

An Introduction to Statistical Mechanics and Thermodynamics

Second Edition

Robert H. Swendsen

OXFORD GRADUATE TEXTS

AN INTRODUCTION TO STATISTICAL MECHANICS AND THERMODYNAMICS

An Introduction to Statistical Mechanics and Thermodynamics

Second Edition

Robert H. Swendsen



OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP, United Kingdom

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide. Oxford is a registered trade mark of Oxford University Press in the UK and in certain other countries

© Robert H. Swendsen 2020

The moral rights of the author have been asserted

First Edition published in 2012 Second Edition published in 2020

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without the prior permission in writing of Oxford University Press, or as expressly permitted by law, by licence or under terms agreed with the appropriate reprographics rights organization. Enquiries concerning reproduction outside the scope of the above should be sent to the Rights Department, Oxford University Press, at the address above

> You must not circulate this work in any other form and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press 198 Madison Avenue, New York, NY 10016, United States of America

> British Library Cataloguing in Publication Data Data available

Library of Congress Control Number: 2019945687

ISBN 978-0-19-885323-7

DOI: 10.1093/oso/9780198853237.001.0001

Printed and bound by CPI Group (UK) Ltd, Croydon, CR0 4YY

Links to third party websites are provided by Oxford in good faith and for information only. Oxford disclaims any responsibility for the materials contained in any third party website referenced in this work. To the memory of Herbert B. Callen, physicist and mentor, and to my wife, Roberta L. Klatzky, without whom this book could never have been written.

Preface to the Second Edition

In preparing to write the preface to the second edition of this book, I realized anew my debt to the work of Herbert B. Callen. He was not only my thesis advisor and my friend, but it was through his teaching and his book on thermodynamics that I first understood the subject in any depth. I take this opportunity once again to acknowledge how much his pedagogy and advice have meant to my work.

The postulational approach to thermodynamics, which is primarily based on his work and that of his thesis advisor, László Tisza, provides a clear basis for the theory. It is not difficult to understand but can seem rather abstract when first encountered as a student – as, indeed, it did to me many years ago. Many professors have told me that they thought that Callen's book was too daunting to give to their students, but that it was the book that they consulted for thermodynamics.

Part I of my book originated as an introduction to Callen's *Thermodynamics* in my teaching. One difficulty that I had found as a student was that Callen's book started off presenting entropy and the postulates of thermodynamics in the first chapter, and temperature as a partial derivative of the entropy in the second chapter. I had only a vague idea at the time of what the entropy was, and its partial derivative with respect to energy was a complete mystery. I have tried to avoid this difficulty in my own teaching of thermodynamics by presenting the students with an explicit calculation of the entropy of a classical ideal gas. All assumptions are stated, and all mathematics is explained. I felt – and my students generally agreed – that they were then ready to understand Callen's postulates.

Part II developed from my notes for teaching from Callen's textbook. I found that while the ideas in Callen's postulates provided a great foundation for thermodynamics, their specific form was less than ideal. For the first edition of this book, I separated them into six new postulates, each of which expressed a separate idea. I also generalized the postulates to include non-homogeneous systems.

I gave an explicit guide to the use of Jacobians in deriving thermodynamic identities, which I have not found anywhere else, but which my students have found to be easy to apply. Callen mentioned Jacobians in his first edition, but not in his second. Similarly, I simplified the derivation of Maxwell relations, with the result that my students have regarded them (correctly) as being easy to derive.

I also gave an explicit derivation of the stability criteria for second partial derivatives with respect to intensive variables because many students had difficulty with them.

Parts III (classical statistical mechanics) and IV (quantum statistical mechanics) used computer calculations extensively. They allowed many calculations to be carried out explicitly. I firmly believe that the future of physics will rely heavily on the computer, and I think that computation is currently being neglected in university curricula.

The second edition has come into being because I have discovered how to clarify the presentation of many of the central concepts, especially in the derivation of the entropy in Part 1. Along the way, I have corrected a significant number of typographical errors.

In Part I, Chapters 4 and 6, I have more clearly distinguished generic variables from variables describing particular systems used in derivations. My previous labeling convention did not cause any problems in the classes I taught, but it has caused confusion with some readers. I have also generalized the derivation of the entropy from treating only two systems at a time to deriving the entropy simultaneously for all systems that might interact.

In the second edition, I have again changed the list of postulates to include the possibility of negative temperatures. Callen had mentioned negative temperatures in his book, but had excluded them in the interest of simplicity.

In Chapter 11, I have expanded the review of the Carnot cycle with two new illustrations. This chapter now also contains a discussion of negative temperatures, and how they affect the analysis of heat engines.

Massieu functions were mentioned by Callen, but not developed. I did the same in the first edition. I have expanded the treatment of Massieu functions in Chapter 12, after realizing that they are much more useful than I had previously thought. They are essential when considering negative temperatures because the corresponding entropy is not monotonic.

The discussion of the Nernst Postulate (Third Law of Thermodynamics) in Chapter 18 includes a discussion of why zero temperature would not be possible to attain if classical mechanics were valid instead of quantum mechanics. In fact, it would be *more* difficult to attain very low temperatures if the Nernst Postulate were not valid.

A new chapter (Chapter 21) has been added to discuss the consequences of including the widths of the energy and particle-number distributions in the calculation of the entropy. It is both a more realistic assumption and gives better expressions for the entropy. These results are based on new work since the publication of the first edition of this book.

In Chapters 28 on Bose-Einstein statistics and 29 on Fermi-Dirac statistics, I've introduced numerical calculations based on work with a former student, Tyson Price. The numerical results show many of the thermal properties of Bose and Fermi gases more clearly and simply than would be possible with analytic calculations alone.

The Index has been thoroughly updated and expanded.

My recommendations for a programming language to use for the computational problems have changed. I still advocate the use of Python, although not VPython. I have found that plots using MatPlotLib are much better, as well as being easier for students (and professors) to program. On the other hand, I have found that students prefer the freedom to use a wide variety of programming languages, and I have never insisted that they use Python.

I would like to thank my colleagues, Markus Deserno and Michael Widom, for their helpful comments based on their own experiences from using my book to teach both undergraduate and graduate courses in thermal physics.

I would also like to thank my former students, William Griffin, Lachlan Lancaster, and Michael Matty, for their contributions to some of the results presented here. I would

especially like to thank Michael Matty for his extensive constructive criticism of the text and his contributions to my class. Finally, I would like to thank Karpur Shukla for many useful conversations.

As in the first preface, I would like to thank my wife, Roberta L. Klatzky, for her unwavering support.

Robert H. Swendsen Pittsburgh, April 2019

Preface to the First Edition

Habe Muth dich deines eigenen Verstandes zu bedienen. (Have the courage to think for yourself.)

Immanuel Kant, in Beantwortung der Frage: Was ist Aufklärung?

The disciplines of statistical mechanics and thermodynamics are very closely related, although their historical roots are separate. The founders of thermodynamics developed their theories without the advantage of contemporary understanding of the atomic structure of matter. Statistical mechanics, which is built on this understanding, makes predictions of system behavior that lead to thermodynamic rules. In other words, statistical mechanics is a conceptual precursor to thermodynamics, although it is a historical latecomer.

Unfortunately, despite their theoretical connection, statistical mechanics and thermodynamics are often taught as separate fields of study. Even worse, thermodynamics is usually taught first, for the dubious reason that it is older than statistical mechanics. All too often the result is that students regard thermodynamics as a set of highly abstract mathematical relationships, the significance of which is not clear.

This book is an effort to rectify the situation. It presents the two complementary aspects of thermal physics as a coherent theory of the properties of matter. My intention is that after working through this text a student will have solid foundations in both statistical mechanics and thermodynamics that will provide direct access to modern research.

Guiding Principles

In writing this book I have been guided by a number of principles, only some of which are shared by other textbooks in statistical mechanics and thermodynamics.

- I have written this book for students, not professors. Many things that experts might take for granted are explained explicitly. Indeed, student contributions have been essential in constructing clear explanations that do not leave out 'obvious' steps that can be puzzling to someone new to this material.
- The goal of the book is to provide the student with conceptual understanding, and the problems are designed in the service of this goal. They are quite challenging, but the challenges are primarily conceptual rather than algebraic or computational.

- I believe that students should have the opportunity to program models themselves and observe how the models behave under different conditions. Therefore, the problems include extensive use of computation.
- The book is intended to be accessible to students at different levels of preparation. I do not make a distinction between teaching the material at the advanced undergraduate and graduate levels, and indeed, I have taught such a course many times using the same approach and much of the same material for both groups. As the mathematics is entirely self-contained, students can master all of the material even if their mathematical preparation has some gaps. Graduate students with previous courses on these topics should be able to use the book with self-study to make up for any gaps in their training.
- After working through this text, a student should be well prepared to continue with more specialized topics in thermodynamics, statistical mechanics, and condensed-matter physics.

Pedagogical Principles

The over-arching goals described above result in some unique features of my approach to the teaching of statistical mechanics and thermodynamics, which I think merit specific mention.

Teaching Statistical Mechanics

- The book begins with *classical* statistical mechanics to postpone the complications of quantum measurement until the basic ideas are established.
- I have defined ensembles in terms of probabilities, in keeping with Boltzmann's vision. In particular, the discussion of statistical mechanics is based on Boltzmann's 1877 definition of entropy. This is not the definition usually found in textbooks, but what he actually wrote. The use of Boltzmann's definition is one of the key features of the book that enables students to obtain a deep understanding of the foundations of both statistical mechanics and thermodynamics.
- A self-contained discussion of probability theory is presented for both discrete and continuous random variables, including all material needed to understand basic statistical mechanics. This material would be superfluous if the physics curriculum were to include a course in probability theory, but unfortunately, that is not usually the case. (A course in statistics would also be very valuable for physics students—but that is another story.)
- Dirac delta functions are used to formulate the theory of continuous random variables, as well as to simplify the derivations of densities of states. This is not the way mathematicians tend to introduce probability densities, but I believe that it is by far the most useful approach for scientists.

