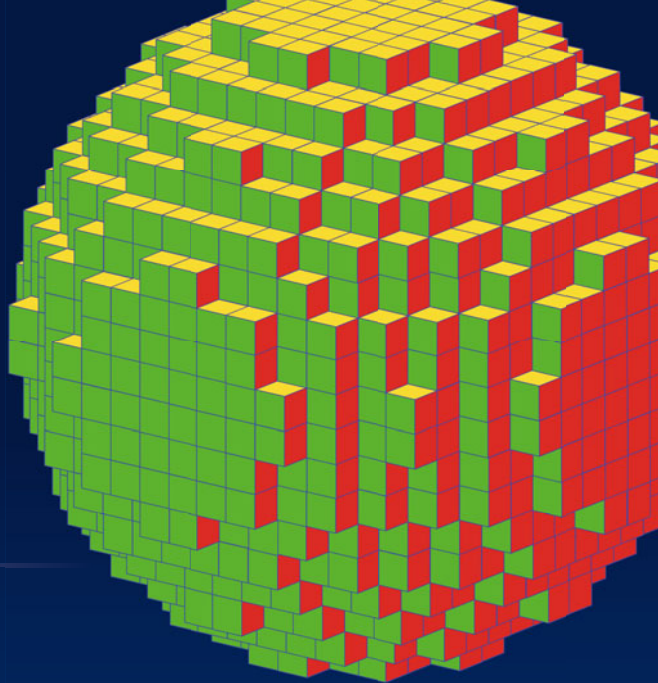


Don Koks



Microstates, Entropy and Quanta

An Introduction to Statistical Mechanics



Springer

Microstates, Entropy and Quanta

Don Koks

Microstates, Entropy and Quanta

An Introduction to Statistical Mechanics

ISBN 978-3-030-02428-4

ISBN 978-3-030-02429-1 (eBook)

<https://doi.org/10.1007/978-3-030-02429-1>

Library of Congress Control Number: 2018960736

© Springer Nature Switzerland AG 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Cover art: Central to statistical mechanics is the idea of counting the states accessible to a system. When these states exist within some continuous space, they cannot be counted. Instead, we “tile” the space into cells, with each cell defining a state, and then we count those cells. The ball on the front cover is a schematic of this tiling of the velocity space of a free particle that moves in three spatial dimensions. For real particles, the cells are so much smaller than the size of the ball that, to all intents and purposes, the ball is a smooth sphere. The number of cells can then easily be found from the sphere’s volume.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

*For my ancestors, and all those
who have gone before.*

Preface

Another book on introductory statistical mechanics? You might think that a century-old subject would have nothing left unsaid; but that is perhaps not the case. Unlike most other fields of physics, one can compare a dozen books on statistical mechanics and find a dozen different approaches to the discipline. At one extreme are authors who revel in arcane abstraction, but whose books go mostly unread. At the other extreme are very readable books that lack the mathematics to carry the reader very far beyond a set of physical assumptions. Most readers are looking for something in between; but that space is vast and subjective, with plenty of room for another book to aim for the “Goldilocks Zone” of being just right. That’s why I wrote this book: I think that the field of introductory statistical mechanics still has plenty of scope for an author to try a different mix of mathematical exposition and physical reasoning.

The physics part of this mix is a build of statistical mechanics from the ground up, anchored to a bedrock of physical concepts. With this approach, I hope to have revealed the necessity and importance of the subject’s core ideas, such as entropy and temperature. The mathematics part of the mix has been an emphasis on a strong logical reasoning that has a clean outline, yet avoids the notational clutter and obscure discussions that are so often associated with statistical mechanics, and which can make it so hard to learn.

Thus, beside the calculations of representative physical quantities, you will find here various mathematical analyses that I believe are important to physicists. Much of this mathematical foundation is given in the first chapter, such as details of integrating the gaussian function, and the correct use of infinitesimals, partial derivatives, and units of measurement. By the time you reach that chapter’s end, you might be wondering whether you are really reading a book on statistical mechanics after all. And yet, you will encounter those topics time and again as you work through the rest of the book.

The choice of how and where to begin describing a subject is always highly author dependent. The concepts that I introduce methodically, as needed,

are sometimes merely postulated with a breezy stroke of the pen in books that announce themselves as introductions. Postulatory approaches to other subjects can certainly work well; for instance, I admire Feynman's approach to electromagnetism in his *Lectures on Physics*, since, although he postulates Maxwell's equations at the very start, we never lose sight of the physics in his discussions. In contrast, I struggle to see *any* physics at all in some postulatory approaches to statistical mechanics, which can so easily ignore the difficult questions that interest physicists.

I commence the subject of statistical mechanics with an archetypal observation: why does a drop of ink placed in a bathtub disperse? Once dispersed, might it ever re-assemble into a drop? This question showcases the importance of counting the number of ways in which a system's constituents can be arranged, and leads to statistical mechanics proper via its fundamental postulate. That discussion demands knowledge of the concept of energy, a concept that was useful and intriguing to early astronomers studying planetary orbits, but whose wider application was not well understood in the early days of thermodynamics, 150 years ago. With a more modern understanding of energy (or perhaps "acceptance" is a better word, since we still don't know what it is—if, indeed, asking what it is has any meaning), we are in a good position to write down the laws of thermodynamics. Then, we can explore heat engines, chemical processes and equilibria, and heat flow. The flow of heat is a stepping stone to appreciating diverse related areas, such as particle diffusion and, in fact, the signal processing performed in a modern radar receiver.

But no system is ever truly isolated; and the question of how to analyse a system in contact with the wider world brings us to the Boltzmann distribution, with examples in paramagnetism, atomic energy levels, molecular and crystal heat capacities, and data-transmission theory. The Boltzmann distribution also sheds light on the motion of gas particles. I use that theory to explore an atmosphere, as well as the molecular details of viscosity and thermal conductivity.

Quantum ideas then emerge, via Einstein's and Debye's theories of heat capacity. The notion of fermions and bosons forms a springboard to the study of electronic heat capacity, electrical conduction, thermal noise in electric circuits, the spectra of light produced by hot bodies, some cosmology, the greenhouse effect, and the modern technologies of light-emitting diodes and the laser.

I have sprinkled the text with occasional short digressions, discussing topics such as the factorial function in number theory, the energy-momentum tensor in relativity, a little bit of signal processing, and decrying the shortcomings of modern analytical astronomy. Hopefully, these asides will only enrich your interest without being a distraction.

Unlike some books on statistical mechanics, I have chosen to discuss a lot of material before introducing the Boltzmann distribution. Thus, in those pre-Boltzmann chapters, I invoke the equipartition theorem to approximate the

particles in an ideal gas of temperature T as each having, say, translational energy $\frac{3}{2}kT$. Later, when studying the Boltzmann distribution, we learn that only their *average* translational energy is $\frac{3}{2}kT$. Some authors will avoid this initial simplification by introducing the Boltzmann distribution very early on. But I think that using the simple approximation initially, and leaving Boltzmann for later, is useful pedagogically.

A subject as old as statistical mechanics is bound to carry baggage picked up along the way, created as a normal part of its development, when physicists and chemists were searching for the best path through the new forest they had discovered. The choice of what might best be discarded is a little subjective. I have tended to minimise the use of phrases and topics that appear to be generally confusing, unattractive, or not useful.

For example, I cannot imagine that the conventional vernacular that describes various flavours of ensemble, along with free energies, partition functions, and Maxwell relations, does anything to attract new adherents to statistical mechanics. Pseudo wisdom that you can find in books on the subject, such as “The trick to solving this problem is to use the *grand* canonical ensemble”, is apt to give the impression that statistical mechanics is all about finding the right trick using the right ensemble to get the right answer. The language of ensembles is not especially deep, and after explaining what it means and how it’s used, I tend to avoid it, because the “correct ensemble to use” should be clear from the context being studied; it is not some arbitrary choice that we make. Free energies have a range of uses in thermodynamics (and I certainly use them in this book), but they are probably more relevant to the history of the subject, when early physicists and chemists worked hard to ascertain the nature of what was then a cutting-edge new quantity called energy. Nowadays, we view free energies as useful combinations of more fundamental parameters such as energy, temperature, and entropy. I think that the use of partition functions can be minimised in a book on introductory statistical mechanics: they are *sufficient* but not *necessary* to the field; and yet, all too often, books seem to suggest that the partition function is the answer to every problem. Lastly, Maxwell relations are a useful but straightforward application of basic partial derivatives to thermodynamics, and that they have a name at all is probably just historical. More generally, long parades of partial derivatives and an endless swapping and tabulation of independent variables appear in so many books on statistical mechanics. These relics of history are best left to museums. No one *really* uses them!

One deliberate nuance in some of my numerical calculations should be explained: I tend not to choose nicely rounded versions of some parameters that turn up again and again. For example, I use 298 kelvins for room temperature and 9.8 m/s^2 for Earth’s gravity, instead of the simpler-looking rounded values of 300 kelvins and 10 m/s^2 . The reason here is that, when you see “298” and “9.8” in a calculation, you will perhaps find it easier to digest the various parameters quickly through recognising those quirky numbers at a glance, as opposed to seeing the more generic-looking numbers 300 and 10.

Also, whenever I have a quantity written in both non-bold and bold font in one context—such as “ a ” and “ $\mathbf{a}$ ”—then a should be understood to be the length of the vector $\mathbf{a}$.

This book has benefitted from the contributions of my family, friends, and colleagues—although, of course, I claim full ownership of my often-strong opinions about physics in general. My undergraduate lecturers at Auckland University, Graeme Putt and Paul Barker, provided my first instruction in thermodynamics and statistical mechanics in the mid-1980s, and so laid out the backbone for a later set of lectures of my own, which became this book. All manner of details were donated by others. Brad Alexander gave me a computer scientist’s view of entropy. Colin Andrew discussed scuba diving and ocean pressure. Shayne Bennetts listened to my views on the principle of detailed balance. Encouragement and some discussion of grammar came from Ine Brummans. The modern puzzle that is liquid helium was spelled out for me by Peter McClinton. I discussed some ideas of presentation with Steven Cohen. Occasional technical discussions took place with Scott Foster. Roland Keir contributed his knowledge of physical chemistry. Harry Koks informed me of some evolved wording in combinatorics, and Rudolf Koks explained osmosis in humans. Mark Krieg improved my grammar. Hans Laue discussed atmospheric temperature. Nadine Pesor helped me settle on the use of some jargon. Robert Purvinskis was a sounding board on occasion. Andy Rawlinson gave feedback on many ideas. Keith Stowe helped untangle some knotty problems in the subject. Vivienne Wheaton prompted some early deliberation on the Boltzmann distribution. The feedback of two anonymous early referees certainly helped make a better final product. Springer’s proof reader, Marc Beschler, gave a final and detailed burnish to my words. And the entire text was much improved by the careful reading and many thoughtful suggestions of Alice von Trojan.

Beyond that, I thank Springer’s Tom Spicer for having the confidence to allow the project to go ahead, and Cindy Zitter for the details of making it happen.

Adelaide, Australia
August 2018

Don Koks

Contents

Preface	vii
1 Preliminary Ideas of Counting, and Some Useful Mathematics	1
1.1 The Spreading of an Ink Drop	2
1.1.1 Identical-Classical Particles	6
1.2 Wandering Gas Particles	7
1.3 Fluctuations in the Binomial Distribution	17
1.3.1 Expected Value and Standard Deviation of a Random Variable	18
1.3.2 The Random Walk	23
1.4 Gaussian Approximation of the Binomial Distribution	25
1.5 Integrals of the Gaussian Function	30
1.5.1 Calculating the Error Function Numerically	36
1.5.2 The 3-Dimensional Gaussian	37
1.6 Increases and Infinitesimals	39
1.6.1 Basis Vectors	48
1.6.2 The Probability Density	50
1.7 Exercising Care with Partial Derivatives	53
1.8 Exact and Inexact Differentials	60
1.9 Numerical Notation, Units, and Dimensions	64
1.9.1 Units versus Dimensions	67
1.9.2 Function Arguments Must Be Dimensionless	75
1.9.3 Distinguishing Between an Entity and its Representation	80
2 Accessible States and the Fundamental Postulate of Statistical Mechanics	83
2.1 States and Microstates	83

2.2	Energy Spacing of States	87
2.3	Position–Momentum and Phase Space	91
2.4	Microstates Are Cells of Phase Space	95
2.4.1	A System’s Quadratic Energy Terms	112
2.4.2	When Particles are Identical Classical	114
2.5	The Density of States	116
2.6	Ω_{tot} for Massless Particles	119
3	The Laws of Thermodynamics	125
3.1	The Concept of Energy for a Central Force	125
3.2	Force and Potential Energy	132
3.3	Interaction Types and the Zeroth Law of Thermodynamics	136
3.4	The First Law of Thermodynamics	138
3.4.1	Expressions for Quasi-Static Mechanical Work	140
3.4.2	The δC Term and Chemical Potential	146
3.5	The Definition of Temperature	148
3.5.1	Accessible Microstates for Thermally Interacting Systems	149
3.5.2	Temperature and the Equipartition Theorem	152
3.6	The Ideal Gas and Temperature Measurement	154
3.6.1	Measuring Temperature: the Constant-Volume Gas Thermometer	159
3.6.2	Temperature of Our Upper Atmosphere	161
3.7	The Non-Ideal Gas and van der Waals’ Equation	162
3.8	Entropy and the Second Law of Thermodynamics	167
3.8.1	Entropy of an Ideal Gas of Point Particles	170
3.8.2	The Canonical Example of Entropy Growth	171
3.8.3	Reversible and Cyclic Processes	174
3.8.4	The Use of Planck’s Constant for Quantifying Entropy	176
3.9	Can Temperature Be Negative?	177
3.10	Intensive and Extensive Variables, and the First Law	181
3.11	A Non-Quasi-static Process	183
3.12	The Ideal-Gas Law from Entropy	185
3.13	Relation of Entropy Increase to Interaction Direction	187
3.14	Integrating the Total Energy	191
3.14.1	Swapping the Roles of Conjugate Variables	193
3.14.2	Maxwell Relations	198
3.15	Excursus: Pressure and Temperature of a Star’s Interior	200

4	The First Law in Detail	207
4.1	The First Term: Thermal Interaction	207
4.1.1	The Third Law of Thermodynamics	216
4.1.2	Heat Flow and the Thermal Current Density	220
4.1.3	The Continuity Equation	224
4.1.4	The Heat Equation, or Diffusion Equation	226
4.2	The Second Term: Mechanical Interaction	233
4.2.1	Heat Engines and Reversibility	233
4.2.2	The Joule–Thomson Process	238
4.3	The Third Term: Diffusive Interaction	247
4.3.1	Pressure and Density of the Atmosphere	247
4.3.2	Pressure and Density of the Ocean	251
4.3.3	Pressure and Density from the Chemical Potential	255
4.3.4	Phase Transitions and the Clausius–Clapeyron Equation	257
4.3.5	Chemical Equilibrium	270
5	The Non-Isolated System: the Boltzmann Distribution	275
5.1	The Boltzmann Distribution	277
5.1.1	The Exponential Atmosphere Again	278
5.2	Paramagnetism	279
5.3	Energy Levels, States, and Bands	283
5.4	Hydrogen Energy Levels	287
5.5	Excitation Temperature	290
5.6	Diatomic Gases and Heat Capacity	291
5.6.1	Quantised Rotation	292
5.6.2	Quantised Vibration	298
5.7	Another Look at the Hydrogen Atom	300
5.8	Equipartition for a System Contacting a Thermal Bath	303
5.8.1	Fluctuation of the System’s Energy	306
5.9	The Partition Function in Detail	307
5.10	Entropy of a System Contacting a Thermal Bath	312
5.11	The Brandeis Dice	319
5.12	Entropy and Data Transmission	324
6	The Motion of Gas Particles, and Transport Processes	333
6.1	The Maxwell Velocity Distribution	338
6.1.1	Alternative Derivation of the Velocity Distribution	342
6.2	The Maxwell Speed Distribution	344
6.2.1	Alternative Derivation of the Speed Distribution	345

6.3	Representative Speeds of Gas Particles	346
6.4	Doppler Broadening of a Spectral Line	350
6.5	Temperature Gradient in a Weatherless Atmosphere	353
6.6	Gaseous Makeup of Planetary Atmospheres	358
6.7	Mean Free Path of Gas Particles	365
6.7.1	Excursus: The Proof of (6.123)	368
6.8	Viscosity and Mean Free Path	371
6.9	Thermal Conductivity and Mean Free Path	376
6.10	Excursus: The Energy–Momentum Tensor	379
7	Introductory Quantum Statistics	385
7.1	Einstein’s Model of Heat Capacity	385
7.2	A Refinement of Einstein’s Model of Heat Capacity	389
7.3	Debye’s Model of Heat Capacity	394
7.4	Gibbs’ Paradox and Its Resolution	403
7.5	The Extent of a System’s Quantum Nature	405
7.5.1	Average de Broglie Wavelength	407
7.6	Fermions and Bosons	409
7.7	Occupation Numbers of Fermion and Boson Gases	418
7.7.1	Calculating $\mu(T)$ and $\bar{n}(E, T)$ for Fermions	421
7.7.2	Calculating $\mu(T)$ and $\bar{n}(E, T)$ for Bosons	424
7.8	Low-Temperature Bosons and Liquid Helium	426
7.9	Excursus: Particle Statistics from Counting Configurations ..	432
7.9.1	Fermi–Dirac and Bose–Einstein from Configurations ..	442
8	Fermion Statistics in Metals	445
8.1	Conduction Electrons’ Contribution to Heat Capacity	445
8.1.1	A More Accurate Approximation of $\bar{n}(E, T)$	454
8.2	Electrical Conductivity of Metals	458
8.3	Thermal Conductivity of Metals	465
8.3.1	The Lorenz Number	467
8.4	Insulators and Semiconductors	468
8.5	Diodes	473
9	Boson Statistics in Blackbody Radiation	481
9.1	Spectrum of Radiation Inside an Oven	482
9.1.1	Mean “Extractable” Energy of an Oscillator, $\varepsilon(f)$	484
9.2	The One-Dimensional Oven: an Electrical Resistor	488
9.2.1	Calculating the Density of Wave States, $g(f)$	489

9.2.2	Excursus: Thermal Noise in a Resistor, and Some Communications Theory	496
9.3	The Three-Dimensional Oven	502
9.4	The End Product: Planck's Law	506
9.4.1	Planck's Law Expressed Using Wavelength	508
9.5	Total Energy of Radiation in the Oven	510
9.6	Letting the Radiation Escape the Oven	511
9.7	Blackbody Radiation	514
9.7.1	The Large-Scale Universe Is a Very Cold Oven	517
9.7.2	Total Power Emitted by a Black Body	520
9.8	The Greenhouse Effect	523
9.9	Photon Absorption and Emission: the Laser	525
Index		535

List of Common Symbols

Chapter 1: Preliminary Ideas of Counting, and Some Useful Mathematics

N_A	Avogadro's number.
C_x^n	Number of combinations (selections) of x objects taken from a total of n objects.
$\mathcal{N}(x; \mu, \sigma^2)$	Normal distribution of x , with mean μ and variance σ^2 .
$\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, P)$	Multi-dimensional normal distribution of $\mathbf{x}$, with mean $\boldsymbol{\mu}$ and covariance matrix P .
$\mathbf{e}_q$	Basis vector for coordinate q .
$\mathbf{u}_q$	Unit-length basis vector for coordinate q .
dx	Infinitesimal, an “exact differential” of state variable x .
$\text{d}A$	Infinitesimal, an “inexact differential” of quantity A that is not a state variable.
$\lambda(x)$	Linear mass density as a function of position x .
M_{mol}	Molar mass.
$[L]_S$	Representation of L in system S .

