and all terms in the sum (9·12·11) beyond the first few are negligible.

Remark Note that in writing down (9·12·11) we have not worried about any angular-momentum component parallel to the axis of the dumbbell. The reason is that the moment of inertia about this axis is very small. Any state with such an angular-momentum component different from zero would, in analogy to (9·12·10), have very high energy compared to kT and can therefore be neglected.

On the other hand, suppose that the temperature T is reasonably large and the moment of inertia is not too small, so that $\hbar^2 J(J+1)(2AkT)^{-1} \ll 1$. (This is the case for many diatomic molecules in which the spacing between the rotational energy levels (9·12·10) is of the order of 10^{-4} ev. Exceptions are molecules such as H_2 below room temperature, because these have such small moments of inertia.) Then the spacing of rotational-energy levels is small compared to kT . This implies that the rotation of the molecule could then also be treated by classical statistical mechanics. Mathematically this means

that successive terms in the sum (9.12.11) differ from each other by relatively small amounts so that this sum can be approximated by an integral. Thus one can write, putting $u = J(J + 1)$,

$$\xi_r \approx \int_0^{\infty} du e^{-(\beta \hbar \nu_{12A})u} = \frac{2A}{\beta \hbar^2}$$

or $\xi_r \approx \frac{2AkT}{\hbar^2} \quad (9.12.12)$

If the two nuclei of the molecule are identical, then we must again be concerned about their essential indistinguishability (just as we were concerned about the factor $N!$ in the translational partition function). In the classical limit, where (9.12.12) is valid, the indistinguishability is easily handled. Turning the molecule end-for-end is the same as interchanging the two identical nuclei. We have counted such a turning over by 180° as a distinct state in calculating (9.12.12), and properly so for unlike nuclei. But it is *not* to be counted as a distinguishable state for *like* nuclei; in this case (9.12.12) is too large by a factor of 2. Hence one should generally put

$$\xi_r = \frac{2AkT}{\hbar^2 \sigma} \quad (9.12.13)$$

where $\sigma = \begin{cases} 1 & \text{if the nuclei are unlike} \\ 2 & \text{if they are identical} \end{cases} \quad (9.12.14)$

In the case where the quasi-classical treatment of rotation is not applicable (e.g., for H_2 at low temperatures) the situation is more complicated and leads to the involvement of nuclear spins with the rotation in a very intimate way. We shall forego discussing the interesting peculiarities arising in such cases.

Remark In the classical limit where (9.12.13) is applicable

$$\ln \xi_r = -\ln \beta + \text{constant}$$

Hence the mean energy of rotation is given by

$$\epsilon_r = -\frac{\partial}{\partial \beta} \ln \xi_r = \frac{1}{\beta} = kT \quad (9.12.15)$$

This is indeed what one would get from the classical equipartition theorem applied to the two degrees of freedom which represent classical rotation, namely, rotation about the two orthogonal principal axes which are perpendicular to the line joining the two nuclei. (We already mentioned in our last remark that the rotation *about* the line joining the nuclei cannot be treated in the classical limit.)

Vibration Finally, the nuclei are also free to vibrate relative to each other about their equilibrium separation R_0 . The potential energy of the nuclei as a function of their separation R is given by the electronic ground-state energy $\epsilon_{el}(R)$ of Fig. 9.12.1. Near its minimum it can be expanded in the form

$$\epsilon_{el}(R) = -\epsilon_D' + \frac{1}{2}b\xi^2 \quad (9.12.16)$$

where $b = \frac{\partial^2 \epsilon_{\text{el}}(R_0)}{\partial R^2}$ and $\xi \equiv R - R_0$ (9 · 12 · 17)

The kinetic energy of vibration of the nuclei relative to their center of mass is given by

$$K = \frac{1}{2}\mu^*\dot{R}^2 = \frac{1}{2}\mu^*\xi^2 \quad (9 \cdot 12 \cdot 18)$$

By (9 · 12 · 16) and (9 · 12 · 18), one would obtain classically simple harmonic motion of angular frequency

$$\omega = \sqrt{\frac{b}{\mu^*}} \quad (9 \cdot 12 \cdot 19)$$

Quantum mechanically, (9 · 12 · 16) and (9 · 12 · 18) yield the Hamiltonian of a simple harmonic oscillator whose possible vibrational energy levels are given by

$$\epsilon_v = \hbar\omega(n + \frac{1}{2}) \quad (9 \cdot 12 \cdot 20)$$

Here the possible quantum states are labeled n , which can assume all values $n = 0, 1, 2, 3, \dots$

Hence the vibrational partition function is

$$\xi_v = \sum_{n=0}^{\infty} e^{-\beta\hbar\omega(n+\frac{1}{2})} \quad (9 \cdot 12 \cdot 21)$$

We already evaluated this simple geometric series in (7 · 6 · 10). Thus

$$\xi_v = \frac{e^{-\frac{1}{2}\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}} \quad (9 \cdot 12 \cdot 22)$$

For most diatomic molecules at ordinary temperatures $\hbar\omega$ is so large (of the order of 0.1 ev) that $\beta\hbar\omega \ll 1$. In that case (9 · 12 · 21) reduces to its first term

$$\xi_v \approx e^{-\frac{1}{2}\beta\hbar\omega}$$

The vibrational degrees of freedom can then certainly *not* be treated classically.