- Entropy is presented as a logical consequence of applying probability theory to systems containing a large number of particles, instead of just an equation to be memorized.
- The entropy of the classical ideal gas is derived in detail. This provides an explicit example of an entropy function that exhibits all the properties postulated in thermodynamics. The example is simple enough to give every detail of the derivation of thermodynamic properties from statistical mechanics.
- The book includes an explanation of Gibbs' paradox—which is not really paradoxical when you begin with Boltzmann's 1877 definition of the entropy.
- The apparent contradiction between observed irreversibility and time-reversalinvariant equations of motion is explained. I believe that this fills an important gap in a student's appreciation of how a description of macroscopic phenomena can arise from statistical principles.

Teaching Thermodynamics

- The four fundamental postulates of thermodynamics proposed by Callen have been reformulated. The result is a set of six thermodynamic postulates, sequenced so as to build conceptual understanding.
- Jacobians are used to simplify the derivation of thermodynamic identities.
- The thermodynamic limit is discussed, but the validity of thermodynamics and statistical mechanics does not rely on taking the limit of infinite size. This is important if thermodynamics is to be applied to real systems, but is sometimes neglected in textbooks.
- My treatment includes thermodynamics of non-extensive systems. This allows me to include descriptions of systems with surfaces and systems enclosed in containers.

Organization and Content

The principles I have described above lead me to an organization for the book that is quite different from what has become the norm. As was stated above, while most texts on thermal physics begin with thermodynamics for historical reasons, I think it is far preferable from the perspective of pedagogy to begin with statistical mechanics, including an introduction to those parts of probability theory that are essential to statistical mechanics.

To postpone the conceptual problems associated with quantum measurement, the initial discussion of statistical mechanics in Part I is limited to classical systems. The entropy of the classical ideal gas is derived in detail, with a clear justification for every step. A crucial aspect of the explanation and derivation of the entropy is the use of Boltzmann's 1877 definition, which relates entropy to the probability of a macroscopic state. This definition provides a solid, intuitive understanding of what entropy is all about. It is my experience that after students have seen the derivation of the entropy of the classical

ideal gas, they immediately understand the postulates of thermodynamics, since those postulates simply codify properties that they have derived explicitly for a special case.

The treatment of statistical mechanics paves the way to the development of thermodynamics in Part II. While this development is largely based on the classic work by Herbert Callen (who was my thesis advisor), there are significant differences. Perhaps the most important is that I have relied entirely on Jacobians to derive thermodynamic identities. Instead of regarding such derivations with dread—as I did when I first encountered them—my students tend to regard them as straightforward and rather easy. There are also several other changes in emphasis, such as a clarification of the postulates of thermodynamics and the inclusion of non-extensive systems; that is, finite systems that have surfaces or are enclosed in containers.

Part III returns to classical statistical mechanics and develops the general theory directly, instead of using the common roundabout approach of taking the classical limit of quantum statistical mechanics. A chapter is devoted to a discussion of the apparent paradoxes between microscopic reversibility and macroscopic irreversibility.

Part IV presents quantum statistical mechanics. The development begins by considering a probability distribution over all quantum states, instead of the common *ad hoc* restriction to eigenstates. In addition to the basic concepts, it covers black-body radiation, the harmonic crystal, and both Bose and Fermi gases. Because of their practical and theoretical importance, there is a separate chapter on insulators and semiconductors. The final chapter introduces the Ising model of magnetic phase transitions.

The book contains about a hundred multi-part problems that should be considered as part of the text. In keeping with the level of the text, the problems are fairly challenging, and an effort has been made to avoid 'plug and chug' assignments. The challenges in the problems are mainly due to the probing of essential concepts, rather than mathematical complexities. A complete set of solutions to the problems is available from the publisher.

Several of the problems, especially in the chapters on probability, rely on computer simulations to lead students to a deeper understanding. In the past I have suggested that my students use the C++ programming language, but for the last two years I have switched to VPython for its simplicity and the ease with which it generates graphs. An introduction to the basic features of VPython is given in in Appendix A. Most of my students have used VPython, but a significant fraction have chosen to use a different language—usually Java, C, or C++. I have not encountered any difficulties with allowing students to use the programming language of their choice.

Two Semesters or One?

The presentation of the material in this book is based primarily on a two-semester undergraduate course in thermal physics that I have taught several times at Carnegie Mellon University. Since two-semester undergraduate courses in thermal physics are rather unusual, its existence at Carnegie Mellon for several decades might be regarded as surprising. In my opinion, it should be the norm. Although it was quite reasonable to teach two semesters of classical mechanics and one semester of thermodynamics to undergraduates in the nineteenth century—the development of statistical mechanics was just beginning—it is not reasonable in the twenty-first century.

However, even at Carnegie Mellon only the first semester of thermal physics is required. All physics majors take the first semester, and about half continue on to the second semester, accompanied by a few students from other departments. When I teach the course, the first semester covers the first two parts of the book (Chapters 1 through 18), plus an overview of classical canonical ensembles (Chapter 18) and quantum canonical ensembles (Chapter 22). This gives the students an introduction to statistical mechanics and a rather thorough knowledge of thermodynamics, even if they do not take the second semester.

It is also possible to teach a one-semester course in thermal physics from this book using different choices of material. For example:

- If the students have a strong background in probability theory (which is, unfortunately, fairly rare), Chapters 3 and 5 might be skipped to include more material in Parts III and IV.
- If it is decided that students need a broader exposure to statistical mechanics, but that a less detailed study of thermodynamics is sufficient, Chapters 14 through 17 could be skimmed to have time to study selected chapters in Parts III and IV.
- If the students have already had a thermodynamics course (although I do not recommend this course sequence), Part II could be skipped entirely. However, even if this choice is made, students might still find Chapters 9 to 18 useful for review.

One possibility that I do not recommend would be to skip the computational material. I am strongly of the opinion that the undergraduate physics curricula at most universities still contain too little instruction in the computational methods that students will need in their careers.

Acknowledgments

This book was originally intended as a resource for my students in Thermal Physics I (33-341) and Thermal Physics II (33-342) at Carnegie Mellon University. In an important sense, those students turned out to be essential collaborators in its production.

I would like to thank the many students from these courses for their great help in suggesting improvements and correcting errors in the text. All of my students have made important contributions. Even so, I would like to mention explicitly the following students: Michael Alexovich, Dimitry Ayzenberg, Conroy Baltzell, Anthony Bartolotta, Alexandra Beck, David Bemiller, Alonzo Benavides, Sarah Benjamin, John Briguglio, Coleman Broaddus, Matt Buchovecky, Luke Ceurvorst, Jennifer Chu, Kunting Chua, Charles Wesley Cowan, Charles de las Casas, Matthew Daily, Brent Driscoll, Luke Durback, Alexander Edelman, Benjamin Ellison, Danielle Fisher, Emily Gehrels, Yelena Goryunova, Benjamin Greer, Nils Guillermin, Asad Hasan, Aaron Henley, Maxwell

Hutchinson, Andrew Johnson, Agnieszka Kalinowski, Patrick Kane, Kamran Karimi, Joshua Keller, Deena Kim, Andrew Kojzar, Rebecca Krall, Vikram Kulkarni, Avishek Kumar, Anastasia Kurnikova, Thomas Lambert, Grant Lee, Robert Lee, Jonathan Long, Sean Lubner, Alan Ludin, Florence Lui, Christopher Magnollay, Alex Marakov, Natalie Mark, James McGee, Andrew McKinnie, Jonathan Michel, Corey Montella, Javier Novales, Kenji Oman, Justin Perry, Stephen Poniatowicz, Thomas Prag, Alisa Rachubo, Mohit Raghunathan, Peter Ralli, Anthony Rice, Svetlana Romanova, Ariel Rosenburg, Matthew Rowe, Kaitlyn Schwalje, Omar Shams, Gabriella Shepard, Karpur Shukla, Stephen Sigda, Michael Simms, Nicholas Steele, Charles Swanson, Shaun Swanson, Brian Tabata, Likun Tan, Joshua Tepper, Kevin Tian, Eric Turner, Joseph Vukovich, Joshua Watzman, Andrew Wesson, Justin Winokur, Nanfei Yan, Andrew Yeager, Brian Zakrzewski, and Yuriy Zubovski. Some of these students made particularly important contributions, for which I have thanked them personally. My students' encouragement and suggestions have been essential in writing this book.

Yutaro Iiyama and Marilia Cabral Do Rego Barros have both assisted with the grading of Thermal Physics courses, and have made very valuable corrections and suggestions.

The last stages in finishing the manuscript were accomplished while I was a guest at the Institute of Statistical and Biological Physics at the Ludwig-Maximilians-Universität, Munich, Germany. I would like to thank Prof. Dr. Erwin Frey and the other members of the Institute for their gracious hospitality.

Throughout this project, the support and encouragement of my friends and colleagues Harvey Gould and Jan Tobochnik have been greatly appreciated.

I would also like to thank my good friend Lawrence Erlbaum, whose advice and support have made an enormous difference in navigating the process of publishing a book.

Finally, I would like to thank my wife, Roberta (Bobby) Klatzky, whose contributions are beyond count. I could not have written this book without her loving encouragement, sage advice, and relentless honesty.

My thesis advisor, Herbert Callen, first taught me that statistical mechanics and thermodynamics are fascinating subjects. I hope you come to enjoy them as much as I do.