Chapter 2: Accessible States and the Fundamental Postulate of Statistical Mechanics

D	Number of internal variables in which a particle can store its energy.
$\Omega(E)$	Number of microstates that each have energy E .
$\Omega_{\text{tot}}(E)$	Total number of microstates that each have energy somewhere in the range 0 to E .
ν	Number of quadratic energy terms of a particle, meaning the number of quadratic coordinates that describe the particle's energy. (In other texts, this is called the number of degrees of

	freedom of the particle.) The particle need not be an atom; it could be a molecule.
$\Omega_{\text{tot}}^{\text{ic}}(E)$	$\Omega_{\text{tot}}(E)$ for identical-classical particles.
$g(E), g(f)$	Density of states as functions of energy E and frequency f .

Chapter 3: The Laws of Thermodynamics

$\mathbf{u}_r$	Unit-length radial vector.
$b(r)$	A function describing the central force as a function of radial distance r .
U	Potential energy of a particle.
$\text{d}Q$	Thermal energy put into a system.
$\text{d}W$	Mechanical work performed on a system.
$\text{d}C$	Energy brought into a system by incoming particles or environmental changes.
$\mathbf{E}, \mathbf{p}$	Electric field and electric dipole moment.
$\mathbf{B}, \boldsymbol{\mu}$	Magnetic field and magnetic dipole moment.
μ	Chemical potential.
γ_i	$\nu_i N_i / 2$, where ν_i is the number of quadratic energy terms (“degrees of freedom” in other texts) per particle in system i , and N_i is the number of particles in system i .
N	Number of particles.
n	Number of moles.
$\hat{E}_i$	Value of energy E_i at which $\Omega(E_i)$ peaks.
$\Omega_{\text{tot}}^{\text{dist}}$	Ω_{tot} for distinguishable particles.
S^{dist}	Entropy of distinguishable particles.
S^{ic}	Entropy of identical-classical particles.
F	Helmholtz energy.
G	Gibbs energy.
H	Enthalpy.
κ	Coefficient of isothermal compressibility.
β	Coefficient of thermal expansion.

Chapter 4: The Three Interactions of the First Law

C_P, C_V	Heat capacities at constant pressure and volume.
C^{sp}	Specific heat capacity.
C^{mol}	Molar heat capacity.

γ	C_P/C_V , which also equals $C_P^{\text{sp}}/C_V^{\text{sp}}$ and $C_P^{\text{mol}}/C_V^{\text{mol}}$.
μ_{JT}	Joule–Thomson coefficient.
a, b	Van der Waals parameters.
$a_{\text{mol}}, b_{\text{mol}}$	Van der Waals parameters for the molar form of van der Waals' equation.
$\mathbf{J}$	Current density, also known as flux density.
κ	Thermal conductivity.
$(\mathbf{a}, \mathbf{b})$	Angle between vectors $\mathbf{a}$ and $\mathbf{b}$ (between 0 and π).
I	Heat current.
R	Thermal resistance.
ϱ	Thermal resistivity.
ϱ_E	Energy content per unit volume.
ϱ_m	Mass per unit volume.
K	Diffusion constant.
$*$	Convolution operator.
ϱ	Mass per unit volume.
ν	Number of particles per unit volume.
B	Bulk modulus.
ϕ	Ratio of salt particles to total number of salt and water particles.
$L_{\text{vap}}^{\text{mol}}, L_{\text{fusion}}^{\text{mol}}$	Molar latent heats of vaporisation and fusion.

Chapter 5: The Non-Isolated System: the Boltzmann Distribution

$p_{\text{level } n}$	Probability that a system occupies any state at energy level n .
$p_{\text{state } n}$	Probability that a system occupies a specific state n .
β	$1/(kT)$.
E_n, V_n	Energy and volume of a hydrogen atom at energy level n .
Z	Partition function.
T_e	Excitation temperature of a system.
T_R, T_V	Characteristic temperatures of the onsets of rotation and vibration.
$\bar{E}$	Abbreviated version of $\bar{E}_s$, the mean energy of the system.

Chapter 6: The Motion of Gas Particles and Transport Processes

$v, \mathbf{v}$	Speed and velocity of a particle.
d^3v	$\text{d}v_x \text{d}v_y \text{d}v_z$.

$\mathcal{N}_{\text{vel}}(\mathbf{v}) d^3v$	Infinitesimal number of particles with velocities in the range $\mathbf{v}$ to $\mathbf{v} + d\mathbf{v}$.
$\mathcal{N}_x(v_x) dv_x$	Infinitesimal number of particles with x velocities in the range v_x to $v_x + dv_x$.
$\mathcal{N}_{\text{sp}}(v) dv$	Infinitesimal number of particles with speeds in the range v to $v + dv$.
N_{tot}	Total number of particles.
$d\Omega_{\text{tot}}$	Number of microstates in the energy range E to $E + dE$.
$\mathcal{N}_z(z, v_z) dz dv_z$	Infinitesimal number of particles with heights in z to $z + dz$, and z velocities in v_z to $v_z + dv_z$.
$\mathcal{N}_{\text{sp}}(z, v) dz dv$	Infinitesimal number of particles with heights in z to $z + dz$, and speeds in v to $v + dv$.
λ	Mean free path.
ν	Number of particles per unit volume.
σ	Collision cross section.
η	Coefficient of viscosity.
κ	Thermal conductivity.

Chapter 7: Introductory Quantum Statistics

n	The energy level of a one-dimensional oscillator, and also the number of quantum particles per state, in which each state denotes one dimension of oscillation of a single molecule in a crystal.
$\bar{E}$	Mean energy of a crystal molecule (a quantised oscillator that can oscillate in three dimensions). Eventually redefined to exclude zero-point energy.
$\bar{E}_{1D}$	Mean energy of a one-dimensional quantised oscillator.
T_E, T_D	Einstein and Debye temperatures.
$\bar{n}$	Occupation number of a crystal, the arithmetic mean of n : the mean number of quantum particles present per 1D-oscillator in the crystal. A function of temperature.
E_{tot}	Total energy of all oscillators in the crystal.
$\bar{n}(E, T)$	Occupation number treated as a function of energy and temperature: the mean number of quantum particles per state.
N	Number of quantum particles with energies up to E .
λ	De Broglie wavelength.
E	Energy of a state, which that state “bestows” on each particle occupying it.
p_n	Probability of n quantum particles being present in a state.

N	Total number of quantum particles of all energies (in a later section to N immediately above).
C	A constant for a gas of massive particles, encoding spin, volume, and particle mass.
E_F	Fermi energy.
T_c	Critical temperature of liquid helium.
N	Total number of balls to be placed on shelves.
n_i	Number of balls on shelf i .

Chapter 8: Fermion Statistics in Metals

$C_V^{\text{mol}}(\text{electrons})$	Valence-electron contribution to a crystal's molar heat capacity.
$\bar{E}$	Mean energy of one valence electron.
E	Energy of one valence electron.
N	Total number of valence electrons.
$\bar{n}(E, T)$	Occupation number of valence electrons.
T_F	Fermi temperature of valence electrons.
v_F	Fermi speed of valence electrons.
α	A number in the region of 1 or 2, modelling the characteristic width of the fall-off of the Fermi–Dirac occupation number with energy.
ϱ	Electrical resistivity.
κ	Thermal conductivity.
N_e	Number of electrons in a conduction band.

Various parameters are also defined in (8.65).

Chapter 9: Boson Statistics in Blackbody Radiation

$\varrho(f)$	Spectral energy density as a function of frequency f .
$\varepsilon(f)$	Mean energy of a single oven-wall oscillator of frequency f .
λ_0	Wavelength corresponding to the peak radiated power from a black body.
σ	Stefan–Boltzmann constant.
$\mathcal{N}$	Number of frequency- f photons produced by any process in a laser.

List of Common Constants

Avogadro's number	N_A	6.022×10^{23} (or $6.022 \times 10^{23} \text{ mol}^{-1}$)
Boltzmann's constant	k	$1.381 \times 10^{-23} \text{ J/K}$
Gas constant	$R = N_A k$	8.314 J/K (or $\text{J K}^{-1} \text{mol}^{-1}$)
Planck's constant	h	$6.626 \times 10^{-34} \text{ J s}$
	$\hbar = h/(2\pi)$	$1.0546 \times 10^{-34} \text{ J s}$
Speed of light in vacuum	c	$2.998 \times 10^8 \text{ m/s}$
Proton mass		$1.67 \times 10^{-27} \text{ kg}$
Electron mass		$9.11 \times 10^{-31} \text{ kg}$
Electron charge	$-e$	$-1.602 \times 10^{-19} \text{ C}$
Electron volt		$1 \text{ eV} = 1.602 \times 10^{-19} \text{ J}$
Room temperature		$298 \text{ K (} 25^\circ\text{C)}$
Ground temperature		$288 \text{ K (} 15^\circ\text{C)}$
Air temperature (mean atmospheric)		$253 \text{ K (} -20^\circ\text{C)}$
Mass of a generic air molecule		$4.8 \times 10^{-26} \text{ kg}$
Molar mass of air		29.0 g
Molar mass of sodium		23.0 g
Earth gravity		9.8 m/s^2
Atmospheric pressure at sea level		$101,325 \text{ Pa}$
Copper's molar mass		63.5 g
Copper's Fermi temperature		$81,000 \text{ K}$
Number density of copper's valence electrons		$8.47 \times 10^{28} \text{ m}^{-3}$



Chapter 1

Preliminary Ideas of Counting, and Some Useful Mathematics

In which we set the stage for counting the number of ways in which a system can arrange itself, derive some mathematical and statistical results that will be useful later, examine the meaning of infinitesimals and partial derivatives, and study how to use units correctly and efficiently.

The modern subject of statistical mechanics is built on a single, simple idea: that much of the physics and chemistry of many-particle systems can be deduced from ideas of counting and probability. Classical physics contains no randomness, and yet a probabilistic description turns out to be very capable of analysing systems composed of a large number of entities. This is not mysterious; after all, if we flip a coin a large number of times, we fully expect that roughly half of the flips will yield heads, even though predicting the outcome of each flip is so complex as to be effectively impossible. Physical systems that can be analysed in a similar way are all around us, and they are too complex to analyse from first principles using Newton's laws. Historically, this idea of representing large, complex systems by a small set of parameters gave rise to thermodynamics, which relied on concepts such as pressure and temperature. Although thermodynamics could analyse pressure by postulating the existence of large numbers of moving atoms (or molecules), it was silent on what lay behind temperature. Only with the advent of statistical mechanics did we gain insight into this and many other areas relevant to large—and not so large—systems. And, in a modern age, statistical mechanics has spawned the field of quantum statistical mechanics, which explains the inner workings of much of our modern technology.

To begin to investigate statistical mechanics, we must know something about counting in the broader sense, and the purpose of this first chapter is to present a set of useful tools to do that. Physicists are generally more fascinated by the physical world than by counting combinations and permutations, and so some of the calculations of this first chapter might come across as a little arduous. But they lay a secure foundation on which statistical mechanics rests: this being the idea that the averaging that results from combining many atoms to make up our world is what makes that world predictable, despite our lack of knowledge of what each individual atom is doing at any moment.

Along the way, we will investigate several other mathematical topics that will come in handy in the chapters to follow.

1.1 The Spreading of an Ink Drop

Imagine a billiards table on which you place a single ball. You give the ball a particular initial velocity, parallel to the sides, so that it bounces back and forth on a single line from one end to the other. In the absence of friction, it will roll on this line forever, and its motion will be as simple and ordered as it could possibly be.

Now add more balls, one by one, trying to arrange for some given motion each time a ball is added. Although they continue to be governed by Newton's laws, the range of motions available to the balls becomes phenomenally larger each time we put another ball on the table. It very quickly becomes clear that the entire motion of the balls *effectively* becomes more and more random, even though it is not random at all. When we replace the balls by, say, molecules of water in a filled bathtub, the range of motions available to the molecules becomes so large that it surpasses any test or theory that mathematicians have devised to ascertain or even define randomness.

The *ergodic assumption* of classical mechanics postulates that as the number of particles in our system is increased, with their initial conditions always arranged to be highly randomised, the system will tend more and more to spend an equal amount of time in the vicinity of each arrangement of particles that is allowed by the external constraint of fixed energy.

The essence of statistical mechanics can be seen at work when we carefully place a drop of ink into a bathtub of pure water. Slowly but surely, the drop spreads out. Suppose that the bathtub is made from an esoteric material that insulates the mixture of water and ink from the wider world in every physical way. This material prevents any heat from flowing between the outside world and the water and ink. No vibration can pass through the material to or from the outside world, and particles can neither be shed from the material nor adhere to it. The energy of the water-ink mixture is thus a fixed quantity, and the ergodic assumption applies. That is, if we wait long enough, the drop will essentially reform at some time in the distant future; and, in fact, it will eventually pass arbitrarily close to any configuration that we care to describe. There will come a time when the spread of ink takes on the appearance of Vermeer's "Girl with a Pearl Earring", the complete text of every book ever written or yet to be written—and indeed, the same texts, but with one or two or 3005 letters upside down—and every sculpture and picture ever made or that will ever be made.

But we know from experience that this re-forming of the ink into a drop or dispersing into the appearance of a Vermeer painting is not likely to happen any time soon: we can be confident that no one has ever seen an ink drop form spontaneously in a tub full of a water-ink mixture, or even a *teaspoon* full of a water-ink mixture. Of course, the esoteric material that forms the tub walls that we described above simply does not exist in Nature, and so no bathtub is ever completely insulated from its environment. But this interaction with

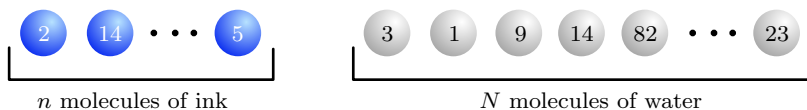


Fig. 1.1 We can count distinguishable molecules of ink and water by systematically laying them out along a line

the wider world does not affect the basic idea that only a comparatively few arrangements of the water–ink molecules are recognisable to us as forming any sort of pattern. So, as the ink molecules interact with the water molecules and become dispersed, we ask the following question: in how many ways can n molecules of ink be mixed with N molecules of water, while still retaining the visual appearance of a distinct ink drop?