Remark Note that even at $T = 0$ the nuclei still have a zero-point energy $\frac{1}{2}\hbar\omega$ in their lowest vibrational state. Hence ϵ_D' in Fig. 9 · 12 · 1 is not equal to the dissociation energy ϵ_D which must be provided at $T = 0$ to dissociate the molecule into two atoms at rest at an infinite distance from each other. Instead, one has (see Fig. 9 · 12 · 1)

$$\epsilon_D = \epsilon_D' - \frac{1}{2}\hbar\omega \quad (9 \cdot 12 \cdot 23)$$

We have now calculated all the essential ingredients necessary for the evaluation of the partition function (9 · 12 · 5) for an ideal gas of diatomic molecules. (Some examples of applications will be found in the problems.) If the nuclei of the molecules have spin, then ξ in (9 · 12 · 5) must also be multiplied by the possible number of nuclear-spin states. If one deals with molecules consisting of more than two atoms, the decomposition (9 · 12 · 4) or (9 · 12 · 5) is

in general still valid, but the rotational and vibrational partition functions ζ_r and ζ_v become more complicated.

BLACK - BODY RADIATION

9 · 13 *Electromagnetic radiation in thermal equilibrium inside an enclosure*

Let us consider the electromagnetic radiation (or in quantum-mechanical language, the assembly of photons) which exists in thermal equilibrium inside an enclosure of volume V whose walls are maintained at the absolute temperature T . In this situation photons are continuously absorbed and reemitted by the walls; it is, of course, by virtue of these mechanisms that the radiation inside the container depends on the temperature of the walls. But, as usual, it is not at all necessary to investigate the exact mechanisms which bring about the thermal equilibrium, since the general probability arguments of statistical mechanics suffice to describe the equilibrium situation.

Let us regard the radiation as a collection of photons. These must, of course, be considered as indistinguishable particles. The total number of photons inside the enclosure is not fixed, but depends on the temperature T of the walls. The state s of each photon can be specified, in a manner discussed below, by the magnitude and direction of its momentum and by the direction of polarization of the electric field associated with the photon. The radiation field existing in thermal equilibrium inside the enclosure is completely described if one knows the mean number \bar{n}_s of photons in each possible state. The calculation of this number is precisely the problem already solved in (9 · 3 · 5). The result is the Planck distribution

$$\bar{n}_s = \frac{1}{e^{\beta \epsilon_s} - 1} \quad (9 \cdot 13 \cdot 1)$$

where ϵ_s is the energy of a photon in state s .

To make this result more concrete we have to consider in greater detail how the state of each photon is specified. Since we are dealing with electromagnetic radiation, the electric field $\boldsymbol{\epsilon}$ (or each component thereof) satisfies the wave equation

$$\nabla^2 \boldsymbol{\epsilon} = \frac{1}{c^2} \frac{\partial^2 \boldsymbol{\epsilon}}{\partial t^2} \quad (9 \cdot 13 \cdot 2)$$

This is satisfied by (the real part of) plane wave solutions of the form

$$\boldsymbol{\epsilon} = \mathbf{A} e^{i(\boldsymbol{\kappa} \cdot \mathbf{r} - \omega t)} = \boldsymbol{\epsilon}_0(\mathbf{r}) e^{-i\omega t} \quad (9 \cdot 13 \cdot 3)$$

(where \mathbf{A} is any constant), provided that the wave vector $\boldsymbol{\kappa}$ satisfies the condition

$$\kappa = \frac{\omega}{c}, \quad \kappa \equiv |\boldsymbol{\kappa}| \quad (9 \cdot 13 \cdot 4)$$

Remark Note that the spatial part $\mathbf{E}_0(\mathbf{r})$ on the right side of (9·13·3) satisfies the time independent wave equation

$$\nabla^2 \mathbf{E}_0 + \frac{\omega^2}{c^2} \mathbf{E}_0 = 0$$

which is, for each component of \mathbf{E}_0 , of exactly the same form as the time-independent Schrödinger equation (9·9·8) for a nonrelativistic particle.

If the electromagnetic wave is regarded as quantized, then the associated photon is described in the familiar way as a relativistic particle of energy ϵ and momentum \mathbf{p} given by the familiar relations

$$\left. \begin{aligned} \epsilon &= \hbar\omega \\ \mathbf{p} &= \hbar\mathbf{k} \end{aligned} \right\} \quad (9 \cdot 13 \cdot 5)$$

Thus (9·13·4) implies that

$$|\mathbf{p}| = \frac{\hbar\omega}{c} \quad (9 \cdot 13 \cdot 6)$$

Since an electromagnetic wave satisfies the Maxwell equation $\nabla \cdot \mathbf{E} = 0$, it follows by (9·13·3) that $\mathbf{k} \cdot \mathbf{E} = 0$, i.e., that \mathbf{E} is transverse to the direction of propagation determined by the vector \mathbf{k} . For each \mathbf{k} , there are thus only two possible components of \mathbf{E} , perpendicular to \mathbf{k} , which can be specified. In terms of photons this means that, for each \mathbf{k} , there are two possible photons corresponding to the two possible directions of polarization of the electric field \mathbf{E} .