Robert H. Swendsen Pittsburgh, January 2011

Contents

| 1 Introduction | | uction | 1 |
|----------------|-----|--------------------------------|---|
| | 1.1 | Thermal Physics | 1 |
| | 1.2 | What are the Questions? | 2 |
| | 1.3 | History | 2 |
| | 1.4 | Basic Concepts and Assumptions | 4 |
| | 1.5 | Plan of the Book | 6 |
| | | | |

Part I Entropy

| 2 | The C | lassical Ideal Gas | 11 |
|---|--------|---|----|
| | 2.1 | Ideal Gas | 11 |
| | 2.2 | Phase Space of a Classical Gas | 12 |
| | 2.3 | Distinguishability | 13 |
| | 2.4 | Probability Theory | 13 |
| | 2.5 | Boltzmann's Definition of the Entropy | 14 |
| | 2.6 | $S = k \log W$ | 14 |
| | 2.7 | Independence of Positions and Momenta | 15 |
| | 2.8 | Road Map for Part I | 15 |
| 3 | Discre | te Probability Theory | 16 |
| | 3.1 | What is Probability? | 16 |
| | 3.2 | Discrete Random Variables and Probabilities | 18 |
| | 3.3 | Probability Theory for Multiple Random Variables | 19 |
| | 3.4 | Random Numbers and Functions of Random Variables | 21 |
| | 3.5 | Mean, Variance, and Standard Deviation | 24 |
| | 3.6 | Correlation Function | 25 |
| | 3.7 | Sets of Independent Random Numbers | 25 |
| | 3.8 | Binomial Distribution | 27 |
| | 3.9 | Gaussian Approximation to the Binomial Distribution | 29 |
| | 3.10 | A Digression on Gaussian Integrals | 30 |
| | 3.11 | Stirling's Approximation for N! | 31 |
| | 3.12 | Binomial Distribution with Stirling's Approximation | 34 |

| | 3.13 | Multinomial Distribution | 35 |
|---|---------|--|----|
| | 3.14 | Problems | 36 |
| 4 | The C | lassical Ideal Gas: Configurational Entropy | 43 |
| | 4.1 | Separation of Entropy into Two Parts | 43 |
| | 4.2 | Probability Distribution of Particles | 44 |
| | 4.3 | Distribution of Particles between Two Isolated Systems that were | |
| | | Previously in Equilibrium | 45 |
| | 4.4 | Consequences of the Binomial Distribution | 46 |
| | 4.5 | Actual Number versus Average Number | 47 |
| | 4.6 | The 'Thermodynamic Limit' | 48 |
| | 4.7 | Probability and Entropy | 48 |
| | 4.8 | The Generalization to $M \ge 2$ Systems | 51 |
| | 4.9 | An Analytic Approximation for the Configurational Entropy | 52 |
| | 4.10 | Problems | 53 |
| 5 | Contir | nuous Random Numbers | 54 |
| | 5.1 | Continuous Dice and Probability Densities | 54 |
| | 5.2 | Probability Densities | 55 |
| | 5.3 | Dirac Delta Functions | 57 |
| | 5.4 | Transformations of Continuous Random Variables | 61 |
| | 5.5 | Bayes' Theorem | 63 |
| | 5.6 | Problems | 65 |
| 6 | The C | lassical Ideal Gas: Energy Dependence of Entropy | 70 |
| | 6.1 | Distribution for the Energy between Two Subsystems | 70 |
| | 6.2 | Evaluation of Ω_p | 72 |
| | 6.3 | Distribution of Energy between Two Isolated Subsystems that were | |
| | | Previously in Equilibrium | 75 |
| | 6.4 | Probability Distribution for Large N | 76 |
| | 6.5 | The Logarithm of the Probability Distribution and the | |
| | | Energy-Dependent Terms in the Entropy | 78 |
| | 6.6 | The Generalization to $M \ge 2$ systems | 79 |
| 7 | Classie | cal Gases: Ideal and Otherwise | 81 |
| | 7.1 | Entropy of a Composite System of Classical Ideal Gases | 81 |
| | 7.2 | Equilibrium Conditions for the Ideal Gas | 82 |
| | 7.3 | The Volume-Dependence of the Entropy | 85 |
| | 7.4 | Asymmetric Pistons | 87 |
| | 7.5 | Indistinguishable Particles | 87 |
| | 7.6 | Entropy of a Composite System of Interacting Particles | 89 |
| | 7.7 | The Second Law of Thermodynamics | 95 |

| 0 | • |
|----------|-----|
| Contents | X1X |

| | 7.8 | Equilibrium between Subsystems | 95 |
|-----|---------|--|-----|
| | 7.9 | The Zeroth Law of Thermodynamics | 97 |
| | 7.10 | Problems | 97 |
| 8 | Tempe | erature, Pressure, Chemical Potential, and All That | 99 |
| | 8.1 | Thermal Equilibrium | 99 |
| | 8.2 | What do we Mean by 'Temperature'? | 100 |
| | 8.3 | Derivation of the Ideal Gas Law | 101 |
| | 8.4 | Temperature Scales | 105 |
| | 8.5 | The Pressure and the Entropy | 106 |
| | 8.6 | The Temperature and the Entropy | 106 |
| | 8.7 | Equilibrium with Asymmetric Pistons, Revisited | 107 |
| | 8.8 | The Entropy and the Chemical Potential | 108 |
| | 8.9 | The Fundamental Relation and Equations of State | 109 |
| | 8.10 | The Differential Form of the Fundamental Relation | 109 |
| | 8.11 | Thermometers and Pressure Gauges | 110 |
| | 8.12 | Reservoirs | 110 |
| | 8.13 | Problems | 111 |
| Par | tll The | ermodynamics | |
| 9 | The Po | ostulates and Laws of Thermodynamics | 115 |
| | 9.1 | Thermal Physics | 115 |
| | 9.2 | Microscopic and Macroscopic States | 117 |
| | 9.3 | Macroscopic Equilibrium States | 117 |
| | 9.4 | State Functions | 118 |
| | 9.5 | Properties and Descriptions | 118 |
| | 9.6 | The Essential Postulates of Thermodynamics | 118 |
| | 9.7 | Optional Postulates of Thermodynamics | 120 |
| | 9.8 | The Laws of Thermodynamics | 123 |
| 10 | Pertur | bations of Thermodynamic State Functions | 124 |
| | 10.1 | Small Changes in State Functions | 124 |
| | 10.2 | Conservation of Energy | 125 |
| | 10.3 | Mathematical Digression on Exact and Inexact Differentials | 125 |
| | 10.4 | Conservation of Energy Revisited | 128 |
| | 10.5 | An Equation to Remember | 129 |
| | 10.6 | Problems | 130 |
| 11 | Therm | nodynamic Processes | 132 |
| | 11.1 | Irreversible, Reversible, and Quasi-Static Processes | 132 |
| | 11.2 | Heat Engines | 133 |

| ; |
|---|
| 5 |

| | 11.3 | Maximum Efficiency | 135 |
|----|--------|---|-----|
| | 11.4 | Refrigerators and Air Conditioners | 136 |
| | 11.5 | Heat Pumps | 137 |
| | 11.6 | The Carnot Cycle | 137 |
| | 11.7 | Alternative Formulations of the Second Law | 139 |
| | 11.8 | Positive and Negative Temperatures | 140 |
| | 11.9 | Problems | 146 |
| 12 | Therm | nodynamic Potentials | 148 |
| | 12.1 | Mathematical Digression: The Legendre Transform | 148 |
| | 12.2 | Helmholtz Free Energy | 152 |
| | 12.3 | Enthalpy | 153 |
| | 12.4 | Gibbs Free Energy | 155 |
| | 12.5 | Other Thermodynamic Potentials | 155 |
| | 12.6 | Massieu Functions | 156 |
| | 12.7 | Summary of Legendre Transforms | 156 |
| | 12.8 | Problems | 157 |
| 13 | The Co | onsequences of Extensivity | 159 |
| | 13.1 | The Euler Equation | 160 |
| | 13.2 | The Gibbs–Duhem Relation | 161 |
| | 13.3 | Reconstructing the Fundamental Relation | 162 |
| | 13.4 | Thermodynamic Potentials | 163 |
| 14 | Therm | nodynamic Identities | 165 |
| | 14.1 | Small Changes and Partial Derivatives | 165 |
| | 14.2 | A Warning about Partial Derivatives | 165 |
| | 14.3 | First and Second Derivatives | 166 |
| | 14.4 | Standard Set of Second Derivatives | 168 |
| | 14.5 | Maxwell Relations | 169 |
| | 14.6 | Manipulating Partial Derivatives | 170 |
| | 14.7 | Working with Jacobians | 174 |
| | 14.8 | Examples of Identity Derivations | 176 |
| | 14.9 | General Strategy | 179 |
| | 14.10 | Problems | 180 |
| 15 | Extrem | num Principles | 184 |
| | 15.1 | Energy Minimum Principle | 185 |
| | 15.2 | Minimum Principle for the Helmholtz Free Energy | 188 |
| | 15.3 | Minimum Principle for the Enthalpy | 190 |
| | 15.4 | Minimum Principle for the Gibbs Free Energy | 191 |
| | 15.5 | Exergy | 192 |