This question is not easily answered, because we must be clear about what a drop is: need it have the same shape as the original drop? Must it be in the same place? It's sufficient for our purposes to count the total number of molecular arrangements that put the ink drop in a chosen place with a chosen shape in the tub, and we will ignore all molecular velocities. Although the molecules' arrangement is three dimensional, we can always construct a three-dimensional grid in the bath, use it to pick the molecules out systematically, and then lay them out along a line. Suppose we've done that for the ink drop and water molecules in such a way for the chosen place and shape of the drop, that the ink molecules all end up on the left end of the line, with the water molecules continuing to the right. Suppose too that the ink and water molecules are distinguishable, meaning we can label the ink molecules $1, 2, \dots, n$ and the water molecules $1, 2, \dots, N$. The result is the arrangement in Figure 1.1. The number of arrangements resembling the chosen drop will then be the product of factorials, $n!N!$. The total number of ways in which all $n + N$ molecules can be laid out is $(n + N)!$. Let's calculate the ratio r of the total number of possible arrangements to the number of “ink drop” arrangements, when $n \ll N$:

$$\begin{aligned}
 r &= \frac{\text{total number of arrangements}}{\text{number of ink-drop arrangements}} = \frac{(n + N)!}{n!N!} \\
 &= \frac{(n + N)(n - 1 + N)(n - 2 + N) \dots (1 + N) \cancel{N!}}{n! \cancel{N!}} \\
 &\simeq N^n / n!.
 \end{aligned} \tag{1.1}$$

We will approximate $n!$ using Stirling's rule:

$$n! \approx n^n e^{-n}, \tag{1.2}$$

and leave a proper discussion of that to Section 1.2. Then,

$$r \approx \frac{N^n}{n^n e^{-n}} = \left(\frac{Ne}{n} \right)^n. \quad (1.3)$$

The number of drops in the bathtub is $(n + N)/n = 1 + N/n$, so

$$\frac{N}{n} \simeq \frac{\text{volume of tub}}{\text{volume of a drop}} \simeq \frac{1000 \text{ mm} \times 600 \text{ mm} \times 500 \text{ mm}}{3 \times 3 \times 3 \text{ mm}^3} \simeq 1.1 \times 10^7. \quad (1.4)$$

What is the value of n ? Assume that the molar mass and density of the ink are the same as those of water: say, 18 g/mol and 1000 kg/m³. Then, calling on Avogadro's number $N_A \simeq 6.022 \times 10^{23}$,

$$\begin{aligned} n &= N_A \times \text{number of moles of ink} = N_A \times \frac{\text{mass of ink}}{\text{molar mass of ink}} \\ &= N_A \times \frac{\text{volume of ink drop} \times \text{density of ink}}{\text{molar mass of ink}} \\ &= 6.022 \times 10^{23} \times \frac{27 \text{ mm}^3 \times 1000 \text{ kg/m}^3}{18 \text{ g}} \\ &= 6.022 \times 10^{23} \times \frac{27 \times 10^{-9} \text{ m}^3 \times 1000 \text{ kg/m}^3}{0.018 \text{ kg}} \\ &\simeq 9.0 \times 10^{20}. \end{aligned} \quad (1.5)$$

Placing these values of N/n and n into (1.3) gives

$$r \simeq (1.1 \times 10^7 \times 2.718)^{9.0 \times 10^{20}} \simeq 10^{6.7 \times 10^{21}}. \quad (1.6)$$

We see here the staggeringly large number of ways in which ink can be spread out, compared to its remaining in a drop form at the place we specified. The number $10^{6.7 \times 10^{21}}$ certainly looks large, but how can we comprehend its true size? Consider that if you voice $10^{6.7 \times 10^{21}}$ as (approximately) “one million million million million...” at a normal talking speed, you will be saying the word “million” for the next *30 million million years*. Alternatively, reflect on the economy of the decimal system: when we represent the number “one million” by a “1” followed by six zeroes, with each zero just 1 cm across, the length of this string of zeroes will be six centimetres—a deceptively compact and efficient scheme for representing the idea of one million things. In contrast, writing $10^{6.7 \times 10^{21}}$ in this decimal system gives a “1” followed by a string of zeroes whose length is about 7100 light years. That’s not *how big* $10^{6.7 \times 10^{21}}$ is; rather, that’s just how long its decimal representation is when written down. The number itself is *inconceivably* large.

Here is another way of picturing the size of $10^{6.7 \times 10^{21}}$. Use (1.2) and some numerical manipulation to write $10^{6.7 \times 10^{21}}$ as a factorial:

$$10^{6.7 \times 10^{21}} \approx (3.3 \times 10^{20})!. \quad (1.7)$$

We will be content to examine the size of the “slightly smaller” number $10^{20}!$, which we’ll do by shuffling cards. Consider that a deck of 3 cards can be shuffled in $3! = 6$ ways. This is not a large number of different arrangements (permutations), of course, and certainly not large enough to form the basis of any real card game.¹ A deck of 4 cards can be shuffled in $4! = 24$ ways—still not large—but the number of ways in which a deck can be shuffled grows quickly with its size. A deck of 10 cards can be shuffled in $10! = 3,628,800$ ways, and if we count each of those arrangements at one count per second, the job will take six weeks to complete. Counting the $11!$ possible arrangements of 11 cards at the same speed will take over a year. With 12 cards, we’ll need 15 years of counting, and with 15 cards, the count will take over 41,000 years. These times are shown in Table 1.1.

Suppose that to be more efficient, we enlist the entire world’s population, with each person magically able to count *one million arrangements per second*. Counting the possible arrangements of 15 cards in this way will take 0.2 milliseconds; 20 cards will take 6 minutes; 25 cards will take 70 years, and 30 cards will take one thousand million years. What about a real deck of 52 cards? Counting its possible arrangements in this way will require 3.7×10^{44} years. If you take a deck of cards and shuffle it well, you can

Table 1.1 Time required to count the possible arrangements of cards in a deck

Number of cards in deck	Time for one person to count arrangements at 1 count/second	Time for world’s population to count arrangements, with each person counting 10^6 /second
3	6 seconds	negligible
4	24 seconds	
5	2 minutes	
10	6 weeks	
11	1.3 years	
12	15 years	0.2 milliseconds
15	41,400 years	
20	see $\rightarrow$	
25		
30		
52		3.7×10^{44} years
1000		2×10^{2544} years
10,000		$10^{35,636}$ years
10^{20}		$10^2 \times 10^{21}$ years

¹ A welcome evolutionary change to the words “permutations” and “combinations” is worth mentioning here. An effort is now being made in some high-school curricula to de-mystify combinatorics by discarding the dry jargon of these terms. Permutations are now being called “arrangements” (think of arranging flowers), and combinations are now “selections” (think of selecting books for a library). I used “arrangements” in some of the above discussion, but decided to retain the old words in other places, since my use of them is a little peripheral to the main subject anyway.

be as good as completely certain that the arrangement you are holding has *never* appeared in human history, and will *never appear again*.

This thought experiment has only taken us up to envisaging the size of $52!$, the number of permutations of a deck of 52 cards; we still need to work up to $10^{20}!$. The amount of time needed for the world's population to count the $10^{20}!$ permutations of 10^{20} cards is simply unfathomable. We can write down the number of years required (which is $10^2 \times 10^{21}$, or $10^{2,000,000,000,000,000,000}$, more accurately $10^{1,956,570,552,000,000,000,000}$), but we gain no real ground; we end up only trying to gain a feel for a new number ($10^2 \times 10^{21}$ years) that looks just like the one we started with ($r \simeq 10^{6.7 \times 10^{21}}$). No matter how we write these numbers down, they have no meaning for us.

Enlisting more people with higher count rates does not help here. Suppose we employ 10^{100} people, who each count 10^{100} permutations per second. Counting the permutations of 1000 cards will take them 10^{2360} years. For 10,000 cards: $10^{35,452}$ years. And for 10^{20} cards: $10^2 \times 10^{21}$ years, where the single significant figure used here makes that number appear to equal that in Table 1.1, but the two numbers of years *are* different. In the end, we have gained nothing by switching the counting from the world's population at one million per second, to 10^{100} people counting 10^{100} per second.

The above discussion of ever-changing ink patterns in a bathtub assumed the tub had no connection to the outside world. But a real bathtub *does* interact with the outside world, and so we can never really treat the water-ink mixture in isolation. The heat of the bathtub walls affects the liquid, tending to keep the water and ink molecules maximally mixed. And, of course, those walls interact with *their* environment, and so on outward, until we find ourselves obliged to consider the universe as a whole. Such considerations form part of the subject of *quantum cosmology*, which sits at the cutting edge of modern physics.

1.1.1 Identical-Classical Particles

The particles above were distinguishable, and thus able to be counted. In contrast, the case of truly *identical* particles is something else, because if real identical particles are crowded together sufficiently densely, they behave quite differently to distinguishable particles. We'll study such identical particles in Chapter 7, where we'll give a meaning to "crowded together". Particles that are identical and are being treated as a set, but are not crowded together, don't have a conventional name. We will call them *identical classical*.

If we have 5 *distinguishable* particles—say, they are numbered—and we wish to deal them into 5 numbered bins, then this can be done in $5!$ ways. But if we have 5 identical-classical particles and wish to deal them into those same bins, then this can be done in only *one* way. There is no notion here of

being able to swap two particles: if we try to swap two particles, we accomplish nothing: at a deep level, the final configuration is unchanged from the initial configuration.

If the ink molecules are identical among themselves, and the water molecules are identical among themselves, then both groups will certainly be identical classical. In that case, the $n!$ permutations of n ink molecules cannot be told apart, and must be counted as a single configuration; and the same is true for the N water molecules. In other words, our calculations above have over-counted by factors of $n!$ and $N!$. We conclude that the number of molecular arrangements in which the ink can form a drop, or a Girl with a Pearl Earring, or anything else, is not $n!N!$, but simply 1: there is only a *single* way in which a given configuration can exist if all of its particles are identical.

Of course, this idea of having over-counted by a total factor of $n!N!$ also applies to the case when the identical-classical particles are dispersed. There, too, we must divide the number $(n + N)!$ of distinguishable configurations by $n!N!$ to give the number of identical-classical configurations.

In summary, for the distinguishable molecules analysed in (1.1)–(1.6),

$$\begin{aligned} \text{total number of configurations} &= (n + N)! \\ \text{number of ink-drop configurations} &= n!N!. \end{aligned} \tag{1.8}$$

For the identical-classical case, we have

$$\begin{aligned} \text{total number of configurations} &= (n + N)!/(n!N!) \\ \text{number of ink-drop configurations} &= 1. \end{aligned} \tag{1.9}$$

Thus, for both cases, the sought-after ratio is

$$r = \frac{\text{total number of configurations}}{\text{number of ink-drop configurations}} = \frac{(n + N)!}{n!N!} \stackrel{(1.6)}{=} 10^{6.7 \times 10^{21}}. \tag{1.10}$$

Although the numbers of configurations are reduced overall for identical-classical particles, the *ratio* of the total number of configurations to the number of ink-blot configurations is unchanged, at $r \simeq 10^{6.7 \times 10^{21}}$.

1.2 Wandering Gas Particles

Central to statistical mechanics is the idea of counting the number of ways in which some interesting and effectively isolated system can take on various configurations. As we saw with the bath example above, counting these configurations exactly is usually a difficult or even impossible task—and one that's not always completely well defined. Physicists must do the best they can to get around such difficulties.

We will focus on calculating something more basic than the number of configurations of an ink drop in water. Given a fixed number of gas particles in a room, what is the chance of there being some specified number of those particles in a specified part of the room? More interesting still: how probable are sizeable *fluctuations* around this number if the particles move about randomly? It will turn out that for systems with large numbers of particles such as we find in everyday life, even incredibly tiny fluctuations are *very* improbable indeed. Perhaps, then, a view of the world that has randomness at its heart can still be compatible with the fact that the world does not look particularly random.

Figure 1.2 shows a room with imaginary partitions that divide it into four parts. It contains a gas of 14 particles that are free to move about, effectively randomly. We will take the particles to be *distinguishable*, meaning that, to all intents and purposes, they have numbers written on them. At any given moment, what is the chance that 3 particles will be found in part 1, 4 particles in part 2, and so on, as shown in the figure? We don't care which numbers are found where; we focus only on 3 particles being found in part 1, 4 particles in part 2, and so on. To begin to solve this problem, it's easier to think of dividing the room into just *two* parts. This is the job of the *binomial distribution*. Given N distinguishable particles, place each of them randomly into either of the two parts (usually called “bins” in combinatorial theory). The chance of a particular particle being allocated to bin 1 is p_1 , and the chance of a particular particle being allocated to bin 2 is $p_2 = 1 - p_1$. When all N particles have been placed in the bins, we ask: what is the chance $P(n_1)$ that n_1 particles are found in bin 1 (and hence $n_2 = N - n_1$ particles are found in bin 2), with no regard for the numbers that are written on them?

Focus first on a particular set of n_1 particles ending up in bin 1: say, particles 1, 2, and 4. We are not concerned with the order of the particles here;

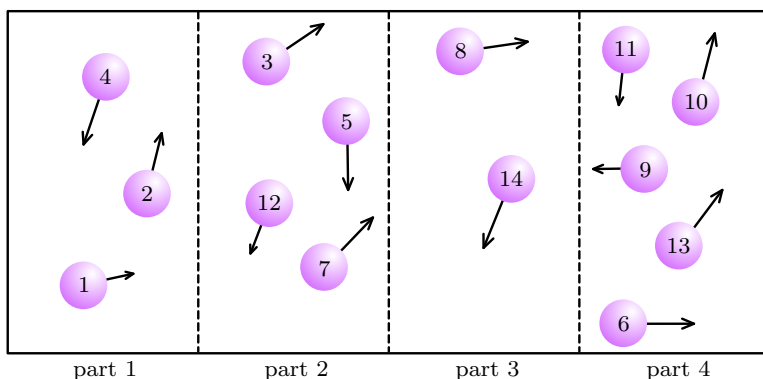


Fig. 1.2 Numbered particles are confined to a room in which they can move freely. If we partition the room as shown, what is the chance that, at some given moment, 3 particles will be found in part 1, 4 particles in part 2, and so on?

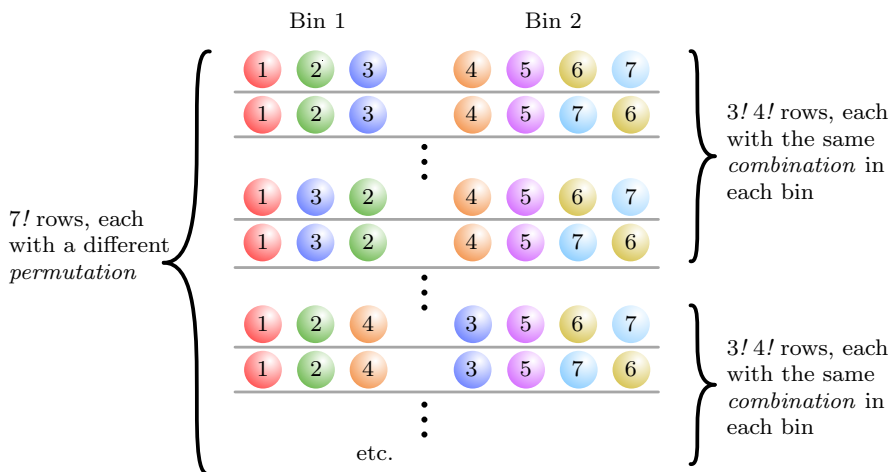


Fig. 1.3 Counting combinations by the “trick” of counting permutations and then correcting for the resultant over-counting

if particles 1, 2, 4 are found in bin 1, we don’t distinguish between referring to them as “1, 2, 4” or “1, 4, 2” or “2, 1, 4”, etc. Recall that this set of three numbers with no regard for their order is a *combination* of those numbers. If we insist on a particular order, then that ordered set of numbers is a *permutation* of those numbers. What, then, is the chance that the combination of particles 1, 2, and 4 ends up in bin 1, with the rest in bin 2? We throw the particles toward the bins one by one. The chance that particle 1 ends up in the correct bin (bin 1) is p_1 . The chance that particle 2 ends up in the correct bin is also p_1 , that particle 3 ends up in the correct bin (bin 2) is p_2 , and so on. Multiplying these probabilities then gives the chance that particles 1, 2, and 4 end up in bin 1 and the rest in bin 2 as $p_1^3 p_2^4$.

Clearly, the chance that a particular combination of n_1 particles ends up in bin 1 and the remaining n_2 particles in bin 2 is $p_1^{n_1} p_2^{n_2}$. To complete our task of finding $P(n_1)$, we need only multiply this individual probability by the total number of such possible combinations. We begin to do this by writing down all possible combinations of 3 particles that can be found in bin 1 and 4 particles in bin 2. In combinatorial theory, when you want to count something, you start listing all possible configurations, and you will often begin to see an efficient way to make this list without actually writing it down fully. In our case, we will use a kind of “trick” of over-counting: we will list all *permutations*, meaning that we *will* distinguish between “1, 2, 4” and “1, 4, 2”, and so on. Later, we will correct for having listed too many possibilities.

Figure 1.3 shows such a listing, with each number allocated a unique colour to aid in the visualisation. With $N = 7$ particles in total, $n_1 = 3$ of which must

appear in bin 1, we list (schematically in the figure) all $7!$ permutations: one permutation per row, with the balls in bin 1 written first, then a space, and then bin 2.

Now, each *combination* of a given set of three numbers in bin 1 (say, 1, 2, 3) and the remainder in bin 2 (hence 4, 5, 6, 7) appears $3!4!$ times, and so the total number of *permutations*, $7!$, over-counts the number of *combinations* by this factor. It follows that the required number of combinations of 3 particles in bin 1 and 4 particles in bin 2 is $7!/(3!4!)$. Alternatively, we could focus on bin 1 and note that there are $7 \times 6 \times 5 = 7!/4!$ ways of putting three numbered particles into it if we take order into account (i.e., count permutations). Then, to count combinations instead, we must correct for the fact that each combination “is equivalent to” $3!$ permutations. We do this by dividing the number of permutations $7!/4!$ by $3!$, to arrive at $7!/(3!4!)$.

For a general N and n_1 (with $n_2 = N - n_1$), the total number of combinations is $N!/(n_1!n_2!)$. We write this as $C_{n_1}^N$, noting that $C_{n_1}^N = C_{n_2}^N$. Each of these combinations occurs with probability $p_1^{n_1} p_2^{n_2}$. The final sought-after probability that we’ll find *any* n_1 particles, without regard for order, in bin 1, and the rest likewise in bin 2, is then

$$P(n_1) = \frac{N!}{n_1!n_2!} p_1^{n_1} p_2^{n_2} = C_{n_1}^N p_1^{n_1} p_2^{n_2}. \quad (1.11)$$

This function is the binomial distribution. (More conventionally, the function notation on the left-hand side of (1.11) will mention N and p_1 , but the simple form $P(n_1)$ suffices for our discussion.)