As in the case of the particle discussed in Sec. 9·9, not all possible values of \mathbf{k} are allowed, but only certain discrete values depending on the boundary conditions. Let us again take the enclosure to be in the form of a parallelepiped with edges L_x, L_y, L_z in length. We suppose that the smallest of these lengths, call it L , is so large that $L \gg \lambda$ where $\lambda = 2\pi/\kappa$ is the longest wavelength of significance in the discussion. Then we can again neglect effects occurring near the walls of the container and can describe the situation in terms of simple traveling waves of the form (9·13·3). To eliminate wall effects it is only necessary to proceed as in Sec. 9·9 by imposing the periodic boundary conditions (9·9·11). The enumeration of possible states is then exactly identical to that in Sec. 9·9, the possible values of \mathbf{k} being those given by (9·9·13).

Let $f(\mathbf{k}) d^3\mathbf{k}$ = the mean number of photons per unit volume, with one specified direction of polarization, whose wave vector lies between \mathbf{k} and $\mathbf{k} + d\mathbf{k}$.

There are, by virtue of (9·9·16), $(2\pi)^{-3} d^3\mathbf{k}$ photon states of this kind per unit volume. Each of these has an energy $\epsilon = \hbar\omega = \hbar c \kappa$. Since the mean number of photons with one definite value of \mathbf{k} in this range is given by (9·13·1),

$$\blacktriangleright \qquad f(\mathbf{k}) d^3\mathbf{k} = \frac{1}{e^{\theta\hbar\omega} - 1} \frac{d^3\mathbf{k}}{(2\pi)^3} \quad (9 \cdot 13 \cdot 7)$$

Obviously $f(\mathbf{k})$ is only a function of $|\mathbf{k}|$.

Let us find the mean number of photons per unit volume of *both* directions of polarization and with angular frequency in the range between ω and $\omega + d\omega$. This is given by summing (9.13.7) over all the volume of κ space contained within the spherical shell of radius $\kappa = \omega/c$ and $\kappa + d\kappa = (\omega + d\omega)/c$ and then multiplying by 2 to include both directions of polarization; i.e., it equals

$$2f(\kappa)(4\pi\kappa^2 d\kappa) = \frac{8\pi}{(2\pi c)^3} \frac{\omega^2 d\omega}{e^{\beta\hbar\omega} - 1} \quad (9.13.8)$$

Let $\bar{u}(\omega; T) d\omega$ denote the mean energy per unit volume (i.e., the mean "energy density") of photons of both directions of polarization in the frequency range between ω and $\omega + d\omega$. Since each photon of this type has an energy $\hbar\omega$, one obtains

$$\bar{u}(\omega; T) d\omega = [2f(\kappa)(4\pi\kappa^2 d\kappa)](\hbar\omega) = \frac{8\pi\hbar}{c^3} f(\kappa) \omega^3 d\omega \quad (9.13.9)$$

or

$$\bar{u}(\omega; T) d\omega = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3 d\omega}{e^{\beta\hbar\omega} - 1} \quad (9.13.10)$$

Note that the significant dimensionless parameter of the problem is

$$\eta \equiv \beta\hbar\omega = \frac{\hbar\omega}{kT} \quad (9.13.11)$$

the ratio of photon energy to thermal energy. Thus one can write \bar{u} in terms of η as

$$\bar{u}(\omega; T) d\omega = \frac{\hbar}{\pi^2 c^3} \left(\frac{kT}{\hbar} \right)^4 \frac{\eta^3 d\eta}{e^\eta - 1} \quad (9.13.12)$$

A plot of \bar{u} as a function of η is shown in Fig. 9.13.1. The curve has a maximum for some value $\eta = \bar{\eta} \approx 3$. Note a simple scaling property. If at tem-