| Contents | |
|----------|-----|
| Contents | XXI |

| | 15.6 | Maximum Principle for Massieu Functions | 193 |
|-----|----------|---|-----|
| | 15.7 | Summary | 194 |
| | 15.8 | Problems | 194 |
| 16 | Stabili | ty Conditions | 195 |
| | 16.1 | Intrinsic Stability | 195 |
| | 16.2 | Stability Criteria based on the Energy Minimum Principle | 196 |
| | 16.3 | Stability Criteria based on the Helmholtz Free Energy Minimum | |
| | | Principle | 199 |
| | 16.4 | Stability Criteria based on the Enthalpy Minimization Principle | 200 |
| | 16.5 | Inequalities for Compressibilities and Specific Heats | 201 |
| | 16.6 | Other Stability Criteria | 201 |
| | 16.7 | Problems | 203 |
| 17 | Phase | Transitions | 205 |
| | 17.1 | The van der Waals Fluid | 206 |
| | 17.2 | Derivation of the van der Waals Equation | 206 |
| | 17.3 | Behavior of the van der Waals Fluid | 207 |
| | 17.4 | Instabilities | 208 |
| | 17.5 | The Liquid–Gas Phase Transition | 210 |
| | 17.6 | Maxwell Construction | 212 |
| | 17.7 | Coexistent Phases | 212 |
| | 17.8 | Phase Diagram | 213 |
| | 17.9 | Helmholtz Free Energy | 214 |
| | 17.10 | Latent Heat | 216 |
| | 17.11 | The Clausius-Clapeyron Equation | 217 |
| | 17.12 | Gibbs' Phase Rule | 218 |
| | 17.13 | Problems | 219 |
| 18 | The No. | ernst Postulate: The Third Law of Thermodynamics | 223 |
| | 18.1 | Classical Ideal Gas Violates the Nernst Postulate | 223 |
| | 18.2 | Planck's Form of the Nernst Postulate | 224 |
| | 18.3 | Consequences of the Nernst Postulate | 224 |
| | 18.4 | Coefficient of Thermal Expansion at Low Temperatures | 225 |
| | 18.5 | The Impossibility of Attaining a Temperature of Absolute Zero | 226 |
| | 18.6 | Summary and Signposts | 226 |
| | 18.7 | Problems | 227 |
| Par | t III Cl | assical Statistical Mechanics | |
| 19 | Ensem | bles in Classical Statistical Mechanics | 231 |

| 19.1 | Microcanonical Ensemble | 232 |
|------|-------------------------|-----|
| | | |

| | 19.2 | Molecular Dynamics: Computer Simulations | 232 |
|----|---------|--|-----|
| | 19.3 | Canonical Ensemble | 234 |
| | 19.4 | The Partition Function as an Integral over Phase Space | 237 |
| | 19.5 | The Liouville Theorem | 238 |
| | 19.6 | Consequences of the Canonical Distribution | 240 |
| | 19.7 | The Helmholtz Free Energy | 241 |
| | 19.8 | Thermodynamic Identities | 242 |
| | 19.9 | Beyond Thermodynamic Identities | 243 |
| | 19.10 | Integration over the Momenta | 244 |
| | 19.11 | Monte Carlo Computer Simulations | 245 |
| | 19.12 | Factorization of the Partition Function: The Best Trick in | |
| | | Statistical Mechanics | 249 |
| | 19.13 | Simple Harmonic Oscillator | 250 |
| | 19.14 | Problems | 252 |
| 20 | Classie | cal Ensembles: Grand and Otherwise | 258 |
| | 20.1 | Grand Canonical Ensemble | 258 |
| | 20.2 | Grand Canonical Probability Distribution | 259 |
| | 20.3 | Importance of the Grand Canonical Partition Function | 261 |
| | 20.4 | $\mathcal{Z}(T, V, \mu)$ for the Ideal Gas | 262 |
| | 20.5 | Summary of the Most Important Ensembles | 263 |
| | 20.6 | Other Classical Ensembles | 264 |
| | 20.7 | Problems | 264 |
| 21 | Refini | ng the Definition of Entropy | 266 |
| | 21.1 | The Canonical Entropy | 266 |
| | 21.2 | The Grand Canonical Entropy | 268 |
| | 21.3 | Problems | 270 |
| 22 | Irreve | rsibility | 271 |
| | 22.1 | What Needs to be Explained? | 271 |
| | 22.2 | Trivial Form of Irreversibility | 272 |
| | 22.3 | Boltzmann's H-Theorem | 272 |
| | 22.4 | Loschmidt's Umkehreinwand | 272 |
| | 22.5 | Zermelo's Wiederkehreinwand | 273 |
| | 22.6 | Free Expansion of a Classical Ideal Gas | 273 |
| | 22.7 | Zermelo's Wiederkehreinwand Revisited | 278 |
| | 22.8 | Loschmidt's Umkehreinwand Revisited | 278 |
| | 22.9 | What is 'Equilibrium'? | 279 |
| | 22.10 | Entropy | 279 |
| | 22.11 | Interacting Particles | 281 |

Part IV Quantum Statistical Mechanics

| 23 | Quant | um Ensembles | 285 |
|----|--------|---|-----|
| | 23.1 | Basic Quantum Mechanics | 286 |
| | 23.2 | Energy Eigenstates | 287 |
| | 23.3 | Many-Body Systems | 290 |
| | 23.4 | Two Types of Probability | 290 |
| | 23.5 | The Density Matrix | 293 |
| | 23.6 | Uniqueness of the Ensemble | 294 |
| | 23.7 | The Planck Entropy | 295 |
| | 23.8 | The Quantum Microcanonical Ensemble | 296 |
| 24 | Quant | um Canonical Ensemble | 297 |
| | 24.1 | Derivation of the QM Canonical Ensemble | 297 |
| | 24.2 | Thermal Averages and the Average Energy | 299 |
| | 24.3 | The Quantum Mechanical Partition Function | 299 |
| | 24.4 | The Quantum Mechanical Entropy | 302 |
| | 24.5 | The Origin of the Third Law of Thermodynamics | 303 |
| | 24.6 | Derivatives of Thermal Averages | 305 |
| | 24.7 | Factorization of the Partition Function | 306 |
| | 24.8 | Special Systems | 308 |
| | 24.9 | Two-Level Systems | 309 |
| | 24.10 | Simple Harmonic Oscillator | 311 |
| | 24.11 | Einstein Model of a Crystal | 313 |
| | 24.12 | Problems | 315 |
| 25 | Black- | Body Radiation | 322 |
| | 25.1 | Black Bodies | 322 |
| | 25.2 | Universal Frequency Spectrum | 322 |
| | 25.3 | A Simple Model | 323 |
| | 25.4 | Two Types of Quantization | 323 |
| | 25.5 | Black-Body Energy Spectrum | 325 |
| | 25.6 | Total Energy | 328 |
| | 25.7 | Total Black-Body Radiation | 329 |
| | 25.8 | Significance of Black-Body Radiation | 329 |
| | 25.9 | Problems | 330 |
| 26 | The H | armonic Solid | 331 |
| | 26.1 | Model of a Harmonic Solid | 331 |
| | 26.2 | Normal Modes | 332 |
| | 26.3 | Transformation of the Energy | 336 |
| | 26.4 | The Frequency Spectrum | 338 |

| | 26.5 | Alternate Derivation: Equations of Motion | 340 |
|----|---------|---|-----|
| | 26.6 | The Energy in the Classical Model | 341 |
| | 26.7 | The Quantum Harmonic Crystal | 342 |
| | 26.8 | Debye Approximation | 343 |
| | 26.9 | Problems | 348 |
| 27 | Ideal (| Quantum Gases | 350 |
| | 27.1 | Single-Particle Quantum States | 350 |
| | 27.2 | Density of Single-Particle States | 352 |
| | 27.3 | Many-Particle Quantum States | 353 |
| | 27.4 | Quantum Canonical Ensemble | 355 |
| | 27.5 | Grand Canonical Ensemble | 355 |
| | 27.6 | A New Notation for Energy Levels | 356 |
| | 27.7 | Exchanging Sums and Products | 357 |
| | 27.8 | Grand Canonical Partition Function for Independent Particles | 358 |
| | 27.9 | Distinguishable Quantum Particles | 359 |
| | 27.10 | Sneaky Derivation of $PV = Nk_BT$ | 360 |
| | 27.11 | Equations for $U = \langle E \rangle$ and $\langle N \rangle$ | 360 |
| | 27.12 | $\langle n_{\alpha} \rangle$ for Bosons | 362 |
| | 27.13 | $\langle n_{\alpha} \rangle$ for Fermions | 362 |
| | 27.14 | Summary of Equations for Fermions and Bosons | 363 |
| | 27.15 | Integral Form of Equations for N and U | 364 |
| | 27.16 | Basic Strategy for Fermions and Bosons | 365 |
| | 27.17 | P = 2U/3V | 365 |
| | 27.18 | Problems | 367 |
| 28 | Bose- | Einstein Statistics | 369 |
| | 28.1 | Basic Equations for Bosons | 369 |
| | 28.2 | $\langle n_{\alpha} \rangle$ for Bosons | 369 |
| | 28.3 | The Ideal Bose Gas | 370 |
| | 28.4 | Low-Temperature Behavior of μ | 371 |
| | 28.5 | Bose-Einstein Condensation | 373 |
| | 28.6 | Below the Einstein Temperature | 373 |
| | 28.7 | Energy of an Ideal Gas of Bosons | 375 |
| | 28.8 | What About the Second-Lowest Energy State? | 376 |
| | 28.9 | The Pressure below $T < T_E$ | 377 |
| | 28.10 | Transition Line in P-V Plot | 378 |
| | 28.11 | A Numerical Approach to Bose-Einstein Statistics | 378 |
| | 28.12 | Problems | 380 |