When more than two bins are present—such as the four in Figure 1.2—the binomial distribution generalises easily to the *multinomial distribution*. Label N distinguishable balls 1, 2, $\dots$, N and allocate each to one of M bins; then apply the approach outlined in Figure 1.3 to count the number of ways of ending up with n_1 particles in bin 1 with no regard for order, n_2 particles in bin 2 with no regard for order, and so on up to n_M particles in bin M with no regard for order. In the same way that each combination in Figure 1.3 appeared $3!4!$ times, now each combination will occur $n_1!n_2! \dots n_M!$ times. So, the total number of *permutations*, $N!$, represents an over-counting by that factor. Hence, the required total number of combinations must be

$$\frac{N!}{n_1!n_2! \dots n_M!}. \quad (1.12)$$

If the chance of a particular particle being allocated to bin i is p_i , then what is the chance $P(n_1, n_2, \dots, n_M)$ of finding n_i particles with no regard for order in bin i , for a given set of $n_1, n_2, \dots, n_M$? [Note that for consistency with the $M = 2$ case in (1.11), we should perhaps exclude n_M from $P(n_1, n_2, \dots, n_M)$, since $n_M = N - n_1 - \dots$ is not an independent variable. But the precise notation here is not so important, as long as we know what we are calculating.] Each relevant combination occurs with probability $p_1^{n_1} p_2^{n_2} \dots p_M^{n_M}$, and so

$$P(n_1, n_2, \dots, n_M) = \frac{N!}{n_1! n_2! \dots n_M!} p_1^{n_1} p_2^{n_2} \dots p_M^{n_M}. \quad (1.13)$$

This is the multinomial distribution.

Here is an example of using the binomial distribution to count the number of ways in which a set of distinguishable particles can arrange themselves.

6 and 60 Particles in a Room

Suppose 6 particles are moving randomly in a room. What is the probability p that any 2 of them are in the front third of the room at some given moment?

$$\begin{array}{cc} \underbrace{\text{2 particles}}_{\text{front } 1/3} & \underbrace{\text{4 particles}}_{\text{back } 2/3} \\ p = \frac{6!}{2! 4!} (1/3)^2 (2/3)^4 \simeq 0.33. \end{array} \quad (1.14)$$

Now multiply the numbers by 10: with 60 particles moving randomly, what is the probability p that any 20 of them are in the front third of the room at some given moment?

$$\begin{array}{cc} \underbrace{\text{20 particles}}_{\text{front } 1/3} & \underbrace{\text{40 particles}}_{\text{back } 2/3} \\ p = \frac{60!}{20! 40!} (1/3)^{20} (2/3)^{40} \simeq 0.11. \end{array} \quad (1.15)$$

In the absence of any other information, we can only estimate that one third of the total number of particles in the above two examples should occupy the front third of the room at any given moment. But that doesn't mean that the chance of *exactly* one third of the total number of particles occupying the front third of the room should be large. As we increase the number of particles in the room, there are simply more possibilities for the ways in which the particles can arrange themselves, and so the chance of exactly one third of the total number occupying the front third of the room decreases.

If the numbers of particles being treated in such examples are made much larger, the factorials that appear in the binomial distribution soon become unwieldy, and we need to find a tractable expression for the factorial. [This will turn out to be the one that was used without justification in (1.3).] Here is a Riemann-sum method for approximating the factorial that is also applicable more broadly to other functions, and so is worth a pause to examine in detail. Figure 1.4 shows a plot of the natural logarithm $y = \ln x$. The area under the curve from $x = 1$ to n is

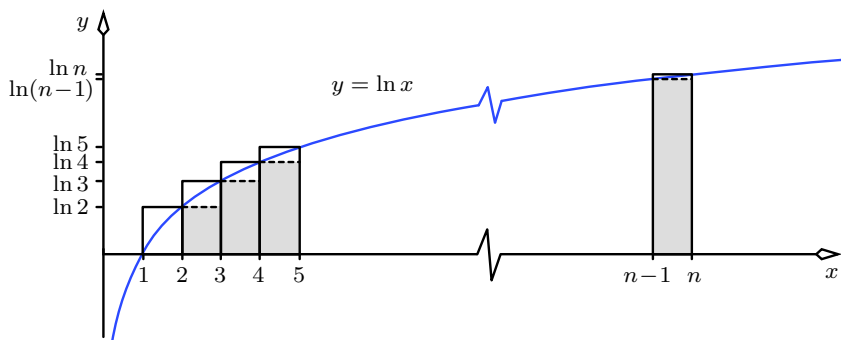


Fig. 1.4 The natural logarithm and associated Riemann sums used to approximate Stirling's rule. The lower-bound Riemann sum is the summed areas of the shaded rectangles, and the upper-bound is the summed areas of the taller rectangles

$$\int_1^n \ln x \, dx = [x \ln x - x]_1^n = n \ln n - n + 1. \quad (1.16)$$

This area is bounded above and below by *Riemann sums*, the summed areas of the vertical strips that terminate all above and all below the curve, as shown in the figure. These bound the area (1.16) in the following way:

$$\underbrace{\ln 2 + \ln 3 + \cdots + \ln(n-1)}_{\text{total shaded area}} < n \ln n - n + 1 < \underbrace{\ln 2 + \ln 3 + \cdots + \ln n}_{\text{total unshaded area}}. \quad (1.17)$$

That is,

$$\ln n! - \ln n < n \ln n - n + 1 < \ln n!. \quad (1.18)$$

Referring to Figure 1.4, we see that a good approximation of the area under the curve, $n \ln n - n + 1$, will be the mean of the upper and lower sums in (1.18):

$$n \ln n - n + 1 \simeq \frac{2 \ln n! - \ln n}{2} = \ln n! - \frac{1}{2} \ln n. \quad (1.19)$$

It follows that

$$\ln n! \simeq (n + \frac{1}{2}) \ln n - n + 1. \quad (1.20)$$

This is one form of *Stirling's rule* for approximating factorials. For any “reasonably large” x not necessarily a whole number, the rule can be stated to higher accuracy as the infinite series

$$\ln x! \sim (x + \frac{1}{2}) \ln x - x + \ln \sqrt{2\pi} + \frac{1}{12x} - \frac{1}{360x^3} + \cdots, \quad (1.21)$$

where

$$“f(x) \sim g(x)” \text{ denotes } \lim_{x \rightarrow \infty} f(x)/g(x) = 1. \quad (1.22)$$

The limit in (1.22) does not imply that $f(x) - g(x) \rightarrow 0$ as $x \rightarrow \infty$; in fact, the difference between the left- and right-hand sides of (1.21) grows without bound as $x \rightarrow \infty$. But the ratio of those sides tends toward 1. Equation (1.21) is often truncated to

$$\ln x! \sim (x + 1/2) \ln x - x + \ln \sqrt{2\pi}. \quad (1.23)$$

This matches (1.20) very well: $\ln \sqrt{2\pi} \simeq 0.92$.

Equation (1.21) is an *asymptotic series*. Such series behave differently from the *convergent* series that are more usually encountered in physics. To see how, consider the well-known convergent series for the exponential function:

$$e^x = 1 + x + x^2/2! + x^3/3! + \dots \quad (1.24)$$

In a convergent series, we fix x and observe convergence of the partial sums as the number of summed terms goes to infinity. For a given x , we can calculate e^x to any accuracy by summing a sufficient number of terms in (1.24); the more terms we sum, the better the approximation to e^x . But in contrast, an asymptotic series such as (1.21) does not converge in this way for any value of x at all. The coefficients of the first few powers of x in (1.21) start out decreasing term by term, but that trend soon reverses as they begin to grow without bound. For any choice of x , those coefficients eventually grow larger at a faster rate than can ever be suppressed by the powers of x in the denominator, and so the series can never converge by our simply summing more terms. Instead, we implement (1.21) by truncating its right-hand side wherever we like, and then we note that summing this finite series produces an increasingly better approximation to $\ln x!$ as x increases. This means we cannot use (1.21) to calculate $\ln x!$ to arbitrary accuracy for any particular x . Precisely where the truncation might best be made to maximise the accuracy of the approximation for minimal computational effort is something of a black art. To summarise:

- In a **convergent series** such as the exponential series (1.24), we fix the value of x and determine the left-hand side to any accuracy by increasing the number of terms summed on the right-hand side.
- In an **asymptotic series** such as the factorial (1.21), we fix the number of terms summed on the right-hand side, and can only be “confident” that their sum is a good approximation of the left-hand side when x is large. A whole field of mathematics exists that investigates the bounds that might be placed on the results of such calculations.

Stirling’s rule is sometimes written by exponentiating both sides of (1.23):

$$x! \sim x^{x+1/2} e^{-x} \sqrt{2\pi}. \quad (1.25)$$

Despite not being exact, (1.25) is the most common expression used to calculate large factorials. We apply it to the following example, similar to those in (1.14) and (1.15).

6000 Particles in a Room

Suppose 6000 particles are moving randomly in a room. What is the probability p that any 2000 of them are in the front third at some given moment?

$$p = \frac{\overbrace{2000 \text{ particles}}^{\text{front } 1/3}}{2000!} \frac{\overbrace{4000 \text{ particles}}^{\text{back } 2/3}}{4000!} (1/3)^{2000} (2/3)^{4000}. \quad (1.26)$$

Evaluate p by applying Stirling's rule (1.23):

$$\begin{aligned} \ln p &\simeq 6000.5 \ln 6000 - \cancel{6000} + \ln \sqrt{2\pi} \\ &\quad - 2000.5 \ln 2000 + \cancel{2000} - \ln \sqrt{2\pi} \\ &\quad - 4000.5 \ln 4000 + \cancel{4000} - \ln \sqrt{2\pi} \\ &\quad + 2000 \ln 1/3 + 4000 \ln 2/3 \\ &\simeq -4.52. \end{aligned}$$

This results in $p = e^{-4.52} \simeq 0.011$. Stirling's rule has been very accurate here: calculating (1.26) on a computer returns the same result for the probability, to six decimal places.

Stirling's rule (1.21) is often written as

$$\ln x! \simeq x \ln x - x, \quad \text{or} \quad x! \simeq x^x e^{-x}. \quad (1.27)$$

For smaller values of x , this function is actually hopelessly inadequate. For example, consider the value of $50!$ formed from the product $50 \times 49 \times 48 \times \dots$:

$$50! \simeq 3.041 \times 10^{64}. \quad (1.28)$$

Compare this value with the accurate form of Stirling's rule, (1.25), and the rough-and-ready form, (1.27):

$$\text{accurate Stirling: } 50! \approx 50^{50.5} e^{-50} \sqrt{2\pi} \simeq 3.036 \times 10^{64}, \quad (1.29)$$

$$\text{rough-and-ready Stirling: } 50! \approx 50^{50} e^{-50} \simeq 0.171 \times 10^{64}. \quad (1.30)$$

The accurate form clearly wins out here over the rough-and-ready form.

The rough-and-ready form (1.27) will set the logarithm of the probability to zero in the example of (1.26), and so it is clearly not accurate there—and

this rough-and-ready form becomes more and more inaccurate as $x \rightarrow \infty$. Nevertheless, compare the accurate form with the rough-and-ready form when $x \approx 10^{24}$, which is a value that we'll later have occasion to use:

$$\begin{aligned} \text{accurate Stirling: } \ln(10^{24}!) &\simeq (10^{24} + 1/2) \ln 10^{24} - 10^{24} + \ln \sqrt{2\pi} \\ &\simeq 5.4262 \times 10^{25}, \\ \text{rough-and-ready Stirling: } \ln(10^{24}!) &\simeq 10^{24} \ln 10^{24} - 10^{24} \\ &\simeq 5.4262 \times 10^{25}. \end{aligned} \tag{1.31}$$

These two estimates are not really the same. The accurate logarithm is greater than the rough-and-ready logarithm by the amount

$$1/2 \ln 10^{24} + \ln \sqrt{2\pi} \simeq 28.55. \tag{1.32}$$

That is, the accurate estimate is greater than the rough-and-ready estimate by a *factor* of $e^{28.55} \simeq 2.5 \times 10^{12}$. It would, of course, normally be unacceptable to regard two such estimates as more or less equal when the first is over two million million times larger than the second! But when working with such enormous numbers as $10^{24}!$, we are usually content with a “very approximate” numerical value of the factorial—even if it *is* wildly incorrect when judged by everyday standards of approximation.

The Factorial in Number Theory

The factorial is one of the most frequently encountered functions in the field of pure mathematics known as number theory. The function is well defined for the natural numbers $1, 2, 3, \dots$, of course, but extending its definition to the real numbers requires making a choice of how its most general definition should behave. The most common approach demands that the function obey “ $n! = n(n-1)!$ ” for as many numbers n as possible. This extends its definition to equate $0!$ with 1, since it makes perfect sense to write $1! = 1 \times 0!$. But we cannot go any further to apply $n! = n(n-1)!$ to negative integers, because writing “ $0! = 0 \times (-1)!$ ” is clearly nonsense. It follows that the factorial of any negative integer is not defined. Other than that, mathematicians are at liberty to define the factorial of any other number in whatever way they please. Several definitions exist that produce “ $n! = n(n-1)!$ ” for whole numbers, but the definition most often used in number theory begins by defining the factorial of any positive *real* number in terms of the real integral

$$x! \equiv \int_0^\infty u^x e^{-u} du. \tag{1.33}$$

This returns the usual values for the whole numbers, as can be seen with an integration by parts. It has asymptotes at the negative integers.

The field of complex analysis takes the definition (1.33) and builds on it to develop a process for defining and calculating the factorial of any *complex* number z , using the theory of *analytic continuation*. The calculations involved are very refined, and involve a fundamentally higher order of difficulty than evaluating expressions such as $\sqrt{2}$ or $\sin 1$, as is evident from the asymptotic nature of Stirling's series (1.21). The central players here are *analytic* functions, which are a particular type of well-behaved function to which complex analysis devotes itself almost exclusively. The core theorem of analytic continuation states the following: given an analytic function $f_1(z)$ defined on some part of the complex plane, if another analytic function $f_2(z)$ can be found that is identical to $f_1(z)$ in some subset of the plane, then $f_2(z)$ is unique in being identical in this way. Thus, if some analytic function can be found (essentially by trial and error) that agrees with $x!$ in some subset of the complex plane, analytic continuation theory guarantees that this function will be unique. Hence, it will be a natural choice of an extension of $x!$ to the complex plane.

Stirling's more accurate rule (1.25) can be used to compute approximate values of factorials of positive real numbers. Its use can be extended to factorials of negative real numbers in the following way. Consider computing $(-4.3)!$. Start with

$$1.7! = 1.7 \times 0.7 \times -0.3 \times -1.3 \times -2.3 \times -3.3 \times (-4.3)!. \quad (1.34)$$

Rearrange this, to obtain

$$\begin{aligned} (-4.3)! &= \frac{1.7!}{1.7 \times 0.7 \times -0.3 \times -1.3 \times -2.3 \times -3.3} \\ &\simeq \frac{1.7^{1.7+1/2} e^{-1.7} \sqrt{2\pi}}{1.7 \times 0.7 \times -0.3 \times -1.3 \times -2.3 \times -3.3} \simeq 0.418. \end{aligned} \quad (1.35)$$

A more accurate value is 0.439. Two simple approximations of $x!$ applicable to x between 1 and 2 are

$$(1 + \delta)! \simeq 0.96 \times 2^\delta \approx 1 + \delta, \quad 0 \leq \delta \leq 1. \quad (1.36)$$

An alternative notation for the more general factorial (1.33) is $\Pi(z) \equiv z!$, which is convenient when working with its derivative, $\Pi'(z)$. You will usually see the equivalent *gamma function* notation $\Gamma(z+1) \equiv z!$ in textbooks. No reason seems to be known—historical or mathematical—for the strange shift of “+1” in its definition. The Π notation tends to make for simpler expressions. If you use the Γ nota-

tion, try to see if you needn't always remind yourself of that “+ 1” shift when calculating something as simple as $\Gamma(5)$. The “+ 1” has long been an irritation for some mathematicians—while others seem to revel in its obscurity!

The Γ notation was used by Bernhard Riemann, who carried out some of the most famous work in number theory, in the mid-nineteenth century. Riemann's *zeta function* leans heavily on the factorial. His *Riemann hypothesis* is a statement concerning the zeroes of the zeta function, and if you can prove or disprove it, you are assured of mathematical immortality. It is an open question as to whether the most commonly used definition of the factorial, (1.33), really does support analyses of the zeta function, or whether the asymptotes at the negative integers resulting from the definition (1.33) only get in the way of everything. Alternative definitions of the factorial don't have asymptotes, and perhaps one of them will one day prove to be a better definition by shedding light on the zeta function and proving Riemann's hypothesis.

1.3 Fluctuations in the Binomial Distribution

To return to our analyses of the binomial distribution, consider the results of (1.14), (1.15), and (1.26). The chance of *exactly* one third of the particles being found in the front third of the room goes to zero as the number of particles is increased,² and indeed, the chance of any other number of particles appearing in the front third is smaller still. Consider a more general room of N particles, with a “sub-room”, a partition, which particles can travel to with probability p , and whose number of particles n we observe at each instant. Given these binomial parameters N and p , we might ask for the most likely value of n : the value that maximises the binomial probability $P(n)$ in (1.11). But another question will prove useful here: we ask for the *mean* number of particles that we would observe to be in the sub-room if we could take measurements over a long period of time. We also ask for the *width* of the probability distribution $P(n)$, meaning the width of the main “bulk” of the probability distribution. This width quantifies how the number of particles in the sub-room is expected to fluctuate around its mean as time passes. A commonly used measure of this width is the *standard deviation* of the probability distribution, which we discuss next.

² Of course, we are not really concerned as to whether the total number of particles in the room is a multiple of 3.

1.3.1 Expected Value and Standard Deviation of a Random Variable

Recall the ergodic assumption described at the start of Section 1.1. This says that the following two numbers will be the same:

1. We place N gas particles into a room, and then measure the number of those particles n that are found in a sub-room at, say, irregular intervals over a long period of time, averaging those measurements.
2. Imagine a large number of identical rooms (each with a sub-room defined), each of which has N particles placed into it in some random way. At a given instant, we measure the number of particles n found in each of the sub-rooms, and average the results.

These two mean values of the number of particles n in the sub-room are hypothesised to be equal, in the limit of long times and a large number of identical rooms. This common value is called $\langle n \rangle$, the “*expected value* of n ”. This name is standard, but something of a misnomer. Consider that an unbiased die presents each face numbering 1 to 6 with equal probability, and the mean of this face number for a number of throws will tend toward 3.5 as that number of throws tends to infinity. And yet we certainly don’t ever *expect* to throw 3.5 on a die. The “expected value” is also not the “most likely value”, because 3.5 is not the most likely value that will appear—it can never appear. The most likely value of a more typical probability distribution is the value at which that probability peaks; and yet the mean rarely coincides with this peak.