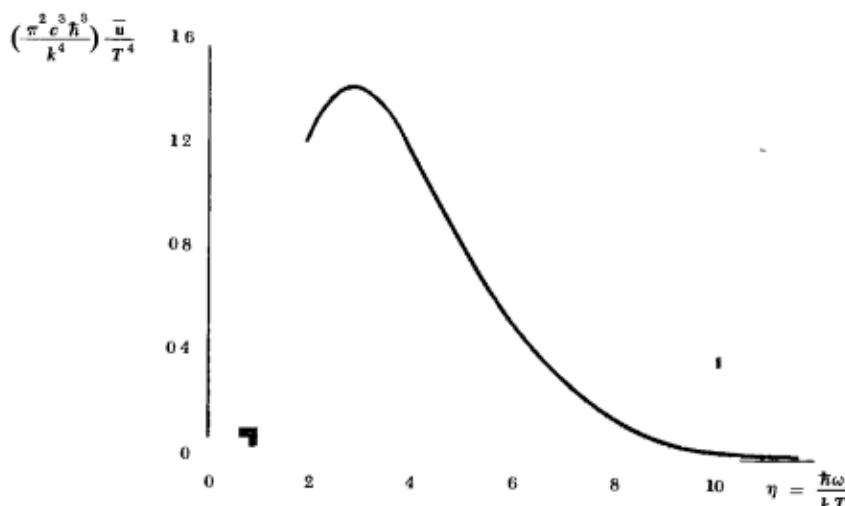


Fig. 9.13.1 The energy density $\bar{u}(\eta)$ (per unit dimensionless frequency range $d\eta$) as a function of $\eta = \hbar\omega/kT$.

perature T_1 the maximum occurs at the angular frequency $\bar{\omega}_1$, and at some other temperature T_2 the maximum occurs at $\bar{\omega}_2$, then one must have

$$\frac{\hbar\bar{\omega}_1}{kT_1} = \frac{\hbar\bar{\omega}_2}{kT_2} = \tilde{\eta}$$

or

$$\frac{\bar{\omega}_1}{T_1} = \frac{\bar{\omega}_2}{T_2} \quad (9 \cdot 13 \cdot 13)$$

This result is known as "Wien's displacement law."

The mean *total* energy density \bar{u}_0 in all frequencies is given by

$$\bar{u}_0(T) = \int_0^{\infty} \bar{u}(T; \omega) d\omega$$

By (9 · 13 · 11) this becomes

$$\bar{u}_0(T) = \frac{\hbar}{\pi^2 c^3} \left(\frac{kT}{\hbar} \right)^4 \int_0^{\infty} \frac{\eta^3 d\eta}{e^\eta - 1} \quad (9 \cdot 13 \cdot 14)$$

The definite integral here is just some constant. Hence one is left with the interesting result that

$$\blacktriangleright \quad \bar{u}_0(T) \propto T^4 \quad (9 \cdot 13 \cdot 15)$$

This statement is called the Stefan-Boltzmann law.

The integral (9 · 13 · 14) can easily be integrated numerically. Although this has no physical importance, it can also be evaluated exactly (see Appendix A · 11). The result is

$$\int_0^{\infty} \frac{\eta^3 d\eta}{e^\eta - 1} = \frac{\pi^4}{15} \quad (9 \cdot 13 \cdot 16)$$

Thus one obtains the explicit expression

$$\bar{u}_0(T) = \frac{\pi^2}{15} \frac{(kT)^4}{(c\hbar)^3} \quad (9 \cdot 13 \cdot 17)$$

The proportionality (9 · 13 · 15) is an obvious result reflecting the fact that space is three-dimensional. This can be seen by the following simple argument. At a temperature T most of the photons must have an energy of the order of kT or less, i.e., they must be photons with wave vector \mathbf{x} of magnitude less than κ' corresponding to an energy

$$\hbar\omega' = \frac{\hbar\kappa'}{c} \approx kT$$

But in three-dimensional space, the number of photon states with $|\mathbf{x}|$ less κ' is proportional to the volume in \mathbf{x} space contained within a sphere of radius κ' . Hence the total mean number \bar{N} of photons at temperature T must be proportional to

$$\bar{N} \propto \kappa'^3 \propto T^3 \quad (9 \cdot 13 \cdot 18)$$

The typical energy of these photons is of the order of kT . Hence it follows that the mean energy density \bar{u}_0 satisfies the proportionality

$$\bar{u}_0 \propto \bar{N}(kT) \propto T^4 \quad (9 \cdot 13 \cdot 19)$$

Calculation of radiation pressure It is of interest to calculate the mean pressure \bar{p} exerted by the radiation on the walls of the enclosure. The pressure contribution from a photon in state s is given by $-\partial\epsilon_s/\partial V$; hence the mean pressure due to all the photons is*

$$\bar{p} = \sum_s \bar{n}_s \left(-\frac{\partial\epsilon_s}{\partial V} \right) \quad (9 \cdot 13 \cdot 20)$$

where \bar{n}_s is given by (9 · 13 · 1). To evaluate $-\partial\epsilon_s/\partial V$, consider for simplicity that the enclosure is a cube of edge lengths $L_x = L_y = L_z \equiv L$ so that its volume is $V = L^3$. With the possible values of κ given by (9 · 9 · 13), one has for a state s specified by the integers n_x, n_y, n_z

$$\epsilon_s = \hbar\omega = \hbar c \kappa = \hbar c (\kappa_x^2 + \kappa_y^2 + \kappa_z^2)^{\frac{1}{2}} = \hbar c \left(\frac{2\pi}{L} \right) (n_x^2 + n_y^2 + n_z^2)^{\frac{1}{2}}$$

or $\epsilon_s = CL^{-1} = CV^{-\frac{1}{3}}, \quad \text{where } C = \text{constant} \quad (9 \cdot 13 \cdot 21)$