| Contents | XXV |
|----------|-----|
| | |

| 29 | Fermi–Dirac Statistics | | 386 |
|----|-------------------------------|---|-----|
| | 29.1 | Basic Equations for Fermions | 386 |
| | 29.2 | The Fermi Function and the Fermi Energy | 387 |
| | 29.3 | A Useful Identity | 388 |
| | 29.4 | Systems with a Discrete Energy Spectrum | 389 |
| | 29.5 | Systems with Continuous Energy Spectra | 390 |
| | 29.6 | Ideal Fermi Gas | 391 |
| | 29.7 | Fermi Energy | 391 |
| | 29.8 | Compressibility of Metals | 392 |
| | 29.9 | Sommerfeld Expansion | 393 |
| | 29.10 | General Fermi Gas at Low Temperatures | 396 |
| | 29.11 | Ideal Fermi Gas at Low Temperatures | 398 |
| | 29.12 | Problems | 399 |
| 30 | Insulators and Semiconductors | | 404 |
| | 30.1 | Tight-Binding Approximation | 404 |
| | 30.2 | Bloch's Theorem | 406 |
| | 30.3 | Nearly-Free Electrons | 408 |
| | 30.4 | Energy Bands and Energy Gaps | 410 |
| | 30.5 | Where is the Fermi Energy? | 411 |
| | 30.6 | Fermi Energy in a Band (Metals) | 412 |
| | 30.7 | Fermi Energy in a Gap | 412 |
| | 30.8 | Intrinsic Semiconductors | 416 |
| | 30.9 | Extrinsic Semiconductors | 416 |
| | 30.10 | Semiconductor Statistics | 418 |
| | 30.11 | Semiconductor Physics | 422 |
| 31 | Phase | Transitions and the Ising Model | 423 |
| | 31.1 | The Ising Chain | 424 |
| | 31.2 | The Ising Chain in a Magnetic Field ($\mathcal{J} = 0$) | 425 |
| | 31.3 | The Ising Chain with $h = 0$, but $\mathcal{J} \neq 0$ | 426 |
| | 31.4 | The Ising Chain with both $\mathcal{J} \neq 0$ and $h \neq 0$ | 428 |
| | 31.5 | Mean-Field Approximation | 432 |
| | 31.6 | Critical Exponents | 436 |
| | 31.7 | Mean-Field Exponents | 437 |
| | 31.8 | Analogy with the van der Waals Approximation | 438 |
| | 31.9 | Landau Theory | 439 |
| | 31.10 | Beyond Landau Theory | 440 |
| | 31.11 | Problems | 441 |

| Appendix | 447 | |
|----------|--------------------------|-----|
| A.1 | MatPlotLib | 447 |
| A.2 | Python | 449 |
| A.3 | Histograms | 449 |
| A.4 | The First Python Program | 450 |
| A.5 | Python Functions | 451 |
| A.6 | Graphs | 452 |
| A.7 | Reporting Python Results | 453 |
| A.8 | Timing Your Program | 455 |
| A.9 | Molecular Dynamics | 455 |
| A.10 | Courage | 456 |
| Index | | 457 |

Introduction

1

If, in some cataclysm, all scientific knowledge were to be destroyed, and only one sentence passed on to the next generation of creatures, what statement would contain the most information in the fewest words? I believe it is the atomic hypothesis (or atomic fact, or whatever you wish to call it) that all things are made of atoms—little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another. In that one sentence you will see an enormous amount of information about the world, if just a little imagination and thinking are applied.

Richard Feynman, in The Feynman Lectures on Physics

1.1 Thermal Physics

This book is about the things you encounter in everyday life: the book you are reading, the chair on which you are sitting, the air you are breathing. It is about things that can be hot or cold; hard or soft; solid, liquid, or gas. It is about machines that work for you: automobiles, heaters, refrigerators, air conditioners. It is even about your own body and the stuff of life. The whole subject is sometimes referred to as thermal physics, and it is usually divided into two main topics: thermodynamics and statistical mechanics.

Thermodynamics is the study of everything connected with heat. It provides powerful methods for connecting observable quantities by equations that are not at all obvious, but are nevertheless true for *all* materials. Statistical mechanics is the study of what happens when large numbers of particles interact. It provides a foundation for thermodynamics and the ultimate justification of why thermodynamics works. It goes beyond thermodynamics to reveal deeper connections between molecular behavior and material properties. It also provides a way to calculate the properties of specific objects, instead of just the universal relationships provided by thermodynamics.

The ideas and methods of thermal physics differ from those of other branches of physics. Thermodynamics and statistical mechanics each require their own particular ways of thinking. Studying thermal physics is not about memorizing formulas; it is about gaining a new understanding of the world.

2 Introduction

1.2 What are the Questions?

Thermal physics seeks quantitative explanations for the properties of macroscopic objects, where the term 'macroscopic' means two things:

- 1. A macroscopic object is made up of a large number of particles.
- 2. Macroscopic measurements have limited resolution.

The apparently vague terms 'large number' and 'limited resolution' take on specific meanings with respect to the law of large numbers that we will study in Chapter 3. We will see that the relative statistical uncertainty in quantities like the energy and the number of particles is usually inversely proportional to the square root of the number of particles. If this uncertainty is much smaller than the experimental resolution, it can be neglected. This leads to thermodynamics, which is a description of the macroscopic properties that ignores small statistical fluctuations.

We will assume that we somehow know what the object is made of; that is, what kinds of atoms and molecules it contains, and how many of each. We will generally speak of 'systems' instead of objects. The difference is that a system can consist of any number of objects, and is somehow separated from the rest of the universe. We will concentrate on systems in equilibrium; that is, systems that have been undisturbed long enough for their properties to take on constant values. We shall be more specific about what 'long enough' means in Chapter 22.

In the simplest case—which we will consider in Part I—a system might be completely isolated by 'adiabatic' walls; that is, rigid barriers that let nothing through. We will also assume, at least at first, that we know how much energy is contained in our isolated system.

Given this information, we will ask questions about the properties of the system. We will first study a simple model of a gas, so we will want to know what temperature and pressure are, and how they are related to each other and to the volume. We will want to know how the volume or the pressure will change if we heat the gas. As we investigate more complicated systems, we will see more complex behavior and find more questions to ask about the properties of matter.

As a hint of things to come, we will find that there is a function of the energy E, volume V, and number of particles N that is sufficient to answer *all* questions about the thermal properties of a system. It is called the entropy, and it is denoted by the letter S.

If we can calculate the entropy as a function of E, V, and N, we know everything about the macroscopic properties of the system. For this reason, S = S(E, V, N) is known as a fundamental relation, and it is the focus of Part I.

1.3 History

Atoms exist, combine into molecules, and form every object we encounter in life. Today, this statement is taken for granted. However, in the nineteenth century, and even into

the early twentieth century, these were fighting words. The Austrian physicist Ludwig Boltzmann (1844–1906) and a small group of other scientists championed the idea of atoms and molecules as the basis of a fundamental theory of the thermal properties of matter, against the violent objections of another Austrian physicist, Ernst Mach (1838– 1916), and many of his colleagues. Boltzmann was right, of course, but the issue had not been settled even at the time of his tragic suicide in 1906. In the tradition of Boltzmann, the intention in this book is to present thermodynamics as a consequence of the molecular nature of matter.

The theory of thermodynamics was developed without benefit of the atomic hypothesis. It began with the seminal work of Sadi Carnot (French physicist, 1792–1832), who initiated the scientific discussion of heat engines. The First Law of Thermodynamics was discovered by James Prescott Joule (English physicist, 1818–1889), when he established that heat was a form of energy and measured the mechanical equivalent of heat. The Second Law of Thermodynamics was first stated in 1850 by Rudolf Clausius (German physicist, 1822–1888), who in 1865 also invented the related concept of entropy. The Second Law can be expressed by the statement that the entropy of an isolated system can increase, but not decrease.

Entropy was a mystery to nineteenth-century scientists. Clausius had given entropy an experimental definition that allowed it to be calculated, but its meaning was puzzling. Like energy, it could not be destroyed, but unlike energy, it could be created. It was essential to the calculation of the efficiency of heat engines (machines that turn the energy in hot objects into mechanical work), but it did not seem to be related to any other physical laws.

The reason why scientists working in the middle of the nineteenth century found entropy so mysterious is that few of them thought in terms of atoms or molecules. Even with molecular theory, explaining the entropy required brilliant insight; without molecular theory, there was no hope.

Serious progress in understanding the properties of matter and the origins of thermodynamic laws on the basis of the molecular nature of matter began in the 1860s with the work of Boltzmann and the American physicist J. Willard Gibbs (1839–1903).

Gibbs worked from a formal starting point, postulating that observable quantities could be calculated from an average over many replicas of an object in different microscopic configurations, and then working out what the equations would have to be. His work is very beautiful (to a theoretical physicist), although somewhat formal. However, it left certain questions unresolved—most notably, what has come to be called 'Gibbs' paradox'. This issue concerned a discontinuous change in the entropy when differences in the properties of particles in a mixture were imagined to disappear continuously. Gibbs himself did not regard this as a paradox. The issues involved are still a matter of debate in the twenty-first century.

Boltzmann devoted most of his career to establishing the molecular theory of matter and deriving the consequences of the existence of atoms and molecules. One of his great achievements was his 1877 definition of entropy—a definition which provides a physical interpretation of the entropy and the foundation of statistical mechanics. Unfortunately, although Boltzmann's paper is often cited, it is rarely read. This had led to frequent misinterpretations of what he meant. At least part of the problem is that he wrote in a fairly dense nineteenth-century German, and an English translation has only appeared recently.¹ We will discuss what he actually wrote in some detail.

Part I of this book is devoted to developing an intuitive understanding of the concept of entropy, based on Boltzmann's definition. We will present an explicit, detailed derivation of the entropy for a simple model to provide insight into its significance. Later parts of the book will develop more sophisticated tools for investigating thermodynamics and statistical mechanics, but they are all based on Boltzmann's definition of the entropy.

1.4 Basic Concepts and Assumptions

This book is concerned with the macroscopic behavior of systems containing many particles. This is a broad topic since everything we encounter in the world around us contains enormous numbers of atoms and molecules. Even in relatively small objects, there are 10^{20} or more atoms.

Atoms and molecules are not the only examples of large numbers of particles in a system. Colloids, for example, can consist of 10^{12} or more microscopic particles suspended in a liquid. The large number of particles in a typical colloid means that they are also well described by statistical mechanics.