But whereas 3.5 is rarely the mean of a given set of throws of a die, and it is *never* expected to appear, it is certainly the *most likely value of the mean of a given set of throws*. If you were given a set of outcomes of the throws of an unbiased die and had to bet in advance what its mean would be (say, to within 0.1), you would be wise to bet on 3.5. So, the “expected value of n ” means “the most likely value of the mean of a set of randomly produced values of n ”. The expected value is also often called the *expectation value*. This is perhaps a better term, because it has no implied meaning based on an everyday use of the word “expectation”, since that word seldom appears in everyday speech. But provided you are aware of the technical meaning, the shorter term “expected” should present no difficulties.

The expected value of n is also often written as $\bar{n}$; in fact, $\bar{n}$ is really the mean of a *given* set of values of n . To see the difference between $\langle n \rangle$ and $\bar{n}$, realise that the expected value $\langle n \rangle$ exists even without a measurement being made, whereas the mean $\bar{n}$ is the average of a set of measurements. Nonetheless, the two notations and names tend to be used interchangeably.

More generally, suppose that a random variable x can take on any of a set of values x_i , with the probability of x_i being p_i . In a great number of

measurements of x , what is $\langle x \rangle$, the most likely value of the mean of those measurements? If a large number N of measurements are taken, the most likely frequency of occurrences of x_i will be about Np_i times. The expected value $\langle x \rangle$ will be the mean of this large number of measurements:

$$\begin{aligned}\langle x \rangle &= \bar{x} \equiv \frac{\sum x_i}{N} = \frac{x_1 + x_1 + \cdots + x_2 + x_2 + x_2 + \cdots}{N} \\ &= \frac{Np_1x_1 + Np_2x_2 + \cdots}{N} = p_1x_1 + p_2x_2 + \cdots\end{aligned}\quad (1.37)$$

The expected value of x is thus

$$\langle x \rangle = \sum_i p_i x_i. \quad (1.38)$$

We now seek a useful measure of the width of a probability distribution. This is essentially how far the x_i typically stray from their expected value (which we'll write for simplicity as their mean $\bar{x}$). We might consider using the mean of the distances between the x_i and $\bar{x}$, known as the *absolute deviation* of x :

$$\text{absolute deviation of } x \equiv \langle |x - \bar{x}| \rangle. \quad (1.39)$$

But it turns out (we'll see why shortly) that a more meaningful definition of the width is the *rms deviation* of x_i from $\bar{x}$, where “rms deviation” means the square root of the mean of the squares of the deviations of x from $\bar{x}$. This rms deviation is more usually called the *standard deviation* of x , written σ_x . For algebraic simplicity, it's more usual to work with σ_x^2 , called the *variance* of x :

$$\sigma_x^2 \equiv \langle (x - \bar{x})^2 \rangle. \quad (1.40)$$

What Makes the Standard Deviation Special?

Here is one explanation for why the width of the probability distribution of x is more meaningfully defined as the standard deviation σ_x from (1.40) rather than the absolute deviation (1.39). Above, we followed the standard route of defining a mean $\bar{x}$, and then used it to define a width, either from (1.39) or (1.40)—or indeed, from any of an infinite number of other choices that might, say, involve higher powers of $x - \bar{x}$. But consider an altogether different approach to this subject. We will reverse the order of presentation and define the measure of width first, and then use it to define a mean $\bar{x}$ —we will assume no prior definition of the mean of x . In particular, suppose we begin with a width σ_x defined by (1.40). This width depends on an unknown quantity called $\bar{x}$: it measures the typical distance that x is away from $\bar{x}$. Now imagine altering the value of

$\bar{x}$ in such a way that σ_x is minimised: this will define the distribution's width in a kind of minimal way. So, we are required to solve $d\sigma_x/d\bar{x} = 0$. It is equivalent, but more convenient, to solve $d(\sigma_x^2)/d\bar{x} = 0$ instead. If our expected values are means taken over some large number N , then

$$\frac{d(\sigma_x^2)}{d\bar{x}} \stackrel{(1.40)}{=} \frac{d}{d\bar{x}} \sum_i \frac{(x_i - \bar{x})^2}{N} = \frac{1}{N} \sum_i 2(x_i - \bar{x}) \stackrel{\text{req}}{=} 0, \quad (1.41)$$

where “ $\stackrel{\text{req}}{=}$ ” is read as “which is required to equal”. It follows from (1.41) that

$$\sum_i (x_i - \bar{x}) = 0, \text{ or } \sum_i x_i - N\bar{x} = 0, \text{ so } \bar{x} = \frac{\sum_i x_i}{N}. \quad (1.42)$$

But this is precisely the expression for $\bar{x}$ that we are familiar with. In other words, the variance (or equivalently, the standard deviation) is special because minimising it with respect to a kind of “centre of mass” of the probability distribution called “ $\bar{x}$ ” produces an expression for $\bar{x}$ that accords with our everyday intuition of what that centre of mass should be: a sum of possible values divided by the total number of those values. For the absolute deviation in (1.39), this minimisation procedure sets $\bar{x}$ to be the *median* of the values of x , meaning their midpoint when laid out from minimum to maximum. The concept of a median is useful as a quantifier of average house prices, because it's insensitive to outliers—after all, the castle on the hill has no relevance to the average house buyer. But generally, it's the arithmetic mean arising from (1.42) that we prefer to deal with in physics, and so the standard deviation becomes the pre-eminent measure of a distribution's width.

It's useful analytically to put the somewhat convoluted-looking expression (1.40) into a different form. An expected value “ $\langle \cdot \rangle$ ” involves a sum, but rather than expand (1.40) into its constituent sums, we appeal instead to the *linearity* of the summation $\sum$. Linear operators are the most important type of operator in physics. An operator L is linear if, for constants a and b ,

$$L(ax + by) = aL(x) + bL(y). \quad (1.43)$$

It suffices to have just two terms on the right-hand side of (1.43); but it's easy to show [by applying (1.43) recursively] that if L is linear, it can be applied in the same way as (1.43) to any number of terms:

$$L(ax + by + cz + \dots) = aL(x) + bL(y) + cL(z) + \dots \quad (1.44)$$

In the case of summation, for constants a and b ,

$$\sum_i ax_i + by_i = a \sum_i x_i + b \sum_i y_i, \quad (1.45)$$

and we conclude that summation is a linear operation. Linearity is the primary property of $\sum$, one that makes short work of showing that the operation of calculating the expected value $\langle \cdot \rangle$ is also linear. We do that in the following way. When x and y are random variables that are each sampled N times, then

$$\langle ax + by \rangle \equiv \frac{1}{N} \sum_i ax_i + by_i = \frac{1}{N} \left(a \sum_i x_i + b \sum_i y_i \right) = a \langle x \rangle + b \langle y \rangle; \quad (1.46)$$

and this result “ $\langle ax + by \rangle = a \langle x \rangle + b \langle y \rangle$ ” is the very definition of linearity. We can now use this linearity of the expected value to simplify (1.40). Write³

$$\begin{aligned} \sigma_x^2 &\equiv \langle (x - \bar{x})^2 \rangle = \langle x^2 - 2x\bar{x} + \bar{x}^2 \rangle = \langle x^2 \rangle - 2\bar{x} \langle x \rangle + \langle \bar{x}^2 \rangle \\ &= \langle x^2 \rangle - 2\bar{x}^2 + \bar{x}^2 \\ &= \langle x^2 \rangle - \bar{x}^2, \text{ alternatively written } \langle x^2 \rangle - \langle x \rangle^2. \end{aligned} \quad (1.47)$$

The variance thus equals “the mean of the square minus the square of the mean”.

Let’s now use these ideas to calculate the expected value and standard deviation of the random variable n that is attached to probability p in the binomial distribution. We were imagining a room of N particles with a sub-room in which particles could appear with probability p , and whose number of particles at any moment is some variable n . We require $\bar{n}$ and σ_n .

It’s reasonable to suppose that the mean number of particles will be $\bar{n} = pN$; for example, if $p = 1/3$ and $N = 300$, then an average number of $\bar{n} = 1/3 \times 300$ particles will be in the sub-room. Here is a different way of deriving that result, which will also soon lend itself to determining σ_n . From first principles (and writing $q \equiv 1 - p$ for convenience),

$$\bar{n} \equiv \sum_{n=0}^N n P(n) \stackrel{(1.11)}{=} \sum_{n=0}^N n C_n^N p^n q^{N-n}. \quad (1.48)$$

This sum looks difficult to evaluate. But we can do it by treating q initially as an independent variable, finding the sum, and then setting $q = 1 - p$ only at the end of the calculation. This is possible to do because the sum (1.48) is completely well defined for *any* p and *any* q ; these two variables need not be related at all. After finding the sum (which will be a function of p and q), we are then free to set q to be whatever we like—such as $1 - p$.

The following procedure evaluates this general sum. Begin with a small space-saving notation: $\partial_p \equiv \partial/\partial p$, and make use of the following two expres-

³ The appearance of both $\bar{x}$ and $\langle x \rangle$ in (1.47) is deliberate: we wish to show how the angle brackets give a linear operation that happens to include the number $\bar{x}$. In the second line of that equation, we recognise that $\langle x \rangle = \bar{x}$.

sions, which can both be proved by induction:

$$(p \partial_p)^k p^n = n^k p^n, \quad (1.49)$$

and

$$(p + q)^N = \sum_{n=0}^N C_n^N p^n q^{N-n}. \quad (1.50)$$

The last expression is the *binomial theorem*. Use (1.49) with $k = 1$ in (1.48):

$$\begin{aligned} \bar{n} &= \sum_{n=0}^N C_n^N n p^n q^{N-n} = \sum_n C_n^N p \partial_p p^n q^{N-n} \\ &= p \partial_p \sum_n C_n^N p^n q^{N-n} \\ &\stackrel{(1.50)}{=} p \partial_p (p + q)^N = pN(p + q)^{N-1} = pN, \end{aligned} \quad (1.51)$$

where we replaced q with $1 - p$ in the last step. This result, “ $\bar{n} = pN$ ”, is just what we intuitively expected a few paragraphs up.

Another application of the above procedure allows us easily to compute the width $\sigma_n^2 = \langle n^2 \rangle - \bar{n}^2$ of the binomial distribution. Begin with

$$\langle n^2 \rangle = \sum_n n^2 P(n) = \sum_n C_n^N n^2 p^n q^{N-n}. \quad (1.52)$$

Now use (1.49) with $k = 2$ to rewrite the last expression in (1.52):

$$\begin{aligned} \langle n^2 \rangle &= (p \partial_p)^2 \sum_n C_n^N p^n q^{N-n} \stackrel{(1.50)}{=} (p \partial_p)^2 (p + q)^N \\ &= p \partial_p [pN(p + q)^{N-1}] = pN[(p + q)^{N-1} + p(N-1)(p + q)^{N-2}] \\ &= pN[1 + p(N-1)] = pN[pN + q] = p^2 N^2 + Npq. \end{aligned} \quad (1.53)$$

Finally, from (1.47), write

$$\begin{aligned} \sigma_n^2 &= \langle n^2 \rangle - \bar{n}^2 \stackrel{(1.53)}{=} p^2 N^2 + Npq - (pN)^2 \\ &= Np(1 - p), \end{aligned} \quad (1.54)$$

where we have replaced q with $1 - p$. This result, $\sigma_n^2 = Np(1 - p)$, is well known as the variance of the binomial distribution.

The standard deviation of a distribution is a measure of its width, and thus is a measure of the fluctuations that can be expected around the distribution's mean value. Define the *relative fluctuation* in n as being

$$\text{relative fluctuation} \equiv \frac{\sigma_n}{\bar{n}} = \frac{\sqrt{Np(1-p)}}{Np} = \sqrt{\frac{1-p}{Np}}. \quad (1.55)$$

Note that the relative fluctuation is inversely proportional to $\sqrt{N}$, where N is the total number of elements in the system. This is an important indicator of randomness in a system, and it appears frequently throughout statistical mechanics.

In (1.14), we examined 6 particles moving randomly in a room, and saw that the chance of finding the expected number 2 in the front third of the room was small. What is the relative fluctuation in this expected number of 2 particles?

$$\sigma_n/\bar{n} = \sqrt{\frac{1-p}{Np}} = \sqrt{\frac{2/3}{N \times 1/3}} = \sqrt{2/N} = \sqrt{2/6} \simeq 0.58. \quad (1.56)$$

The expected number of 2 particles thus fluctuates typically by around 58% (or about 1 particle). Now find the relative fluctuation for the 6000 particles treated in (1.26):

$$\sigma_n/\bar{n} = \sqrt{2/6000} \simeq 0.018. \quad (1.57)$$

The expected number of 2000 particles fluctuates by only 1.8%. This relative fluctuation translates to an absolute fluctuation of $\sigma_n = 36$ particles, but it's the relative fluctuation that interests us. Finally, do the same for a more realistic number of 10^{27} particles in a room:

$$\sigma_n/\bar{n} = \sqrt{2/10^{27}} \simeq 4.5 \times 10^{-14}. \quad (1.58)$$

This relative fluctuation is minuscule. We see that a system with a stupendously large number of elements has *very* small relative fluctuations about its mean, and this makes it *very* predictable. This predictability of a system that is really random is the bedrock into which statistical mechanics is anchored.

1.3.2 The Random Walk

The $\sqrt{N}$ appearing in (1.55) is the classic signature of the *random walk*. If a drunken man stumbles away from a start point, taking steps of uniform length L in arbitrary directions, then on the average, how far from his start point will he be after n steps? To construct a set of distances to be averaged, we could consider one drunk who repeatedly walks away from his start point, only to be “reset” back to it after n steps, after which the process begins anew (without him sobering up). Or we could envisage a whole collection, an *ensemble*, of drunken men all walking away from their own start points, and calculate the average of all of their distances after n steps. It's reasonable to assume that the two averages will be the same, a concept we'll later explore further in Figure 5.1.

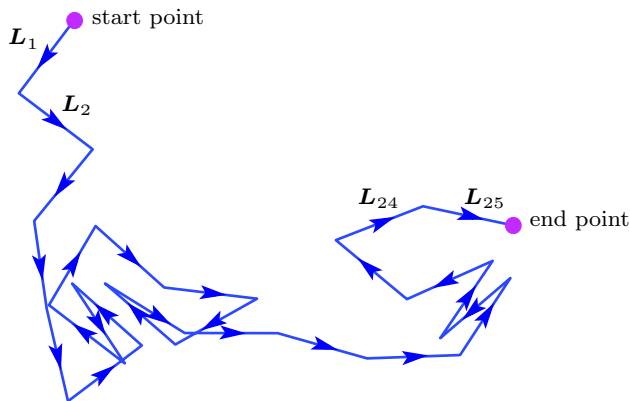


Fig. 1.5 A random walk of 25 steps, each of length L . Over an ensemble of these walks, the root-mean-square distance from start to end positions will be $5L$. The actual distance in the single trial pictured is about $5.2L$

With the start point taken as the origin, let the i^{th} step be a vector $\mathbf{L}_i$. After n steps, the man's position relative to the start is

$$\text{position} = \mathbf{L}_1 + \cdots + \mathbf{L}_n, \quad (1.59)$$

as shown in Figure 1.5. Consider that the square of his final distance from the start point is

$$\begin{aligned} \text{distance}^2 \equiv |\text{position}|^2 &= (\mathbf{L}_1 + \cdots + \mathbf{L}_n) \cdot (\mathbf{L}_1 + \cdots + \mathbf{L}_n) \\ &= nL^2 + \mathbf{L}_1 \cdot \mathbf{L}_2 + \mathbf{L}_1 \cdot \mathbf{L}_3 + \cdots \end{aligned} \quad (1.60)$$

In averaging over the ensemble, the cross terms $\mathbf{L}_i \cdot \mathbf{L}_j$ ($i \neq j$) are just as likely to be positive as negative, since there is no correlation from one step to the next. The cross terms thus make no contribution to the average, and the mean of the squared distances becomes

$$\langle \text{distance}^2 \rangle = nL^2. \quad (1.61)$$

The rms distance from start to end is then

$$\text{rms distance} \equiv \sqrt{\langle \text{distance}^2 \rangle} = \sqrt{n} L. \quad (1.62)$$

We see that the rms distance from the start point increases as the *square root* of the number of steps. Calculating the true mean distance is a good deal more complicated, and that's why the rms value is usually used.

The random walk is featured in the constructive interference of light, and so it figures in the explanation of why laser light is so bright. Suppose we group together n *incoherent* light sources, such as candles or incandescent bulbs, to

combine their light into one bright source. By “incoherent”, we mean that like the steps of the drunken man, the phase of the light waves varies randomly over very short times and distances: this is the case for candles or incandescent bulbs that use a hot filament to generate their light. At any particular place and time, the strength of the electric field due to light from the i^{th} candle or bulb can be represented by a phasor (a vector $\mathbf{E}_i$ of magnitude E that we’ll take as fixed, for simplicity). Light’s intensity is proportional to the square of the amplitude of the total electric field. This amplitude is the length of the sum of all the source phasors. It follows that the total intensity at the point of interest is

$$I_{\text{tot}} \propto \left| \sum_i \mathbf{E}_i \right|^2. \quad (1.63)$$

Because the light bulbs generate uncorrelated phasors $\mathbf{E}_i$ that are added as vectors—just like the uncorrelated steps $\mathbf{L}_i$ of the drunken man—we are back to a random walk, with $\mathbf{E}_i$ in place of $\mathbf{L}_i$. Over the integration time of the human eye, the total intensity of the n incoherent light sources becomes averaged just like an ensemble of random walks; so, its average I_{tot} becomes proportional to nE^2 , or n times the average intensity due to a single candle or light bulb. This accords with our everyday experience of collecting light sources into a group, and it shows why the concepts of rms value and the random walk are so closely allied with our physical perceptions.