Hence $\frac{\partial\epsilon_s}{\partial V} = -\frac{1}{3} CV^{-\frac{4}{3}} = -\frac{1}{3} \frac{\epsilon_s}{V} \quad (9 \cdot 13 \cdot 22)$

Thus (9 · 13 · 20) becomes

$$\bar{p} = \sum_s \bar{n}_s \left(\frac{1}{3} \frac{\epsilon_s}{V} \right) = \frac{1}{3V} \sum_s \bar{n}_s \epsilon_s = \frac{1}{3V} \bar{E}$$

or

► $\bar{p} = \frac{1}{3} \bar{u}_0 \quad (9 \cdot 13 \cdot 23)$

The radiation pressure is thus very simply related to the mean energy density of the radiation.

It is also instructive to calculate the radiation pressure by detailed kinetic arguments similar to those used in Sec. 7 · 13 for computing the mean pressure exerted by a classical gas of particles. Photons impinging upon an element of area dA of the container wall (normal to the z direction) impart to it in unit time a mean z component of momentum $G_z^{(+)}$. In equilibrium, an equal number of photons leaves the wall and gives rise to an equal momentum flow $-G_z^{(+)}$ in the opposite direction. Hence the net force per unit area, or pressure on the wall, is related to the mean rate of change of momentum by

$$\bar{p} = \frac{1}{dA} [G_z^{(+)} - (-G_z^{(+)})] = \frac{2G_z^{(+)}}{dA}$$

Consider, in Fig. 9 · 13 · 2, all photons with wave vector between κ and $\kappa + d\kappa$. There are $2f(\kappa) d^3\kappa$ photons of this kind (of both possible polarizations) per unit volume. Since photons travel with speed c , all photons contained in the cylindrical volume $c dt dA \cos \theta$ strike the area dA in time dt and carry z component of momentum $\hbar\kappa_z$. The total photon momentum arriving at dA per unit time is then

$$G_z^{(+)} = \frac{1}{dt} \int_{\kappa_z > 0} [2f(\kappa) d^3\kappa] (c dt dA \cos \theta) (\hbar\kappa_z)$$

* The same result could be obtained from the general relation $\bar{p} = \beta^{-1}(\partial \ln Z / \partial V)$ by using the partition function for photons in (9 · 5 · 3).

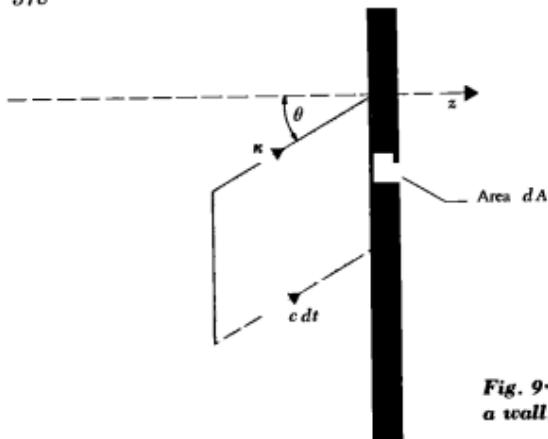


Fig. 9·13·2 Photons impinging upon a wall.

Hence

$$\bar{p} = 2c\hbar \int_{\kappa > 0} [2f(\kappa) d^3\kappa] \frac{\kappa_x^2}{\kappa}$$

where we have put $\cos \theta = \kappa_x/\kappa$. But $f(\kappa)$ depends only on $|\kappa|$, so that the integrand is an even function of κ_x . Thus one can extend the integration over all values of κ and write

$$\bar{p} = c\hbar \int [2f(\kappa) d^3\kappa] \frac{\kappa_x^2}{\kappa} = \frac{1}{3} c\hbar \int [2f(\kappa) d^3\kappa] \frac{(\kappa_x^2 + \kappa_y^2 + \kappa_z^2)}{\kappa}$$

where the last result follows by symmetry, since all directions are equivalent. Since $\kappa_x^2 + \kappa_y^2 + \kappa_z^2 = \kappa^2$, one then obtains

$$\bar{p} = \frac{1}{3} \int [2f(\kappa) d^3\kappa] (c\hbar\kappa) = \frac{1}{3} \bar{u}_0$$

since $c\hbar\kappa$ is simply the energy of a photon of wave vector κ .

9 · 14 Nature of the radiation inside an arbitrary enclosure

The full generality of the results of the preceding section can be made apparent by a few simple physical arguments. Consider an enclosure which has an arbitrary shape and which may contain several bodies within it. Its walls, which may consist of any material, are maintained at an absolute temperature T . The enclosure thus acts as a heat reservoir. From our general arguments of statistical thermodynamics we know that the equilibrium situation of greatest probability, or entropy, is the one where the radiation, as well as the bodies inside the enclosure, are all characterized by the same temperature T .