Another aspect of macroscopic experiments is that they have limited resolution. We will see in Part I that the fluctuations in quantities like the density of a system are approximately inversely proportional to the square root of the number of particles in the system. If there are 10^{20} molecules, this gives a relative uncertainty of about 10^{-10} for the average density. Because it is rare to measure the density to better than one part in 10^{6} , these tiny fluctuations are not seen in macroscopic measurements. Indeed, an important reason for Mach's objection to Boltzmann's molecular hypothesis was that no direct measurement of atoms had ever been carried out in the nineteenth century.

Besides the limited accuracy, it is rare for more than a million quantities to be measured in an experiment, and usually only a handful of measurements are made. Since it would take about 6N quantities to specify the microscopic state of N atoms, thermo-dynamic measurements provide relatively little information about the microscopic state.

Due to our lack of detailed knowledge of the microscopic state of an object, we need to use probability theory—discussed in Chapters 3 and 5—to make further progress. However, we do not know the probabilities of the various microscopic states either. This means that we have to make assumptions about the probabilities. We will make

¹ Sharp, K. and Matschinsky, F. Translation of Ludwig Boltzmann's 'Paper On the Relationship between the Second Fundamental Theorem of the Mechanical Theory of Heat and Probability Calculations Regarding the Conditions for Thermal Equilibrium', (Sitzungberichte der Kaiserlichen Akademie der Wissenschaften. Mathematisch-Naturwissen Classe). Abt. II, LXXVI 1877, pp. 373–435 (Wien. Ber. 1877, 76: 373–435). Reprinted in Wiss. Abhandlungen, Vol. II, reprint 42, pp. 164–223, Barth, Leipzig, 1909. *Entropy* 2015, 17: 1971–2009.

the simplest assumptions that are consistent with the physical constraints (number of particles, total energy, and so on): we will assume that everything we do not know is equally probable.

Based on our assumptions about the probability distribution, we will calculate the macroscopic properties of the system and compare our predictions with experimental data. We will find that our predictions are correct. This is comforting. However, we must realize that it does not necessarily mean that our assumptions were correct. In fact, we will see in Chapter 22 that many different assumptions would also lead to the same predictions. This is not a flaw in the theory but merely a fact of life. Recognizing this fact helps a great deal in resolving apparent paradoxes, as we will see.

The predictions we make based on assumptions about the probabilities of microscopic states lead to the postulates of thermodynamics. These postulates, in turn, are sufficient to derive the very powerful formalism of thermodynamics, as we will examine in Part II.

These same assumptions about the probabilities of microscopic states also lead to the even more powerful formalism of statistical mechanics, which we will investigate in the Parts III and IV of the book.

1.4.1 State Functions

It has long been known that when most macroscopic systems are left by themselves for a long period of time, their measurable properties stop changing and become timeindependent. Simple systems, like a container of gas, evolve into macroscopic states that are well described by a small number of variables. For a simple gas, the energy, volume, and number of particles might be sufficient. These quantities, along with other quantities that we will discuss, are known as 'state functions'.

The molecules in a macroscopic system are in constant motion so that the microscopic state is constantly changing. This fact leads to a fundamental question. How is it that the macroscopic state can be time-independent with precisely defined properties? The answer to this question should become clear in Part I of this book.

1.4.2 Irreversibility

A second fundamental question is this. Even if there exist macroscopic equilibrium states that are independent of time, how can a system evolve toward such a state but not away from it? How can a system that obeys time-reversal-invariant laws of motion show irreversible behavior? This question has been the subject of much debate, at least since Johann Josef Loschmidt's (Austrian physicist, 1821–1895) formulation of the 'reversibility paradox' in 1876. We will present a resolution of the paradox in Chapter 22.

1.4.3 Entropy

The Second Law of Thermodynamics states that there exists a state function called the 'entropy' that is maximized in equilibrium. Boltzmann's 1877 definition of the entropy

6 Introduction

provides an account of what this means. A major purpose of the calculation of the entropy of a classical ideal gas in Part I is to obtain an intuitive understanding of the significance of Boltzmann's entropy.

1.5 Plan of the Book

This book presents thermal physics as a consequence of the molecular nature of matter. The book is divided into four parts, in order to provide a systematic development of the ideas for someone coming to the subject for the first time, as well as for someone who knows the basic material but would like to review a particular topic.

Part I: Entropy The entropy will be introduced by an explicit calculation for the classical ideal gas. The ideal gas is a simple model with macroscopic properties that can be calculated exactly. This allows us to derive the entropy in closed form without any hidden assumptions. Since the entropy of the classical ideal gas exhibits most of the thermodynamic properties of more general systems, it will serve as an introduction to the otherwise rather abstract postulates of thermodynamics. In the last two chapters, Chapters 7 and 8 of Part I, the formal expression for the entropy of a classical gas with interacting particles is obtained, along with general expressions for the temperature, pressure, and chemical potential, which establish the foundations of classical statistical mechanics.

Part II: Thermodynamics The formal postulates of thermodynamics are introduced, based on the properties of the entropy derived in Part I. Our treatment follows the vision of László Tisza (Hungarian physicist, 1907–2009) and Herbert B. Callen (American physicist, 1919–1993). Although the original development of thermodynamics by nineteenth-century physicists was a brilliant achievement, their arguments were somewhat convoluted because they did not understand the microscopic molecular basis of the laws they had discovered. Deriving the equations of thermodynamics from postulates is much easier than following the historical path. The full power of thermodynamics can be developed in a straightforward manner and the structure of the theory made transparent.

Part III: Classical statistical mechanics Here we return to classical statistical mechanics to discuss more powerful methods of calculation. In particular, we introduce the canonical ensemble, which describes the behavior of a system in contact with a heat bath at constant temperature. The canonical ensemble provides a very powerful approach to solving most problems in classical statistical mechanics. We also introduce the grand canonical ensemble, which describes a system that can exchange particles with a large system at a fixed chemical potential. This ensemble will prove to be particularly important when we encounter it again in Chapters 27 through 30 of Part IV, where we discuss quantum systems of indistinguishable particles. Statistical mechanics can derive results that go beyond those of thermodynamics. We discuss and resolve the apparent conflict between time-reversal-invariant microscopic equations with the obvious existence of irreversible behavior in the macroscopic world. Part III will introduce molecular dynamics and Monte Carlo computer simulations to demonstrate some of the modern methods for obtaining information about many-particle systems. We will also refine the calculation of the entropy to include the non-zero width of the energy and particle number distributions.

Part IV: Quantum statistical mechanics In the last part of the book, we develop quantum statistical mechanics, which is necessary for the understanding of the properties of most real systems. After two introductory chapters on basic ideas we will devote chapters to black-body radiation and lattice vibrations. There is a chapter on the general theory of indistinguishable particles, followed by individual chapters on the properties of bosons and fermions. Since the application of the theory of fermions to the properties of insulators and semiconductors has both theoretical and practical importance, this topic has a chapter of its own. The last chapter provides an introduction to the Ising model of ferromagnetism as an example of the theory of phase transitions.

Part I Entropy

2 The Classical Ideal Gas

The laws of thermodynamics, as empirically determined, express the approximate and probable behavior of systems of a great number of particles, or, more precisely, they express the laws of mechanics of such systems as they appear to beings who have not the fineness of perception to enable them to appreciate quantities of the order of magnitude of those which relate to single particles, and who cannot repeat their experiments often enough to obtain any but the most probable results.

J. Willard Gibbs

The purpose of Part I of this book is to provide an intuitive understanding of the entropy, based on calculations for a simple model. The model chosen is the classical ideal gas, for which the entropy can be calculated explicitly and completely without any approximations or hidden assumptions.

The treatment is entirely in terms of the theory of classical mechanics. No quantum mechanical concepts are used. All the ideas will follow directly from the work of Ludwig Boltzmann (1844–1906), and especially his 1877 paper on the Second Law of Thermodynamics. We will use more modern mathematical methods than he did to derive the entropy of the classical ideal gas, but we will make no assumptions with which he was not familiar.

In Chapters 7 and 8 the formal expression for the entropy will be extended to classical systems with interacting particles. Although the expression we obtain can rarely be evaluated exactly, the formal structure will be sufficient to provide a basis for the development of thermodynamics in Part II. The same formal structure will also lead to more powerful methods of calculation for statistical mechanics in Parts III and IV.

2.1 Ideal Gas

What distinguishes an 'ideal' gas from a 'real' gas is the absence of interactions between the particles. Although an ideal gas might seem to be an unrealistic model, its properties are experimentally accessible by studying real gases at low densities. Since even the molecules in the air you are breathing are separated by an average distance of about ten times their diameter, nearly ideal gases are easy to find.

The most important feature that is missing from a classical ideal gas is that it does not exhibit any phase transitions. Other than that, its properties are qualitatively the same as those of real gases, which makes it valuable for developing intuition about statistical mechanics and thermodynamics.

The great advantage of the ideal gas model is that all of its properties can be calculated exactly, and nothing is obscured by mathematical complexity.

2.2 Phase Space of a Classical Gas

Our model of a classical gas consists of N particles contained in some specified volume. Each particle has a well-defined position and momentum. The positions of all particles can be represented as a point in configuration space—an abstract 3N-dimensional space, with axes for every coordinate of every particle. These coordinates can be given in various forms:

$$q = \{\vec{r}_i | i = 1, ..., N\}$$

= $\{x_i, y_i, z_i | i = 1, ..., N\}$
= $\{q_j | j = 1, ..., 3N\}.$ (2.1)

The momenta of all particles can be represented as a point in momentum space an abstract 3N-dimensional space, with axes for every component of the momentum of every particle

$$p = \{ \vec{p}_i | i = 1, ..., N \}$$

= { $p_{x,i}, p_{y,i}, p_{z,i} | i = 1, ..., N \}$
= { $p_j | j = 1, ..., 3N \}.$ (2.2)

The kinetic energy of the *i*-th particle is given by the usual expression, $|\vec{p}_i|^2/2m$, and the total kinetic energy is just the sum of the kinetic energies of all particles,

$$E = \sum_{i=1}^{N} \frac{|\vec{p}_i|^2}{2m}.$$
(2.3)

The complete microscopic state of the system can be described by a point in phase space—an abstract 6*N*-dimensional space with axes for every coordinate and every momentum component for all *N* particles. Phase space is the union of configuration space and momentum space, $\{p, q\}$:

$$\{p,q\} = \{q_j, p_j | j = 1, \dots, 3N\}.$$
(2.4)

Since, by definition, there are no interactions between the particles in an ideal gas, the potential energy of the system is zero.