But if the light sources are *coherent*—which is to say, that the “steps” are correlated from one to the next, such as in a laser (the drunk sobers up!)—then the random-walk picture no longer applies. If n lasers are carefully tuned to be in phase with one another, their electric field phasors at a screen will add constructively, so that

$$I_{\text{tot}} \propto \left| \sum_i \mathbf{E}_i \right|^2 = |n\mathbf{E}_1|^2 = n^2 E^2. \quad (1.64)$$

The average I_{tot} is now n^2 times the average intensity of a single laser, in contrast to the n -fold multiplication of intensity for candles or incandescent bulbs. And even in a single laser, the light sources are really de-exciting atoms that emit their light coherently with each other; so that here again, the sum of squares shows why the coherence of a single laser makes its central spot so very bright.

1.4 Gaussian Approximation of the Binomial Distribution

The above examples of the binomial distribution were simple enough not to tax our calculational abilities too much. But when the systems we exam-

ine involve real-world numbers, the binomial distribution is generally difficult to work with: we quickly find that evaluating the binomial distribution $P(n) = C_n^N p^n (1-p)^{N-n}$ via Stirling's rule is tedious, and the procedure doesn't give any real insight into the shape of $P(n)$. We are interested in examining fluctuations around the mean $\bar{n}$, and it turns out that in the vicinity of the mean, the binomial distribution can be approximated by the normal distribution. The latter is a continuous function, and is thus far more amenable to the tools of calculus. Here we'll examine how this approximation is made.

Our first idea might be to approximate $P(n)$ by a Taylor series expansion about $\bar{n}$, in powers of $n - \bar{n}$. Fitting a Taylor series through a set of points is equivalent to fitting a polynomial exactly through the points: the polynomial is a power series in $n - \bar{n}$. But the binomial distribution can be very peaked, and this sharp peak forces the fitting Taylor polynomial to bend excessively, requiring many terms of the Taylor series to provide a good fit. We can avoid this need to specify several terms of the Taylor series by approximating not $P(n)$ but the *logarithm* of $P(n)$ with a Taylor series, because the logarithm will be far less peaked than the binomial distribution itself.⁴ So, consider the function $f(n) \equiv \ln P(n)$, which we will take to be *continuous*: that is, we'll allow n to assume any real value, because continuous functions are generally far easier to analyse than are discrete ones. We now expand $f(n)$ around $\bar{n}$ in a Taylor series:

$$f(n) = f(\bar{n}) + f'(\bar{n})(n - \bar{n}) + \frac{1}{2!} f''(\bar{n})(n - \bar{n})^2 + \dots \quad (1.65)$$

We'll fit a parabola to the gentler peak of the logarithm, meaning we will truncate the above Taylor series after its term involving $(n - \bar{n})^2$. Thus, we require $f(\bar{n})$ and the derivatives $f'(\bar{n})$ and $f''(\bar{n})$. Recall that $f(n) = \ln P(n)$, so begin with the binomial distribution (1.11), which we now write as

$$P(n) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n}, \quad (1.66)$$

and take its logarithm. Eliminating the factorials using Stirling's rule (1.23) gives

$$\begin{aligned} f(n) &\simeq \ln N! - \ln n! - \ln(N-n)! + n \ln p + (N-n) \ln(1-p) \\ &\simeq (N + 1/2) \ln N - (n + 1/2) \ln n - (N - n + 1/2) \ln(N-n) \\ &\quad - \ln \sqrt{2\pi} + n \ln p + (N-n) \ln(1-p). \end{aligned} \quad (1.67)$$

Differentiating this with respect to n gives

⁴ For example, compare the extreme peak in the sequence “1, 100, 10⁶, 100, 1” with the much softer peak of its base-10 logarithms, “0, 2, 6, 2, 0”.

$$\begin{aligned}
f'(n) &\simeq -\ln n - \frac{1}{2n} + \ln(N-n) + \frac{1}{2(N-n)} + \ln p - \ln(1-p), \\
f''(n) &\simeq \frac{-1}{n} + \frac{1}{2n^2} - \frac{1}{N-n} + \frac{1}{2(N-n)^2}.
\end{aligned} \tag{1.68}$$

In the statistically interesting cases of p not close to 0 or 1, evaluating (1.67) and (1.68) at $n = \bar{n}$ gives the following, where we drop the subscript n on the variance σ_n^2 :

$$f(\bar{n}) \simeq -\ln \sqrt{2\pi\sigma^2}, \quad f'(\bar{n}) \simeq \frac{2p-1}{2\sigma^2}, \quad f''(\bar{n}) \simeq \frac{-1}{\sigma^2}. \tag{1.69}$$

Now place these three items into the Taylor series (1.65). You will find that

$$f(n) \simeq -\ln \sqrt{2\pi\sigma^2} - \frac{1}{2\sigma^2} [(n - \bar{n})^2 - (2p-1)(n - \bar{n})]. \tag{1.70}$$

With p not close to 0 or 1, we can set $2p-1 \approx 0$ to simplify (1.70):

$$f(n) \simeq -\ln \sqrt{2\pi\sigma^2} - \frac{(n - \bar{n})^2}{2\sigma^2}. \tag{1.71}$$

Now, given that $P(n) = e^{f(n)}$, we conclude that for n near $\bar{n}$,

$$P(n) \simeq \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-(n - \bar{n})^2}{2\sigma^2} \quad \begin{cases} \bar{n} = Np, \\ \sigma^2 = Np(1-p). \end{cases} \tag{1.72}$$

Equation (1.72) is the well-known gaussian approximation of the binomial distribution (1.66). The gaussian attains its maximum at the mean number of particles $\bar{n}$. It follows that when the number of particles N in a room is large, determining the mean number $\bar{n}$ that appear in a sub-room with probability p is equivalent to determining where the probability distribution $P(n)$ peaks.

How well does this gaussian function fit the binomial distribution? Figure 1.6 compares the two functions for the binomials in (1.14) and (1.15), where $N = 6$ and 60 , respectively, and $p = 1/3$. The discrete binomial distributions are shown as stem plots, and their gaussian approximations are superimposed as continuous curves. The fits are impressive, especially as N increases.

10²⁷ Particles in a Room

Recall the examples of equations (1.14), (1.15), and (1.26). We'll now place $N = 10^{27}$ particles in a room—which is about the number of air molecules in a real room. What is the chance that one third of them occupy the front third of the room?

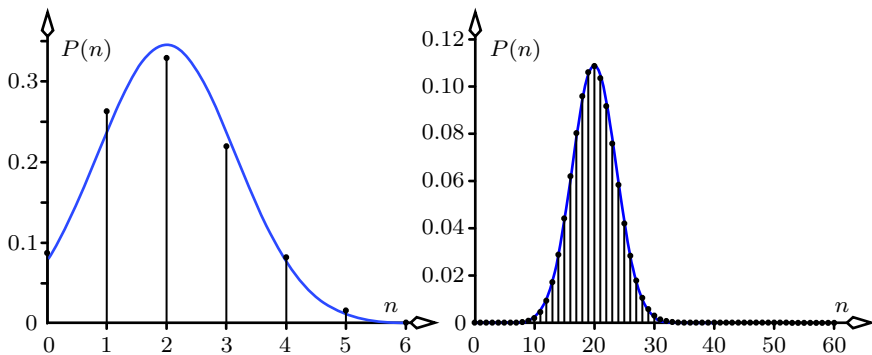


Fig. 1.6 Binomial distributions (stem plots) and their gaussian fits (continuous curves) for: (left) $N = 6$, $p = 1/3$, and (right) $N = 60$, $p = 1/3$. Each binomial was calculated from (1.66), and each gaussian from (1.72)

Just as in the former examples of $N = 6, 60, 6000$, we require $P(\bar{n})$. But N is so large here that, rather than calculate $10^{27}!$, we should instead approximate the binomial using (1.72). The mean number in the front third of the room is $\bar{n} = 1/3 \times 10^{27}$. Equation (1.72) gives

$$P(\bar{n}) \simeq \frac{1}{\sigma\sqrt{2\pi}}, \quad \text{with } \sigma = \sqrt{10^{27} \times 1/3 \times 2/3} = \sqrt{20}/3 \times 10^{13}; \quad (1.73)$$

and so

$$P(\bar{n}) \simeq \frac{1}{\sqrt{20}/3 \times 10^{13} \times \sqrt{2\pi}} \simeq 3 \times 10^{-14}. \quad (1.74)$$

It's not surprising that the chance of *exactly* one third of the 10^{27} particles occupying the front third of the room is minuscule. Now, what is the chance that this number fluctuates upward by 1%?

$$\begin{aligned} P(1.01\bar{n}) &\stackrel{(1.72)}{=} \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-(1.01\bar{n} - \bar{n})^2}{2\sigma^2} = P(\bar{n}) \exp \frac{-(0.01\bar{n})^2}{2\sigma^2} \\ &\simeq 3 \times 10^{-14} \times \exp(-2.5 \times 10^{22}) \approx 10^{-10^{22}}. \end{aligned} \quad (1.75)$$

This is *very* small: even a 1% fluctuation can be treated as never occurring, and it shows how *extremely peaked* the binomial is around its mean. More realistically, we might ask for the chance that the occupation number fluctuates by *at least* 1% up or down. That requires an integration; and with more effort, we can show that the answer is “close” to the number in (1.75). But the real take-home point to note is the extremely peaked nature of the distribution about the mean in the real world.

The precise form of the gaussian fit given in (1.72) is actually something of a lucky fluke. To see why, remember that in this section, we set out to approximate the binomial distribution (1.66). Being a probability distribution, (1.66) obeys

$$\sum_{n=0}^N P_n = 1. \quad (1.76)$$

Hence, the heights of the stems in each plot in Figure 1.6 sum to one. Now imagine widening each stem in that figure into a vertical bar of unit width: this width is dimensionless, which means that the bar's area equals its height. (Technically, we have converted the stem plot into a *histogram*.) We conclude that the total area under all such bars equals one. But if the gaussian (1.72) is to be a good fit to the binomial distribution, then we expect the area under this gaussian curve to be approximately one as well. Realise that n is a continuous variable for the gaussian, in which case this area is

$$\text{area} = \int_0^N \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-(n - \bar{n})^2}{2\sigma^2} dn. \quad (1.77)$$

We could calculate this area to any accuracy with the theory of Section 1.5, but such precision isn't needed. Note only that we are approximating the binomial by a gaussian in the case when N is large, and in that case, $\bar{n} = Np \gg 1$ and $\sigma = \sqrt{Np(1-p)} \ll \bar{n}$. So, effectively *all* of the gaussian's peak—and hence effectively all of its area—is well away from $n = 0$. It follows that the area in (1.77) approximates, to very high accuracy, the area under the entire gaussian. But that area just happens to be exactly one:

$$\text{area} = \int_0^N \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-(n - \bar{n})^2}{2\sigma^2} dn \simeq \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-(n - \bar{n})^2}{2\sigma^2} dn = 1, \quad (1.78)$$

where the value of this integral is calculated ahead in (1.103) and (1.104).

On a final note, we assumed just after (1.68) that p is not close to 0 or 1. But what if p is close to 0 or 1? It turns out that the same gaussian fit still works very well in that case. This is seen in Figure 1.7, which again compares stem plots of binomial distributions with gaussian curves for $N = 6$ and 60, respectively, but now, $p = 1/10$. (Setting p close to 1—say, $p = 9/10$ —gives the same degree of fit, but with the peak shifted to the right-hand end of each plot.) The fit in the $N = 6$ case is not wonderful; but it doesn't have to be, because the gaussian approximation is only used for large N anyway. In contrast, the $N = 60$ fit is very accurate.

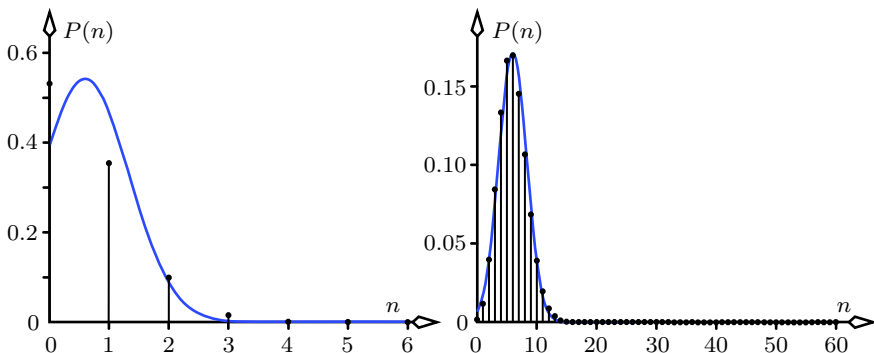


Fig. 1.7 Binomial distributions (stem plots) and their gaussian fits (continuous curves) for: (left) $N = 6$, $p = 1/10$, and (right) $N = 60$, $p = 1/10$. Each binomial was calculated from (1.66), and each gaussian from (1.72)

1.5 Integrals of the Gaussian Function

In Section 1.4, we saw that the binomial distribution can be approximated by the gaussian function. The gaussian function is encountered frequently throughout mathematical physics. Its dominance in probability theory means that even when it doesn't explicitly arise from a first-principles analysis of a physical situation, it can still be used to model very complicated systems. We will use the integral of the gaussian in subsequent chapters, and so it's important for us to get acquainted with its form.

The integral of the basic gaussian function e^{-x^2} is not a simple collection of power functions, exponentials, or sines and cosines; but it is certainly well defined, because the area under the curve $y = e^{-x^2}$ is well defined. Consider for a moment, that the function $1/x$ cannot be integrated using the same easy rule used to integrate every other power of x ; and yet we know that its integral *is* well defined because the area under the curve $y = 1/x$ is well defined, provided we stay away from $x = 0$. We simply *define* the integral of $1/x$ to be a new function—call it $L(x)$ —and then proceed to investigate the properties of $L(x)$ based on the area under the curve $y = 1/x$. We soon find that $L(x)$ behaves identically to a logarithm in *all* respects. That means it must *be* a logarithm. Inverting it then yields the exponential function with its base e .

The same idea applies to integrating e^{-x^2} . We have no rules that allow e^{-x^2} to be integrated so as to produce any function that can be written with a finite number of powers of x , or manipulations of $\sin x$, and so on. In that case, we simply define a new function whose derivative is e^{-x^2} . That function is conventionally called $\sqrt{\pi}/2 \operatorname{erf} x$, where $\operatorname{erf} x$ is the “error function”, a name that springs from its use in the statistical theory of errors. We'll see soon that including the factor of $\sqrt{\pi}/2$ allows $\operatorname{erf} x$ to tend toward the convenient

value of 1 as x tends toward infinity.⁵ We could note that

$$\int_0^x e^{-x^2} dx = \frac{\sqrt{\pi}}{2} (\operatorname{erf} x - \operatorname{erf} 0) \quad (1.79)$$

and now *define* $\operatorname{erf} 0 \equiv 0$, or we could have defined erf originally as

$$\operatorname{erf} x \equiv \frac{2}{\sqrt{\pi}} \int_0^x e^{-x^2} dx, \quad (1.80)$$

in which case it's clear that $\operatorname{erf} 0 = 0$. Next, we can show that erf is an odd function by using the fact that e^{-x^2} is even, to write

$$\begin{aligned} \operatorname{erf}(-x) &= \frac{2}{\sqrt{\pi}} \int_0^{-x} e^{-x^2} dx = \frac{-2}{\sqrt{\pi}} \int_{-x}^0 e^{-x^2} dx \\ &= \frac{-2}{\sqrt{\pi}} \int_0^x e^{-x^2} dx = -\operatorname{erf} x. \end{aligned} \quad (1.81)$$

This proves that erf is odd. Also,

$$\frac{d}{dx} \operatorname{erf} x = \frac{2}{\sqrt{\pi}} e^{-x^2} > 0, \quad (1.82)$$

so that erf is strictly increasing with x . Next, we give a value to $\operatorname{erf} \infty$ as follows. Write $I \equiv \int_0^\infty e^{-x^2} dx$, and note that⁶

$$I^2 = \int_0^\infty \int_0^\infty e^{-x^2} e^{-y^2} dx dy = \int_{\text{first quadrant of } xy \text{ plane}} e^{-(x^2+y^2)} dx dy. \quad (1.83)$$

This last integral converts to polar coordinates as [with a side note in the grey box just after (1.84)]

$$\begin{aligned} I^2 &= \int_0^{\pi/2} \int_0^\infty e^{-r^2} r dr d\theta = \int_0^{\pi/2} d\theta \int_0^\infty dr e^{-r^2} r \\ &= \frac{\pi}{2} \left[\frac{-e^{-r^2}}{2} \right]_0^\infty = \frac{\pi}{4}. \end{aligned} \quad (1.84)$$

⁵ Despite the extremely widespread occurrence of e^{-x^2} in all physical fields, its integral in terms of $\operatorname{erf} x$ is just not something that most physicists commit to memory. That is probably due to the presence of the untidy factor of $\sqrt{\pi}/2$: contrast this with the case of the integral of $1/x$ being simply $\ln x$. And perhaps the language is a hindrance: the name “error function” does no justice to a function that is virtually never used by physicists in the context of the theory of statistical error.

⁶ Strictly speaking, we should not assume a priori that I exists; instead, we should consider a limit as x tends to infinity. But I wish to keep this analysis brief.