The nature of the radiation field existing at this temperature T inside the enclosure can be described in terms of

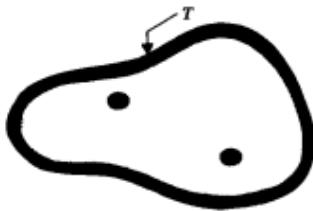


Fig. 9·14·1 Electromagnetic radiation in equilibrium inside an enclosure of arbitrary shape. The radiation must be homogeneous.

$f_\alpha(\kappa, \mathbf{r}) d^3\kappa$ = the mean number of photons per unit volume at the point \mathbf{r} , with wave vector between κ and $\kappa + d\kappa$ and with polarization specified by the index α (i.e., by some unit vector $\mathbf{b}_\alpha \perp \kappa$)

As usual we assume that the dimensions of the enclosure are large compared to the wavelengths $\lambda = 2\pi\kappa^{-1}$ of interest.

If the enclosure is in equilibrium, one can immediately make several general statements about f .

1. The number f is independent of \mathbf{r} ; i.e., the radiation field is homogeneous.

Argument: Suppose that $f_\alpha(\kappa, \mathbf{r})$ were different at two positions in the enclosure. Consider what would happen if two *identical* small bodies at temperature T were placed at these positions. (Imagine these to be surrounded by filters which transmit only frequencies in the specified range $\omega = \omega(\kappa)$ and which transmit only radiation of the specified direction of polarization α .) Since different amounts of radiation would be incident on the two bodies, they would absorb different amounts of energy per unit time and their temperatures would therefore become different. This would contradict the equilibrium condition of maximum entropy according to which the temperature must be uniform throughout the enclosure. Hence

$$f_\alpha(\kappa, \mathbf{r}) = f_\alpha(\kappa) \quad \text{independent of } \mathbf{r}$$

2. The number f is independent of the direction of κ , but depends only on $|\kappa|$; i.e., the radiation field is isotropic.

Argument: Suppose that $f_\alpha(\kappa)$ did depend on the direction κ , e.g., that f is greater if κ points north than if it points east. We could again imagine that two identical small bodies at temperature T (and surrounded by the same filters as before) are introduced into the enclosure as shown in Fig. 9·14·2. Then the body on the north side would have more radiation incident on it and

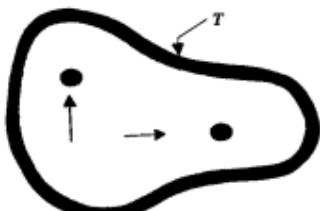


Fig. 9·14·2 The radiation in equilibrium inside the enclosure is isotropic.

thus absorb more power than the body on the east side. This would again lead to a nonpermissible temperature difference being produced between these bodies. Hence we can conclude that

$$f_\alpha(\kappa) = f_\alpha(\kappa), \quad \text{where } \kappa \equiv |\kappa|$$

3. The number f is independent of the direction of polarization of the radiation, i.e., the radiation field in the enclosure is unpolarized.

Argument: Suppose that $f_\alpha(\kappa)$ did depend on the direction of polarization specified by α . Then we could imagine that two identical small bodies at the temperature T are introduced side by side into the enclosure and are surrounded by filters which transmit different directions of polarization. Hence different amounts of radiation would be incident upon these bodies, and a temperature difference would be developed between them in contradiction to the equilibrium condition. Hence

$$f_1(\kappa) = f_2(\kappa)$$

is independent of the polarization index.

4. The function f does not depend on the shape nor volume of the enclosure, nor on the material of which it is made, nor on the bodies it may contain.

Argument: Consider two different enclosures, both at the temperature T , and suppose that the fractions $f_\alpha^{(1)}(\kappa)$ and $f_\alpha^{(2)}(\kappa)$ describing their radiation fields were different. Imagine that we connect the two enclosures through a small hole (containing a filter which transmits only radiation in a narrow frequency range about $\omega(\kappa)$ and of the specified direction of polarization). This would represent an equilibrium situation if both enclosures are at the same temperature T . But if $f^{(1)} > f^{(2)}$, more radiation per unit time would pass from enclosure 1 into enclosure 2 than in the opposite direction. A temperature difference would then develop between the two enclosures, in contradiction to the equilibrium condition of uniform temperature. Hence one concludes that

$$f_\alpha^{(1)}(\kappa) = f_\alpha^{(2)}(\kappa)$$

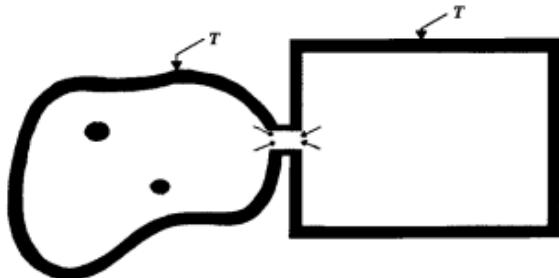


Fig. 9·14·3 Two different enclosures at the same temperature joined through a small hole.