2.3 Distinguishability

Particles will be regarded as distinguishable, in keeping with classical concepts. To be specific, particles are distinguishable when the exchange of two particles results in a different microscopic state. In classical mechanics, this is equivalent to saying that every point in phase space represents a different microscopic state. Distinguishability does not necessarily mean that the particles have different properties; classically, particles were always regarded as distinguishable because their trajectories could, at least in a thought experiment, be followed and the identity of individual particles determined.

On the other hand, it will be important to remember that experiments on macroscopic systems are always assumed to have limited resolution. In both statistical mechanics and thermodynamics, we are concerned with measurements that do not have the sensitivity to resolve the positions or identities of individual atoms.

2.4 Probability Theory

Because of the importance of probability theory in statistical mechanics, two chapters are devoted to the topic. The chapters discuss the basic principles of the probability theory of discrete random variables (Chapter 3) and continuous random variables (Chapter 5).

The mathematical treatment of probability theory has been separated from the physical application for several reasons: (1) it provides an easy reference for the mathematics, (2) it makes the derivation of the entropy more compact, and (3) it is unobtrusive for those readers who are already entirely familiar with probability theory.

If the reader is already familiar with probability theory, Chapters 3 and 5 might be skipped. However, since the methods for the transformation of random variables presented differ from those usually found in mathematics textbooks, these chapters might still be of some interest. It should be noted that we will need the transformation methods using Dirac delta functions, which are rarely found in mathematical texts on probability theory.

The chapters on probability theory are placed just before the chapters in which the material is first needed to calculate contributions to the entropy. Chapter 3 provides the methods needed to calculate the contributions of the positions in Chapter 4, and Chapter 5 provides the methods required to calculate the contributions of the momenta in Chapter 6.

To apply probability theory to the calculation of the properties of the classical ideal gas—or any other model, for that matter—we will have to make assumptions about the probability distribution of the positions and momenta of 10^{20} or more particles. Our basic strategy will be to make the simplest assumptions consistent with what we know about the system and then calculate the consequences.

Another way of describing our strategy is that we are making a virtue of our ignorance of the microscopic states and assume that everything we don't know is equally likely. How this plays out in practice is the subject of the rest of the book.

2.5 Boltzmann's Definition of the Entropy

In 1877, after a few less successful attempts, Boltzmann defined the entropy in terms of the probability of a macroscopic state. His explanation of the Second Law of Thermodynamics was that isolated systems naturally develop from less probable macroscopic states to more probable macroscopic states. Although Boltzmann's earlier efforts to prove this with his famous H-theorem were problematic and highly controversial, his fundamental insight is essentially correct.

In his 1877 paper, Boltzmann also specified that the entropy should be defined in terms of a composite system; that is, a system composed of two or more subsystems with some sort of constraint between them. An example of such a composite system would be a volume of gas divided into two smaller volumes (or subsystems) by a partition. The partition acts as a constraint by restricting the number of particles in each subsystem to be constant. The removal of the partition would then allow the system to develop from a less probable macroscopic state to a more probable macroscopic state for the distribution of particle positions. The final state, after the composite system had come to equilibrium, would correspond to the most probable macroscopic state. According to the Second Law of Thermodynamics, the thermodynamic entropy should also be maximized in the equilibrium state. The comparison of these two properties of the equilibrium state led Boltzmann to associate the entropy with the probability of a macroscopic state, or more precisely with the logarithm of the probability.

In the following chapters, we will make a direct application of Boltzmann's definition to the calculation of the entropy of the classical ideal gas, to within additive and multiplicative constants that we will determine later.

2.6 $S = k \log W$

Boltzmann's achievement has been honored with the inscription of the equation $S = k \log W$ on his tombstone. The symbol S denotes the entropy. The symbol W denotes the German word Wahrscheinlichkeit, which means 'probability'. The natural logarithm is intended, although 'log' is written instead of the more modern 'ln'. Curiously enough, Boltzmann never wrote this equation. It does reflect his ideas, except for the important use of composite systems in his definition. The equation was first written in this form by the German physicist Max Planck (1858–1947) in 1900. The constant k, also written as k_B , is known as the Boltzmann constant.

The symbol W has often been misinterpreted to mean a volume in phase space, which has caused a considerable amount of trouble. This misinterpretation is so common that many scientists are under the impression that Boltzmann defined the entropy as the logarithm of a volume in phase space. Going back to the original meaning of W and Boltzmann's 1877 definition (with composite systems) eliminates much of the confusion about the statistical interpretation of entropy.

The main differences between Boltzmann's treatment of entropy and the one in this book lie in the use of modern mathematical methods and the explicit treatment of the dependence of the entropy on the number of particles.

2.7 Independence of Positions and Momenta

In the derivation of the properties of the classical ideal gas, we will assume that the positions and momenta of the particles are independent. We will present a more formal definition of independence in the Chapter 3, but the idea is that knowing the position of a particle tells us nothing about its momentum, and knowing its momentum tells us nothing about its position. As demonstrated at the beginning of Chapter 4, the independence of the positions and momenta means that their contributions to the total entropy can be calculated separately and simply added to produce the final answer.

2.8 Road Map for Part I

The analysis of the entropy in Part I has been divided into chapters to make it easier to keep track of the different aspects of the derivation.

The concepts and equations of discrete probability theory are developed in Chapter 3, just before they are needed in Chapter 4 to calculate the contributions of the positions to the entropy.

Probability theory for continuous random variables is discussed in Chapter 5, just before its application to the calculation of the contributions of the momenta to the entropy in Chapter 6.

Chapter 7 generalizes the entropy to systems with interacting particles.

Chapter 8 completes the foundations of classical statistical mechanics by relating the partial derivatives of the entropy to the temperature, pressure, and chemical potential.

The following flowchart is intended to illustrate the organization of Part I. The derivation of the entropy of the classical ideal gas follows the arrows down the right-hand side of the flowchart.

Flow chart for Part I



Discrete Probability Theory

The initial state will usually be a very improbable one. The system will run from this state to ever more probable states until it reaches the most probable one, which is the state of thermal equilibrium. If we apply this to the second law, we can identify the quantity that is usually called the entropy with the probability of that state.

Ludwig Boltzmann, 1877

3.1 What is Probability?

The definition of probability is sufficiently problematic to have occasioned something akin to religious wars in academic departments of statistics. There are two basic factions: frequentists and Bayesians.

To understand why the definition of probability can be a problem, consider an experiment with N trials, each of which can result in a success or failure by some criterion. If N_s be the number of successes, then N_s/N is called the frequency of success for those trials.

Frequentists define the probability by the asymptotic frequency of success in the limit of an infinite number of trials,

$$p = \lim_{N \to \infty} \frac{N_s}{N}.$$
(3.1)

This definition looks precise and objective. Indeed, it is found as *the* definition in many physics texts. The major problem is that humans have a finite amount of time available, so that by the frequentist definition we can never determine the probability of anything. Bayesian probability provides a solution to this quandary.

The Bayesian view of probability is based on the work of Thomas Bayes (English mathematician and Presbyterian minister, 1702–1761). Bayesians define probability as a description of a person's knowledge of the outcome of an experiment, based on whatever evidence is at that person's disposal.

One great advantage of the Bayesian definition of probability is that it gives a clear meaning to a statement such as: 'The mass of a proton is $1.672621637(83) \times 10^{-27} kg'$, where the '(83)' is called the uncertainty in the last digit. What does '(83)' mean? Certainly the mass of a proton is an unambiguous number that does not take on

different values for different experiments. Nevertheless, the '(83)' does have meaning as an expression of our uncertainty as to the exact value of a proton's mass.

Bayesian probability is accepted by most statisticians. However, it is in disrepute among some physicists because they regard it as subjective, in the sense that it describes what an individual knows, rather than being absolute truth. However, none of us has access to absolute truth. Bayesian statistics provide an appropriate way to describe the knowledge that an individual does have, based on the experimental results that he or she has access to.

In my opinion, the only reasonable form of objectivity that we can demand is that two observers with the same information will come to the same conclusions, and Bayesian statistics fulfills this requirement.¹

The word 'probability' is used somewhat differently in statistical mechanics. I shall call it 'model probability'. It is most closely related to the frequentist meaning in that the specific values of model probabilities are assumed to be known exactly, as if we had an infinite number of trials. Of course, we really have only finite number of (imprecise) measurements, and as we shall see, we need to fix the probabilities of an enormous number of random variables (in excess of 10^{20}), so that we have no direct access to an experimental confirmation of our assumptions.

The model probability is only assumed to be valid for an equilibrium state. Indeed, all calculations will be for equilibrium, except in Chapter 22, where we discuss irreversibility. Even though we speak of thermodynamics, we limit our discussions to transitions between equilibrium states, primarily because the transitions themselves are too difficult to calculate.

The assumed form of the model probability is usually based on the plausible argument that events that are essentially alike have equal probabilities in equilibrium.