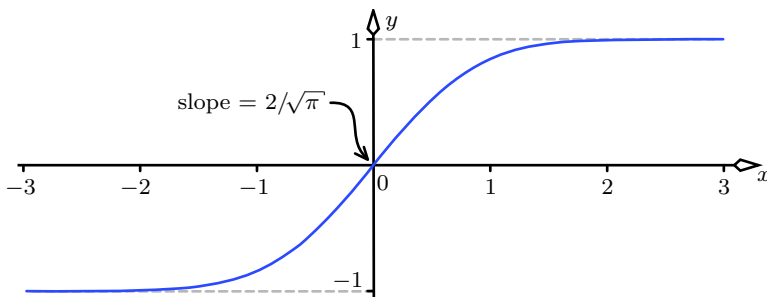


Fig. 1.8 A plot of $y = \operatorname{erf} x$. This has the distinctive shape of a cumulative probability function—which is precisely what it is, up to a scaling and vertical shift

The last double integral in the first line of (1.84) is still a pair of nested integrals, but it reads from left to right instead of the nested outside-to-inside fashion of that line's first double integral. This left-to-right notation (which is quite standard) is not only easy to read, but it automatically factors (1.84) into two separate integrals, which are then easily evaluated on that equation's second line. It's also worth noting that $dx dy$ in (1.83) does not actually *equal* (1.84)'s $r dr d\theta$. These two volume elements describe different infinitesimal-volume cells, and only the weighted aggregation of these cells (namely, the integral) is independent of the choice of coordinates.

We infer from (1.84) that $I = \sqrt{\pi}/2$, or

$$\frac{\sqrt{\pi}}{2} = \int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2} \operatorname{erf} \infty. \quad (1.85)$$

It follows that $\operatorname{erf} \infty = 1$. This simple result is the reason for why the convenience factor of $\sqrt{\pi}/2$ is conventionally included in the definition of erf . Figure 1.8 shows a plot of $y = \operatorname{erf} x$.

As an aside, it turns out that for all *complex* z , $\operatorname{erf}(-z) = -\operatorname{erf} z$, and $\operatorname{erf} z \rightarrow 1$ as $|z| \rightarrow \infty$, provided $|\arg z| < \pi/4$.

We can now evaluate the general one-dimensional gaussian integral

$$\int e^{-ax^2+bx} dx. \quad (1.86)$$

Here, a is a positive real number (positive to ensure we are really dealing with a gaussian function) and b is any real number. Complete the square, by writing

$$\begin{aligned}
 -ax^2 + bx &= -a \left(x^2 - \frac{bx}{a} \right) = -a \left[\left(x - \frac{b}{2a} \right)^2 - \frac{b^2}{4a^2} \right] \\
 &= -a \left(x - \frac{b}{2a} \right)^2 + \frac{b^2}{4a}.
 \end{aligned} \tag{1.87}$$

Change variables from x to y :

$$y \equiv \sqrt{a} \left(x - \frac{b}{2a} \right), \tag{1.88}$$

and now write⁷

$$\begin{aligned}
 \int e^{-ax^2+bx} dx &= e^{b^2/(4a)} \int \frac{e^{-y^2}}{\sqrt{a}} dy = \frac{1}{\sqrt{a}} e^{b^2/(4a)} \frac{\sqrt{\pi}}{2} \operatorname{erf} y \\
 &= \frac{1}{2} \sqrt{\frac{\pi}{a}} e^{b^2/(4a)} \operatorname{erf} \left(\sqrt{a} x - \frac{b}{2\sqrt{a}} \right).
 \end{aligned} \tag{1.89}$$

This very useful result is worth committing to memory. Listed below are some special cases of the gaussian integral that we'll use in the coming chapters. For now, you may wish simply to peruse them, leaving a more detailed reading of each item to the time when it's referenced in the text.

Useful Gaussian Integrals

1. The most basic case is

$$\int_0^\infty e^{-ax^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{a}} \operatorname{erf} \sqrt{a} x \Big|_0^\infty = \frac{1}{2} \sqrt{\frac{\pi}{a}}. \tag{1.90}$$

2. The next is only slightly more complicated:

$$\int_{-\infty}^\infty e^{-ax^2} dx = 2 \int_0^\infty e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}. \tag{1.91}$$

This is the $b = 0$ case of

⁷ Expressions such as $b^2/(4a)$ in these integrals are written by many physicists as “ $b^2/4a$ ”, but this omission of the parentheses actually runs contrary to established convention. For example, you will usually see the Boltzmann distribution's “ $-E/(kT)$ ” written as “ $-E/kT$ ”. But the conventional way to process multiplication and division is strictly from left to right—a protocol that is obeyed by computer languages and digital calculators; so, “ $1/kT$ ” really means $(1/k) \times T$, or T/k . By the same token, “ $1/2$ metre” is universally and correctly understood to mean $1/2 \times$ a metre, or half a metre. It does not mean “ $1/(2 \text{ metres})$ ”. I have consistently followed the standard left-to-right convention throughout this text, and thus I always include parentheses where necessary.

$$\int_{-\infty}^{\infty} e^{-ax^2+bx} dx = \frac{1}{2} \sqrt{\frac{\pi}{a}} e^{b^2/(4a)} \left[\operatorname{erf} \left(\sqrt{a} x - \frac{b}{2\sqrt{a}} \right) \right]_{-\infty}^{\infty} = \sqrt{\frac{\pi}{a}} e^{b^2/(4a)}. \quad (1.92)$$

3. The area under the right-hand tail of a normal distribution is sometimes required in probability theory:

$$\int_x^{\infty} e^{-ax^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{a}} \operatorname{erf} \sqrt{a} x \Big|_x^{\infty} = \frac{1}{2} \sqrt{\frac{\pi}{a}} (1 - \operatorname{erf} \sqrt{a} x). \quad (1.93)$$

The function “ $1 - \operatorname{erf}$ ” is known as erfc , the “complementary error function”:

$$\int_x^{\infty} e^{-ax^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{a}} \operatorname{erfc} \sqrt{a} x. \quad (1.94)$$

4. The next integral is easily evaluated with an educated guess:

$$\int x e^{-ax^2} dx = \frac{-1}{2a} e^{-ax^2}. \quad (1.95)$$

5. This one is evaluated by parts. For clarity, we under-bracket each of the parts.

$$\begin{aligned} \int x^2 e^{-ax^2} dx &= \int \underbrace{x} \underbrace{x e^{-ax^2}} dx = \frac{-x}{2a} e^{-ax^2} + \frac{1}{2a} \int e^{-ax^2} dx \\ &= \frac{-x}{2a} e^{-ax^2} + \frac{1}{4a} \sqrt{\frac{\pi}{a}} \operatorname{erf} \sqrt{a} x. \end{aligned} \quad (1.96)$$

6. An instance of (1.96) that we’ll use later is

$$\int_x^{\infty} x^2 e^{-ax^2} dx = \frac{x}{2a} e^{-ax^2} + \frac{1}{4a} \sqrt{\frac{\pi}{a}} \operatorname{erfc} \sqrt{a} x. \quad (1.97)$$

7. Applying commonly used limits to (1.96),

$$\int_0^{\infty} x^2 e^{-ax^2} dx = \left[\frac{-x}{2a} e^{-ax^2} + \frac{1}{4a} \sqrt{\frac{\pi}{a}} \operatorname{erf} \sqrt{a} x \right]_0^{\infty} = \frac{1}{4a} \sqrt{\frac{\pi}{a}}. \quad (1.98)$$

This last integral could have been evaluated more easily by treating a as a variable and calculating d/da of (1.90):

$$\frac{d}{da} \int_0^{\infty} e^{-ax^2} dx = \frac{d}{da} \frac{1}{2} \sqrt{\frac{\pi}{a}} = \frac{-\sqrt{\pi}}{4} a^{-3/2}. \quad (1.99)$$

The left-hand side of this expression then evaluates via “differentiation [by a] under the integral sign”, to give⁸

$$\int_0^\infty -x^2 e^{-ax^2} dx = \frac{-\sqrt{\pi}}{4} a^{-3/2}. \quad (1.100)$$

Cancelling the minus sign from each side produces (1.98).

8. Another set of commonly used limits for (1.96) is

$$\int_{-\infty}^\infty x^2 e^{-ax^2} dx = 2 \int_0^\infty x^2 e^{-ax^2} dx \stackrel{(1.98)}{=} \frac{1}{2a} \sqrt{\frac{\pi}{a}}. \quad (1.101)$$

9. This integral is evaluated by parts (indicated by under-brackets), with the crossed-out term equalling zero:

$$\begin{aligned} \int_0^\infty x^3 e^{-ax^2} dx &= \int_0^\infty \underbrace{x^2}_{\cancel{\frac{x^2}{2a}}} \underbrace{x e^{-ax^2}}_{\cancel{e^{-ax^2}}} dx = \left[\frac{-x^2}{2a} e^{-ax^2} \right]_0^\infty + \frac{1}{a} \int_0^\infty x e^{-ax^2} dx \\ &\stackrel{(1.95)}{=} \frac{-1}{2a^2} \left[e^{-ax^2} \right]_0^\infty = \frac{1}{2a^2}. \end{aligned} \quad (1.102)$$

10. We can easily check the normalisation of the standard expression for a normal distribution with mean μ and variance σ^2 ,

$$\mathcal{N}(x; \mu, \sigma^2) \equiv \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-(x-\mu)^2}{2\sigma^2}, \quad (1.103)$$

by realising that the function can be shifted by μ with impunity:

$$\begin{aligned} \int_{-\infty}^\infty \mathcal{N}(x; \mu, \sigma^2) dx &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^\infty \exp \frac{-x^2}{2\sigma^2} dx \\ &= \frac{1}{\cancel{\sigma\sqrt{2\pi}}} \frac{1}{2} \cancel{\sqrt{\pi 2\sigma^2}} \operatorname{erf} \frac{x}{\sigma\sqrt{2}} \Big|_{-\infty}^\infty = 1. \end{aligned} \quad (1.104)$$

11. More generally, the two-sided area under a zero-mean normal distribution of standard deviation σ is

$$\int_{-x}^x \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-x^2}{2\sigma^2} dx = \frac{1}{\cancel{\sigma\sqrt{2\pi}}} \frac{1}{2} \cancel{\sqrt{\pi 2\sigma^2}} \operatorname{erf} \frac{x}{\sigma\sqrt{2}} \Big|_{-x}^x = \operatorname{erf} \frac{x}{\sigma\sqrt{2}}. \quad (1.105)$$

1.5.1 Calculating the Error Function Numerically

Defining the error function as an integral is all very well, but can we evaluate it numerically from that definition? The Taylor series for e^{-x^2} is well behaved,⁹ in which case

$$\begin{aligned}\operatorname{erf} x &\equiv \frac{2}{\sqrt{\pi}} \int_0^x e^{-x^2} dx = \frac{2}{\sqrt{\pi}} \int_0^x \left[1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \dots \right] dx \\ &= \frac{2}{\sqrt{\pi}} \left[x - \frac{x^3}{3} + \frac{x^5}{5 \cdot 2!} - \frac{x^7}{7 \cdot 3!} + \dots \right].\end{aligned}\quad (1.106)$$

This series converges quickly for the small values of x that are usually encountered with the error function. To demonstrate, we can use it to reproduce the value of 68% that is frequently quoted in “1-sigma” discussions of the normal distribution. This value is actually the area under the zero-mean normal distribution from $-\sigma$ to σ , and is

$$\int_{-\sigma}^{\sigma} \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-x^2}{2\sigma^2} dx \stackrel{(1.105)}{=} \operatorname{erf} \frac{1}{\sqrt{2}}. \quad (1.107)$$

Evaluate this by setting $x = 1/\sqrt{2}$ in (1.106). This produces the 5 decimal-place result of 0.68269, requiring just 6 terms in the sum. Similarly, the “two-sided n -sigma” value is the area under the zero-mean normal distribution from $-n\sigma$ to $n\sigma$:

$$\int_{-n\sigma}^{n\sigma} \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-x^2}{2\sigma^2} dx \stackrel{(1.105)}{=} \operatorname{erf} \frac{n}{\sqrt{2}}. \quad (1.108)$$

When $n = 2$, setting $x = 2/\sqrt{2}$ in (1.106) gives us 0.95450, needing 12 terms in the sum. For $n = 3$, $x = 3/\sqrt{2}$ in (1.106) gives 0.99730, needing 19 terms. With $n = 4$, we obtain 0.99994 in 30 terms. These figures (or at least the first three, 68%, 95.5%, and 99.7%) are well known to statisticians.

In fact, series expansions of functions are not always the best tool for serious numerical work, but the series for $\operatorname{erf} x$ in (1.106) shows that we can treat the error function on a par with other functions that have more familiar series expansions, such as e^x and $\sin x$. The only difficulty is that when x becomes larger than 2 or 3, the series (1.106) requires many terms to converge. This potential problem occurs in Section 8.4, when we wish to use some kind of series expansion for the function erfc in (1.94):

$$\operatorname{erfc} x = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-x^2} dx. \quad (1.109)$$

⁹ “Well behaved” here means that the series is uniformly convergent on bounded subsets of the real numbers.

This integral tends rapidly to zero for large x , and yet the series in (1.106) cannot practically be applied when x is large, because it converges too slowly. Instead, we can calculate an *asymptotic* series for $\operatorname{erfc} x$ in the following way. Write (1.109) as

$$\frac{\sqrt{\pi}}{2} \operatorname{erfc} x = \int_x^\infty e^{-x^2} dx = \int_x^\infty \frac{-1}{2x} \times -2x e^{-x^2} dx. \quad (1.110)$$

Evaluate the last integral by parts:

$$\begin{aligned} \frac{\sqrt{\pi}}{2} \operatorname{erfc} x &= \frac{-1}{2x} e^{-x^2} \Big|_x^\infty - \int_x^\infty \frac{1}{2x^2} e^{-x^2} dx \\ &= \frac{e^{-x^2}}{2x} - \int_x^\infty \frac{-1}{4x^3} \times -2x e^{-x^2} dx. \end{aligned} \quad (1.111)$$

Again, evaluate the last integral by parts, yielding

$$\frac{\sqrt{\pi}}{2} \operatorname{erfc} x = \frac{e^{-x^2}}{2x} - \frac{e^{-x^2}}{4x^3} + \int_x^\infty \frac{-3}{8x^5} \times -2x e^{-x^2} dx, \quad (1.112)$$

and so on—the integrations by parts carry on indefinitely to produce a series in odd powers of $1/x$. This is an asymptotic series, meaning that any truncation of it will be increasingly accurate as $x \rightarrow \infty$. In fact, even when it's truncated after just one term to give

$$\operatorname{erfc} x \simeq \frac{e^{-x^2}}{x\sqrt{\pi}}, \quad (1.113)$$

it is already accurate to a few percent when x is 4 or 5. For example, consider that $\operatorname{erfc} 10 = 2.088 \times 10^{-45}$. Evaluating (1.113) with $x = 10$ yields 2.099×10^{-45} . But now write $\operatorname{erfc} 10 = 1 - \operatorname{erf} 10$, and then use $x = 10$ in (1.106): you will find that evaluating the series is not easy! We'll put (1.113) to good use in Section 8.4.

1.5.2 The 3-Dimensional Gaussian

We will occasionally require a definite triple integral of a gaussian function. The simplest such integral could be considered as being performed over all values of the spatial dimensions x, y, z :

$$I = \int_{\text{all space}} e^{-a(x^2+y^2+z^2)} dx dy dz. \quad (1.114)$$

This separates into a product of three identical integrals:

$$\begin{aligned}
 I &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-ax^2 - ay^2 - az^2} dx dy dz = \int_{-\infty}^{\infty} e^{-ax^2} dx \int_{-\infty}^{\infty} e^{-ay^2} dy \int_{-\infty}^{\infty} e^{-az^2} dz \\
 &= \left[\int_{-\infty}^{\infty} e^{-ax^2} dx \right]^3 \stackrel{(1.91)}{=} (\pi/a)^{3/2}.
 \end{aligned} \tag{1.115}$$

An alternative approach to evaluating I will shortly turn out to be useful for another task: we could change variables in the original three-dimensional integral from cartesian x, y, z to spherical polar r, θ, ϕ , mimicking the approach of (1.84) by writing [see the grey box just after (1.84)]

$$\begin{aligned}
 I &= \int_0^{2\pi} \int_0^{\pi} \int_0^{\infty} e^{-ar^2} r^2 \sin \theta dr d\theta d\phi \\
 &= \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin \theta \int_0^{\infty} dr r^2 e^{-ar^2} \\
 &\stackrel{(1.98)}{=} 2\pi \times 2 \times \frac{1}{4a} \sqrt{\frac{\pi}{a}} = (\pi/a)^{3/2},
 \end{aligned} \tag{1.116}$$

just as we found in (1.115).