Thus we arrive at the result that in thermal equilibrium $f_{\alpha}(\kappa)$ depends *only* on the temperature T of the enclosure. In particular, it also follows that f is the same for an arbitrary cavity as it is for the rectangular parallelepiped cavity which we used for simplicity in the discussion of Sec. 9·13.

9·15 *Radiation emitted by a body at temperature T*

In the preceding sections we considered the very simple *equilibrium* situation of electromagnetic radiation confined within an enclosure at absolute temperature T . We arrived at a set of interesting results on the basis of very general arguments of statistical mechanics without having to be concerned about the detailed mechanisms of absorption or emission of radiation by the walls. The results of this equilibrium discussion can, however, serve as a basis for treating much more general cases. Consider, for example, a body maintained at some elevated absolute temperature T ; as a concrete example, think of the hot filament of a light bulb hanging from the ceiling. We know that this body emits electromagnetic radiation, and we may be interested in how much energy per unit time (or power) $\Phi_e(\omega) d\omega$ this body emits by radiation in the frequency range between ω and $\omega + d\omega$. The situation envisaged here is certainly *not* one of equilibrium; the walls of the room are at a much lower temperature than the light-bulb filament, and there is a continuous transfer of energy by radiation from the hot filament to the colder walls. It might seem, therefore, that we can no longer use the methods of equilibrium statistical mechanics to discuss this problem and that we need to undertake a detailed investigation of the processes whereby the atoms in the body emit radiation. This would indeed be a formidable problem in quantum mechanics and electromagnetic theory! It is possible, however, to circumvent completely such an analysis by reverting to very clever general arguments based on the equilibrium situation. The method of approach consists of imagining the radiating body to be in an equilibrium situation inside an enclosure containing radiation at this temperature T , and then investigating the conditions that must prevail so that the equilibrium is indeed maintained. The fundamental argument used here is one of "detailed balance"; i.e., one argues that if the body is to remain in equilibrium, then each process of emission by the body must be balanced by an inverse process of absorption of incident radiation. But the radiation incident on the body in an *equilibrium* situation is easily calculated from the results of the preceding sections dealing with the simple case of an ideal gas of photons. Thus one can immediately find the power emitted in such a process *without* engaging in the vastly more complicated calculation of how a collection of interacting atoms in the body emit radiation. Now that we have outlined the general nature of the arguments to be used, let us make them more precise.

Bodies as emitters and absorbers of radiation Consider an arbitrary body at absolute temperature T . The electromagnetic radiation emitted by

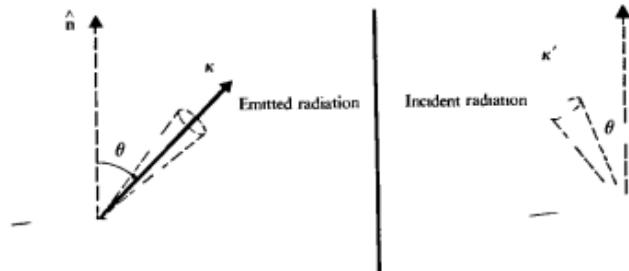


Fig. 9-15-1 Diagram illustrating emission and absorption of radiation by a body.

this body can be described in terms of the energy per unit time, or power, emitted by the body. Specifically one can define its "emissivity" as

$\varphi_e(\kappa; \alpha) d\omega d\Omega$ = the power, per unit area of the body, emitted with polarization α into a range about κ (i.e., into an angular frequency range between ω and $\omega + d\omega$, and into a solid angle $d\Omega$ about the direction κ)

The emissivity depends on the nature of the body and on its temperature.

Having seen how to describe a body as an emitter of radiation, let us now try to describe it as an absorber of radiation. For this purpose, consider radiation of polarization α and with a wave vector in a small range about κ' (i.e., with angular frequency in the range between ω and $\omega + d\omega$ and propagating in a direction lying within the solid-angle range $d\Omega$ about κ'). Suppose that radiation of this type is *incident* on the body so that power $\varphi_i(\kappa', \alpha) d\omega d\Omega$ is incident per unit area of the body. Of this a fraction $a(\kappa', \alpha)$ is absorbed by the body. (By conservation of energy the rest of the incident power is then reflected into various directions if we assume the body to be sufficiently thick that none of the incident radiation is transmitted through it.) The parameter $a(\kappa', \alpha)$ (sometimes called the "absorptivity") is characteristic of the particular body and depends, in general, also on its temperature T . This parameter describes the properties of the body as an absorber of radiation.