Another common argument uses the property of ergodicity, which is an additional assumption. A system is ergodic if, starting from an arbitrary point in phase space, the trajectory of the system will pass arbitrarily close to every allowable point in phase space. This trajectory is necessarily a one-dimensional subset of a 6*N*-dimensional space. Regions in phase space that are visited equally frequently are then assigned equal probabilities. The argument to ergodicity is very popular, but it has certain weaknesses. For one thing, ergodicity by itself does not determine which regions will be visited equally frequently. Next, ergodicity has only been proven for a very limited number of systems (not the ideal gas). Such an argument also has the difficulty that it requires an enormous amount of time for a trajectory to spread through phase space—of the same order of magnitude as the Poincaré recurrence time discussed in Chapter 22.

A somewhat better argument, although still not proven for general systems, starts with an initial probability distribution that is spread over a non-zero region in phase space, which is a much more realistic assumption. If the time development of such a distribution

¹ This interpretation is sometimes called 'objective Bayesian probability' to distinguish it from 'subjective Bayesian probability'. Some workers do not regard the 'prior' in the Bayesian formalism as describing the information available to an experimenter before the current experiment, but rather as being an arbitrary choice. With this interpretation, Bayesian probability is subjective—and useless for physics.

leads to a unique prediction for observable quantities, that is used to justify the model probability. It has the advantage of converging *much* faster, as shown in Chapter 22 for the classical ideal gas.²

We can then use our assumed model probability to calculate the predicted outcome of an experiment. If measurements agree with our predictions, we can say that our model is consistent with the experiment. This is not the same thing as saying that our model probabilities are correct, and we will see in Chapter 22 that many different distributions can lead to the same predictions. Nevertheless, agreement with experiment is always comforting.

Statistical mechanics is based on simple assumptions, expressed as model probabilities, that lead to a wide variety of predictions in excellent agreement with experiments. How this is done is the main subject of this book.

3.2 Discrete Random Variables and Probabilities

We begin by defining a set of elementary events

$$A = \left\{ a_j | j = 1, N_A \right\} \tag{3.2}$$

and assigning a probability $P(a_j)$ to each event. The combination of random events and their probabilities is called a 'random variable'. If the number of elementary events is finite or countable, it is called a 'discrete random variable'.

The probabilities must satisfy the conditions that

$$0 \le P(a_j) \le 1 \tag{3.3}$$

for all a_i . An impossible event has probability zero, and a certain event has probability 1.

Random events can be anything: heads/tails, red/black, and so on. If the random events are all numbers, the random variable is called a 'random number'.

Elementary events are defined to be exclusive—one, and only one elementary event can occur at each trial—so that the probabilities must also satisfy the normalization condition

$$\sum_{j=1}^{N_A} P(a_j) = 1.$$
(3.4)

² In general, the probability distribution after an irreversible process is not exactly the same as the model probability, although it gives identical predictions for the probabilities of macroscopic variables. This supports the use of model probabilities for equilibrium calculations.

To simplify notation, we will often write this equation as

$$\sum_{a} P(a) = 1, \tag{3.5}$$

suppressing explicit mention of the number of elementary events.

3.3 Probability Theory for Multiple Random Variables

If more than one thing can happen at each trial we can describe the situation with two or more sets of random events. For example, both an event from A and an event from

$$B = \{b_k | k = 1, N_B\}$$
(3.6)

might occur. We can then define a joint probability $P(a_j, b_k)$ —or more simply, P(a, b)—which must satisfy

$$0 \le P(a,b) \le 1 \tag{3.7}$$

and

$$\sum_{a}\sum_{b}P(a,b) = 1.$$
(3.8)

3.3.1 Marginal and Conditional Probability

Naturally, if we have P(a, b) we can retrieve the information for either A or B alone. The marginal probability of A is defined by

$$P_A(a) = \sum_b P(a,b) \tag{3.9}$$

with a similar expression for $P_B(b)$. A nice feature of marginal probabilities is that they automatically satisfy the positivity and normalization criteria in eqs. (3.3) and (3.4).

The name marginal probability comes from the practice of calculating it in the margins of a table of probabilities. Table 3.1 shows an example of two random variables that each take on two values.

Conditional probability is the probability of an event a, given that event b has occurred. It is denoted by P(a|b), and is related to the full probability of both A and B occurring by the equations:

$$P(a,b) = P(a|b)P_B(b) \tag{3.10}$$

$$= P(b|a)P_A(a). \tag{3.11}$$

20 Discrete Probability Theory

Table 3.1 Example of a table of probabilities for independent random variables: The events of A (labeled '3' and '4') are listed down the left column, and events of B (labeled '1' and '2') across the top. The values of the probabilities are given in the four center squares. The margins contain the marginal probabilities, as defined in eq. (3.9).

| $A \setminus B$ | 1 | 2 | $P_A(a)$ |
|---------------------|-----|------|----------|
| 3 | 1/2 | 1/4 | 3/4 |
| 4 | 1/6 | 1/12 | 1/4 |
| $\overline{P_B(b)}$ | 2/3 | 1/3 | 1 |

If $P_B(b) \neq 0$, the conditional probability P(a|b) can be written as

$$P(a|b) = \frac{P(a,b)}{P_B(b)}.$$
(3.12)

Combining eqs. (3.10) and eqs. (3.11) gives us Bayes' theorem

$$P(a|b) = \frac{P(b|a)P_A(a)}{P_B(b)}.$$
(3.13)

We will discuss some of the consequences of Bayes' theorem in Section 5.5.

3.3.2 Independent Variables

A particularly interesting case occurs when the probability distribution can be written as a product:

$$P(a,b) = P_A(a)P_B(b).$$
 (3.14)

When eq. (3.14) is true, the two random variables are said to be independent because, if $P_B(b) \neq 0$, the conditional probability P(a|b) is then independent of b,

$$P(a|b) = \frac{P(a,b)}{P_B(b)} = \frac{P_A(a)P_B(b)}{P_B(b)} = P_A(a),$$
(3.15)

and P(b|a) is independent of a.

Table 3.1 gives an example of independent random variables, and Table 3.2 gives an example of random variables that are not independent.

| $\overline{C \backslash D}$ | 1 | 2 | $P_C(c)$ |
|-----------------------------|-----|------|----------|
| 3 | 0.5 | 0.25 | 0.75 |
| 4 | 0.1 | 0.15 | 0.25 |
| $P_D(d)$ | 0.6 | 0.4 | 1 |

Table 3.2 An example of a table of probabilities for random variables that are not independent. The table is arranged in the same way as Table 3.1.

3.3.3 Pairwise Independence and Mutual Independence

If we have more than two random variables, the set of random variables might satisfy two kinds of independence: pairwise or mutual.

Pairwise independence means that the marginal distribution of any *pair* of random variables can be written as the product of the marginal distributions of the individual random variables.

Mutual independence means that the marginal distribution of any *subset* of random variables can be written as the product of the marginal distributions of the individual random variables.

It is obvious that mutual independence implies pairwise independence. Whether the converse is true is the subject of a problem in this chapter.

3.4 Random Numbers and Functions of Random Variables

Given an arbitrary random variable $A = \{a_j | j = 1, \dots, N_A\}$, we can define a numerical function on the set of elementary events, $F = \{F(a_j) | 1, \dots, N_A\}$. The set of random numbers F, together with their probabilities, is then also a random number.

If we introduce the Kronecker delta function

$$\delta_{x,y} = \begin{cases} 1 & x = y \\ 0 & x \neq y \end{cases}$$
(3.16)

then we can write the probability distribution of F compactly,

$$P_F(f) = \sum_a \delta_{f,F(a)} P_A(a). \tag{3.17}$$

As a simple illustration, consider a random variable that takes on the three values -1, 0, and 1, with probabilities $P_A(-1) = 0.2$, $P_A(0) = 0.3$, and $P_A(1) = 0.5$. Define a

22 Discrete Probability Theory

function F(a) = |a|, so that F takes on the values 0 and 1. The probabilities $P_F(f)$ are found from eq. (3.17),

$$P_F(0) = \sum_{a=-1}^{1} \delta_{0,F(a)} P_A(a)$$

= $\delta_{0,F(-1)} P_A(-1) + \delta_{0,F(0)} P_A(0) + \delta_{0,F(1)} P_A(1)$
= $0 + P(0) + 0 = 0.3$ (3.18)

$$P_F(1) = \sum_{a=-1} \delta_{1,F(a)} P_A(a)$$

= $\delta_{1,F(-1)} P_A(-1) + \delta_{1,F(0)} P_A(0) + \delta_{1,F(1)} P_A(1)$
= $P_A(-1) + 0 + P_A(1) = 0.2 + 0.5 = 0.7.$ (3.19)

We can also use the Kronecker delta to express arbitrary combinations of random numbers to form new compound random numbers. For example, if X and Y are random numbers, and G(x,y) is an arbitrary function, we can define a new random variable, Z. The probability distribution of Z is given by a sum over all combinations of the events of X and Y, with the Kronecker delta picking out the ones that correspond to particular events of Z

$$P_{Z}(z) = \sum_{x} \sum_{y} \delta_{z,G(x,y)} P(x,y).$$
(3.20)

A warning is necessary because the limits on the sums in eq. (3.20) have been suppressed. The only difficult thing about actually doing the sums is in keeping track of those limits. Since being able to keep track of limits will also be important when we get to continuous distributions, we will illustrate the technique with the simple example of rolling two dice and asking for the probability distribution of their sum.

Example: Probability of the Sum of Two Dice

We will assume that the dice are honest, which they tend to be in physics problems, if not in the real world. Let $X = \{x | x = 1, 2, 3, 4, 5, 6\}$ be the random number representing the outcome of the first die, with Y a corresponding random number for the outcome of the second die. The sum S = X + Y. The values taken on by S range from 2 to 12. Since all elementary events are equally likely,

$$P(x,y) = P_X(x)P_Y(x) = \left(\frac{1}{6}\right)\left(\frac{1}{6}\right) = \frac{1}{36},$$
(3.21)