The principle of detailed balance We pointed out earlier that it would be quite difficult to calculate directly quantities such as the power φ_e radiated by a body at temperature T . To circumvent this problem, we imagine that the body under consideration is placed inside an enclosure at the same temperature T so as to be in equilibrium with the radiation field existing therein. The characteristics of this radiation field are well known from our previous simple discussion based on equilibrium statistical thermodynamics. Let us now, however, consider more closely the various mechanisms whereby the equilibrium state of the body in this enclosure is actually maintained. Under these circumstances the body emits radiation. On the other hand, radiation is continually incident upon the body which absorbs a certain fraction of it. In the equilibrium situation the energy of the body must remain unchanged. Hence

we can conclude that these processes must balance so that

$$\text{Power radiated by body} = \text{power absorbed by body.} \quad (9 \cdot 15 \cdot 1)$$

We shall, however, want to make statements which are much stronger than this simple condition of over-all energy balance by asserting that the processes which maintain the equilibrium also balance each other in *detail*. For example, it might be conceivable that in one frequency range the body radiates more power than it absorbs, while in another frequency range it radiates less power than it absorbs, in such a way that the over-all energy balance (9·15·1) is preserved. A simple physical argument shows, however, that this cannot be the case. Imagine that the body is surrounded by a shield (a "filter") which absorbs completely all radiation except that, in one small element of area, it is completely transparent to radiation of one direction of polarization and of one narrow frequency range between ω and $\omega + d\omega$. The presence of this shield cannot affect such intrinsic parameters of the body as its emissivity or absorptivity; nor can it, by the arguments of the preceding section, affect the nature of the radiation in the enclosure. Since the equilibrium situation can equally well exist in the presence of the shield, it follows that the energy balance (9·15·1) must hold for this particular element of area, direction of polarization, and frequency range. Since any kind of shield could have been used, one thus arrives at the "principle of *detailed balance*," which asserts that in equilibrium the power radiated and absorbed by the body must be equal for *any* particular element of area of the body, for *any* particular direction of polarization, and for *any* frequency range.



Fig. 9·15·2 A body located inside an enclosure and surrounded by a shield which is only transparent in one small element of area to radiation of one direction of polarization and of one narrow frequency range.

***Microscopic discussion.** The principle of detailed balance is a very fundamental result based on considerations more general than those pertaining to ensembles representing systems in thermal equilibrium. The basic justification of the principle rests on the fundamental laws of microscopic physics, e.g., the Schrödinger equation of quantum mechanics and Maxwell's equations of electromagnetic theory. Consider a single isolated system consisting of several weakly interacting parts (e.g., a body and electromagnetic radiation). In the absence of interaction between these parts the system can be in any one of its quantum states labeled by indices r, s, \dots . The presence of the interaction causes transitions between these states. From the fundamental microscopic laws one can compute the resulting transition probability w_{rs} per unit time from state r to state s . But these microscopic laws are all

invariant under reversal of the time from t to $-t$. Under such a time reversal a state r goes over into a state r^* , etc. (e.g., a state of a particle labeled by its momentum $\mathbf{p} = \hbar\mathbf{k}$ goes over into one labeled by momentum $-\mathbf{p}$). If we call the "reverse" transition the one from state s^* to r^* , then the invariance of the microscopic laws under time reversal implies that

$$w_{s^*r^*} = w_{rs} \quad (9 \cdot 15 \cdot 2)$$

This expresses the "principle of microscopic reversibility." For example, consider the process of emission of a photon with wave vector \mathbf{k} . The reverse process obtained by reversing the sign of the time t is the absorption of a photon of wave vector $-\mathbf{k}$. The microscopic reversibility (9·15·2) asserts that these two processes occur with equal probability.

Once one knows the transition probability for the occurrence of a process in a single system, one can readily calculate the rate of occurrence of this process when one deals with a statistical ensemble of such systems. Let us consider the process of transition from a set A of states labeled by r to some set B of states labeled by s . Let P_r denote the probability in the ensemble that the system is in state r . Then the probability W_{AB} of occurrence of the process $A \rightarrow B$ in the ensemble is given by

$$W_{AB} = \sum_r \sum_s P_r w_{rs} \quad (9 \cdot 15 \cdot 3)$$

Here one sums over all the initial states r in the set A from which the system can start out, each of these states being weighted by the probability that the system is found in this state; then one sums this probability over all the possible set B of final states s . Similarly, one can write for the rate of occurrence of the reverse process

$$W_{B^*A^*} = \sum_{s^*} \sum_{r^*} P_{s^*} w_{s^*r^*} \quad (9 \cdot 15 \cdot 4)$$

But our fundamental statistical postulate asserts that, in an equilibrium situation, an isolated system is found with equal probability in any accessible state of the ensemble. All the probabilities P_r are then equal to the same value P . Hence one obtains by (9·15·2)

$$W_{B^*A^*} = P \sum_{s^*} \sum_{r^*} w_{s^*r^*} = P \sum_s \sum_r w_{rs}$$

so that

$$W_{B^*A^*} = W_{AB} \quad (9 \cdot 15 \cdot 5)$$

This is the principle of detailed balance. In words it asserts that in the statistical ensemble representing a system in equilibrium, the probability of occurrence of *any* process must be equal to the probability of occurrence of the reverse process. By a process we mean transitions from one set of states of the system to another such set of states, the probability of the process being proportional to the number of such transitions occurring per unit time. The reverse process is the one that would result if the sign of the time were reversed, in particular, if all velocities were reversed so that everything went backward in time.