Another useful integral is calculated over all values of each of the three *cartesian* components of the vector $\mathbf{v}$, which has length v :

$$I = \iiint_{-\infty}^{\infty} v e^{-av^2} d^3v, \tag{1.117}$$

where $d^3v \equiv dv_x dv_y dv_z$. Again, we can use polar coordinates:

$$\begin{aligned}
 \iiint_{-\infty}^{\infty} v e^{-av^2} d^3v &= \iiint v e^{-av^2} v^2 \sin \theta dv d\theta d\phi \\
 &= \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin \theta \int_0^{\infty} dv v^3 e^{-av^2} \stackrel{(1.102)}{=} 2\pi \times 2 \times \frac{1}{2a^2} = \frac{2\pi}{a^2}.
 \end{aligned} \tag{1.118}$$

We will also encounter an instance of the normalised three-dimensional gaussian probability density. This generalises the one-dimensional version (1.103) to

$$\mathcal{N}(x, y, z; \mu_x, \sigma_x^2, \dots, \mu_z, \sigma_z^2) = \frac{1}{\sigma_x \sigma_y \sigma_z (2\pi)^{3/2}} \exp \left[\frac{-(x - \mu_x)^2}{2\sigma_x^2} - \frac{(y - \mu_y)^2}{2\sigma_y^2} - \frac{(z - \mu_z)^2}{2\sigma_z^2} \right]. \quad (1.119)$$

This expression describes a peaked density function that is symmetrical about the x , y , and z axes. Although we won't need any further analysis of this, consider writing

$$\mathbf{x} \equiv \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad \boldsymbol{\mu} \equiv \begin{bmatrix} \mu_x \\ \mu_y \\ \mu_z \end{bmatrix}, \quad P \equiv \begin{bmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_y^2 & 0 \\ 0 & 0 & \sigma_z^2 \end{bmatrix}. \quad (1.120)$$

This renders (1.119) as

$$\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, P) = \frac{1}{\sqrt{\det P} (2\pi)^{3/2}} \exp \left[\frac{-1}{2} (\mathbf{x} - \boldsymbol{\mu})^t P (\mathbf{x} - \boldsymbol{\mu}) \right], \quad (1.121)$$

where $\det P$ is the determinant of the *covariance matrix* P . We can rotate and/or stretch the shape of the peak by defining a new set of three variables by way of a linear transformation $\mathbf{x} \rightarrow L\mathbf{x}$ for some matrix L . The gaussian function can then be written more generally in n dimensions as

$$\exp(-\mathbf{x}^t A \mathbf{x} + \mathbf{b}^t \mathbf{x}) \quad (1.122)$$

for some real symmetric $n \times n$ matrix A and real column vector $\mathbf{b}$ that are each built from L, P , and $\boldsymbol{\mu}$. Analogously to (1.92), it can be shown, using the procedure of “orthogonal diagonalisation” found in linear algebra, that the integral of (1.122) over all n dimensions is

$$\int_{\text{all space}} \exp(-\mathbf{x}^t A \mathbf{x} + \mathbf{b}^t \mathbf{x}) \, dx_1 \dots dx_n = \frac{\pi^{n/2}}{\sqrt{\det A}} \exp \frac{1}{4} \mathbf{b}^t A^{-1} \mathbf{b}. \quad (1.123)$$

1.6 Increases and Infinitesimals

In the chapters to come, we will often make use of *infinitesimals* to simplify all manner of analyses. Infinitesimals, also known as *differentials*, occur widely in physics. They have been used to represent “infinitely small” quantities from the earliest days of calculus. In the last century or so, calculus evolved its modern “epsilon-delta” language of limits and Riemann sums, and this language has given rise to a more sophisticated use of the infinitesimal. While it can still be treated as an “infinitely small” quantity, the infinitesimal's modern use is to reduce the length of the rigorous but rather long-winded

limit analyses of calculus. Any expression written using infinitesimals can easily be re-expressed in the modern language of limits; but this modern language can be tedious and tends to clutter real calculations without giving anything useful in return. Infinitesimals offer a way to reduce that clutter, and they will become indispensable throughout our study of statistical mechanics.

To investigate what an infinitesimal really is, we begin by establishing a standard phrasing for quantifying the core idea of calculus, which is “how quantities change”. Physicists, mathematicians, and engineers routinely give the word “change” two distinct meanings that, unfortunately, clash mathematically. The first meaning is the everyday meaning: when a quantity “changes by 2”, its value might go up or down by 2; the actual direction is not apparent and not always important. In this book, we will assign *only* this everyday meaning to the word “change”. Occasionally, we are not concerned as to whether some quantity has increased or decreased, and thus to describe it as “changing” is quite sufficient. So, when we write “ x changes by 2” in the pages to come, it will signify that we don’t care whether x went up by 2 or whether it went down by 2; we only care that x is 2 away from where it was.

But sometimes we require more information, and simply saying that a quantity has changed is not enough. If your bank tells you that “your balance has changed by \$1000”, the first thing you will ask is whether that balance has *increased* by \$1000 or *decreased* by \$1000: the answer makes all the difference! You would almost certainly prefer if they told you more specifically that “your balance has increased by \$1000”, or perhaps “your balance has decreased by \$1000”; these words “increase” and “decrease” specify what we really want to know. And that brings in the second meaning routinely given to the word “change” by scientists: that it is purely an increase. With this meaning, “the change in x is 2” means that x increased by 2.

But I see no advantage to be gained by adding that second meaning of “increase” to the everyday word “change”, because we already have a very specific word meaning “increase”: that very specific word is (you might have guessed) “increase”. Unlike the everyday meaning of “change”, the word “increase” denotes an addition, and the word “decrease” denotes a subtraction; and these meanings give them a mathematical precision and power that the word “change”, with its everyday meaning, lacks. If we say “ x increased by 2”, then we certainly mean the value of x went *up* by 2. If we say “ x decreased by 2”, then we mean the value of x went *down* by 2. But now call upon *negative* numbers. Adding and subtracting negative numbers are perfectly legal mathematical operations. Saying “ x increased by -2 ” means the value of x went up by -2 ; or in other words, the value of x went down by 2 (because -2 was added to x). And saying “ x decreased by -2 ” means the value of x went down by -2 ; or in other words, the value of x went up by 2 (because -2 was subtracted from x).

Mathematically, the word “decrease” might be seen as superfluous, because a decrease of c means an increase of $-c$. But “decrease” is no more superfluous

than “subtract” (subtracting c is equivalent to adding $-c$), and, of course, no one would think of avoiding the word “subtract” in everyday or mathematical speech. Also, linguistically, no one will dispute that in everyday conversation, it is far more preferable to say “my weight decreased by 2 kg” than it is to use the equivalent phrase “my weight increased by -2 kg”.

The distinction between the not very mathematically useful “change” and the very mathematically useful “increase/decrease” is similar to—and as important as—the distinction between “distance” and “displacement”:

$$\begin{aligned}\text{distance} &\equiv |\text{displacement}|, \\ \text{change} &\equiv |\text{increase}| = |\text{decrease}|.\end{aligned}\tag{1.124}$$

The mathematical symbol for “the increase in” is Δ . This symbol is almost universally voiced as “the change in”, but the above discussion shows that this carries no precise mathematical meaning, because most scientists use “change” ambiguously with both of the two meanings above. The precise and unique mathematical meaning of Δ is “the *increase* in”. Δ is sometimes incorrectly used to denote a decrease, and such use tends to make wrong signs appear in the resulting expressions. When caught in that situation, some authors simply insert ad-hoc minus signs to fix things. Incorrect mathematics is not a subjective thing; it does not come down to a question of semantics. It appears, then, that many users of the word “change” are confused by it, both linguistically and mathematically. This is why I suggest that the word “change” should not be given two concurrent and clashing meanings.

Writing Δ implies that the concepts of “initial” and “final” are understood: ΔA correctly denotes $A_{\text{final}} - A_{\text{initial}}$:

$$\begin{aligned}\Delta A &\equiv A_{\text{final}} - A_{\text{initial}} \equiv \text{increase or gain in } A, \\ -\Delta A &= \text{decrease or loss in } A.\end{aligned}\tag{1.125}$$

Δ is the primal symbol with this meaning, and the related symbols of calculus such as d , $\partial/\partial x$, and ∇ inherit that meaning. When we write $\partial f/\partial x$ and ∇f in the context of partial derivatives, we are referring to increases in f that happen while (or because) other quantities are being increased. Given a function $y = f(x)$, it is always correct to write $y + \Delta y = f(x + \Delta x)$ irrespective of the behaviour of x and y . If you find yourself ever wanting to examine $f(x - \Delta x)$, be sure to determine whether this is really what you want, because it almost certainly is not. In general, $f(x - \Delta x)$ has no simple relationship to either $y + \Delta y$ or $y - \Delta y$. The expression $x - \Delta x$ will probably never appear in any correct nontrivial analysis.¹⁰

¹⁰ I am being deliberately vague by saying “probably”, because we can always construct a *trivial* example involving $x - \Delta x$: if we locate x as a point on a number line, and then move to either side to construct a new point $x + \Delta x$, then $x - \Delta x$ is the mirror-image point an equal distance from x , on the other side of x . Nonetheless, the first new point we construct *must* be written as $x + \Delta x$, and never $x - \Delta x$.

The convention that “ Δ = increase, $-\Delta$ = decrease” also applies to vectors, and indeed to anything else: by $\Delta \mathbf{v}$, we mean $\mathbf{v}_{\text{final}} - \mathbf{v}_{\text{initial}}$, the “increase in $\mathbf{v}$ ”. The idea of a vector increasing might not be as intuitive as it is for numbers, because the *length* of $\mathbf{v}$ needn’t change as $\mathbf{v}$ evolves. But magnitude or length are not part of the definition of “increase”. The key phrases are

$$\begin{aligned}\Delta \text{ and } d &= \text{increase} = \text{final} - \text{initial}, \\ -\Delta \text{ and } -d &= \text{decrease} = \text{initial} - \text{final},\end{aligned}\tag{1.126}$$

and they apply to vectors as well as to numbers. The phrase “ $\Delta \mathbf{v}$ is the increase in vector $\mathbf{v}$ ” is certainly a more sophisticated use of the word “increase” than is the case for numbers; but then again, a vector is a more sophisticated object than a number.

Remembering (1.126) will always bring you to the correct result, as well as often fixing ungainly language. For example, consider the following clumsy phrase: “the negative of the change in A ”, which can be found occasionally in physics books. It’s clear that whoever wrote the phrase was translating the expression “ $-\Delta A$ ” into English. We see immediately that what is really meant is the much clearer phrase “the decrease in A ”. The blundering words “the negative of the change in” have no place in physics, nor in everyday speech.

A useful and important property of Δ is its linearity. We can prove that it is linear by considering its action on the following function $f(x, y)$:

$$f = ax + by,\tag{1.127}$$

where a and b are constants. We ask: how does f increase when x and y increase? By definition, increases Δx and Δy give rise to an increase Δf :

$$f + \Delta f = a(x + \Delta x) + b(y + \Delta y).\tag{1.128}$$

Subtracting (1.127) from (1.128) yields

$$\Delta f = a\Delta x + b\Delta y.\tag{1.129}$$

Now combine (1.127) and (1.129) as

$$\Delta(ax + by) = a\Delta x + b\Delta y.\tag{1.130}$$

We conclude from (1.43) that the operation of finding the increase is linear.

Infinitesimal quantities such as dE , where E is energy, are used extensively in statistical mechanics. To see what they represent, recall the most basic definition of the derivative of a function $f(x)$:

$$f'(x) \equiv \lim_{\Delta x \rightarrow 0} \frac{\Delta f}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.\tag{1.131}$$

Suppose that we insert a Taylor expansion of $f(x + \Delta x)$ into (1.131), to see where it will take us:

$$\begin{aligned} f'(x) &= \lim_{\Delta x \rightarrow 0} \frac{f(x) + f'(x) \Delta x + \frac{1}{2!} f''(x) \Delta x^2 + \cdots - f(x)}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} f'(x) + \frac{1}{2!} f''(x) \Delta x + \cdots \\ &= f'(x), \quad \text{as expected.} \end{aligned} \tag{1.132}$$

This circular-looking piece of algebra contains the essence of what infinitesimals are about. The Taylor expansion is an infinite series in powers of Δx , but once we have divided by Δx in (1.132), all terms except $f'(x) \Delta x$ are destined to vanish in the limit as Δx tends to zero. This means we can just as well abbreviate (1.132) by retaining only $f'(x) \Delta x$ in the Taylor series. When doing this, we write Δx as “ dx ” to indicate that we are “not bothering to write down the higher-order terms”—but that we are aware that they *are* invisibly present. Equation (1.131) then becomes the very streamlined, and deceptively obvious-looking expression

$$f'(x) \equiv \frac{df}{dx} \equiv \frac{f(x + dx) - f(x)}{dx} = \frac{f(x) + f'(x) dx - f(x)}{dx}. \tag{1.133}$$

This last expression is an economical and elegant way of writing (1.131) that has the process of taking the limit built in. Notice that (1.133) is *exact*: we are not making any approximation by “dropping higher-order terms”, because, in fact, we have *not* dropped higher-order terms. Those terms are all invisibly present, sitting on the shoulders of dx : by writing “ dx ”, we really mean Δx plus terms of higher order in Δx along with a statement of an eventual division by Δx and a limit being taken as $\Delta x \rightarrow 0$. So, when speaking of an infinitesimal, or an “infinitesimally small quantity”, we are really referring to the end result of a limit process applied to the non-infinitesimal Δx . This implies that it makes no sense to write a power series of an infinitesimal. Whereas the expression “ $\Delta x + \Delta x^2$ ” is certainly meaningful, the expression “ $dx + dx^2$ ” has no meaning at all.

The Delta Function

The idea of a sort of invisible procedure in (1.133) also appears in the theory of the *delta function*, which finds frequent use in Fourier analysis. This function, $\delta(x)$, is conventionally defined to be an infinitely tall spike at $x = 0$, and zero elsewhere, with $\int_{-\infty}^{\infty} \delta(x) dx \equiv 1$. It is usually treated as a function in its own right, but more correct is the idea that any expression involving the delta function implies a limit being taken of a sequence of similar expressions that each replace the delta with a

bell-shaped function. These bell-shaped functions become increasingly narrower and higher at each subsequent term in the sequence.

If the increase Δ is known to be small, it is often approximated by using differentiation. For example, given $y = x^2$, by how much does y increase when x decreases by 0.01? We don't need to visualise the parabolic shape of $y = x^2$ here and ponder whether and where y is going up when x is going down. We simply say, "We require Δy (the increase in y) when $-\Delta x$ (the decrease in x) equals 0.01". We then need only relate Δy to Δx :

$$\begin{aligned}\Delta y &= y + \Delta y - y = (x + \Delta x)^2 - x^2 \\ &= 2x \Delta x + \Delta x^2 \\ &= 2x \times -0.01 + (-0.01)^2 = -0.02x + 10^{-4}.\end{aligned}\tag{1.134}$$

Hence, y increases by $-0.02x + 10^{-4}$; it's entirely equivalent to say that y decreases by $0.02x - 10^{-4}$. In a context where 0.01 is considered "small", the 10^{-4} is negligible. We can then work to a high approximation by saying "We require dy (the small increase in y) when $-dx$ (the small decrease in x) equals 0.01". As an equation, this is

$$dy = 2x dx = 2x \times -0.01 = -0.02x.\tag{1.135}$$

That is, y increases by $-0.02x$ or, equivalently, y decreases by $0.02x$. The many references to "increase" and "decrease" here are deliberate, and their point is to show that, given the most abstruse and convoluted question as to how quantities are changing, everything is easily unravelled by applying the simple rule (1.125).

A classic example of approximating Δ with d appears in the field of laser physics, where a band of frequencies of light is regularly related to the corresponding band of wavelengths. To see how *very small* intervals of each relate to the other, we can treat each interval as having infinitesimal width. Then, starting with the usual wave expression $v = f\lambda$ (where v is the speed of light in the laser medium), we write $f = v/\lambda$, then differentiate to obtain

$$df = -v/\lambda^2 d\lambda.\tag{1.136}$$

For small but non-infinitesimal bands, this becomes

$$\Delta f \simeq -v/\lambda^2 \Delta\lambda.\tag{1.137}$$

This last expression says that the *increase* in frequency, Δf , approximately equals v/λ^2 times the *decrease* in wavelength, $-\Delta\lambda$. Or equivalently, the *decrease* in frequency, $-\Delta f$, approximately equals v/λ^2 times the *increase* in wavelength, $\Delta\lambda$. Laser physicists are well aware that an increase/decrease

in frequency corresponds to a decrease/increase in wavelength, and so are apt to leave the minus sign out; but then, in place of the correct expression $|\Delta f| \simeq v/\lambda^2 |\Delta\lambda|$, they write the incorrect “ $\Delta f \simeq v/\lambda^2 \Delta\lambda$ ”, which has neither a minus sign nor absolute-value bars. This apparent economy of notation must then constantly be “corrected for” in the relevant laser mathematics, by ad-hoc insertions of minus signs to keep things on track.

Words with a sense of direction, such as increase, decrease, gain and loss, are crucial in enabling us to translate a task into useful mathematical language, and examples abound throughout physics. The force $\mathbf{F}$ on a mass equals that mass’s increase in momentum $d\mathbf{p}$ divided by the amount dt that time increases during the interaction: $\mathbf{F} = d\mathbf{p}/dt$. The electric field $\mathbf{E}$ equals the spatial rate of decrease of the electric potential Φ plus the temporal rate of decrease of the magnetic potential $\mathbf{A}$: so, $\mathbf{E} = -\nabla\Phi - \partial\mathbf{A}/\partial t$. Ohm’s rule is written as $V = IR$, but noting that V is an alternative notation for $-\Delta\Phi$ shows that V is a *drop* in electric potential Φ across a resistance. This allows us to relate Ohm’s rule to Maxwell’s equations, as well as apply Kirchhoff’s laws correctly around an electric circuit. And we will encounter another example of this notion of “ d = infinitesimal increase” in Section 9.4.1, when we convert a spectrum as a function of frequency to a function of wavelength. Richard Feynman once remarked that physics is all about knowing where to put the minus sign; but, at least in the case of Δ , you need never divine whether something is growing or diminishing: the correct sign will always appear automatically when you remember that Δ means “the increase in”.

Although it is “the increase in” that carries the symbol Δ , it’s important to realise that the procedures of calculus don’t single out “increase” as being more special than “decrease”. The slope of the curve $y = f(x)$ is dy/dx , meaning the infinitesimal increase in y over the infinitesimal increase in x ; but note that this equals the ratio of infinitesimals $-dy/-dx$, which is the infinitesimal *decrease* in y over the infinitesimal *decrease* in x .

Treating the infinitesimals as separate quantities in the numerator and denominator of a derivative gives us a physical intuition that might not be apparent from simply interpreting the derivative as the slope of a function. For example, a material’s *compressibility* is defined as the ratio of the fractional decrease $-dV/V$ in its volume resulting from an increase dP in applied pressure:

$$\text{compressibility } \kappa \equiv \frac{-dV/V}{dP} \equiv \frac{1}{\text{bulk modulus } B}. \quad (1.138)$$

Compressibility is always a positive number. In contrast, simply viewing compressibility as “ $-1/V \times dV/dP$ ” and interpreting it as “the slope of a substance’s volume-versus-pressure curve divided by minus its volume” is not helpful at all in giving us an intuitive feel for compressibility.

Another example of the translation of the language of increase/decrease into mathematics occurs when we push on a piston to compress gas in a

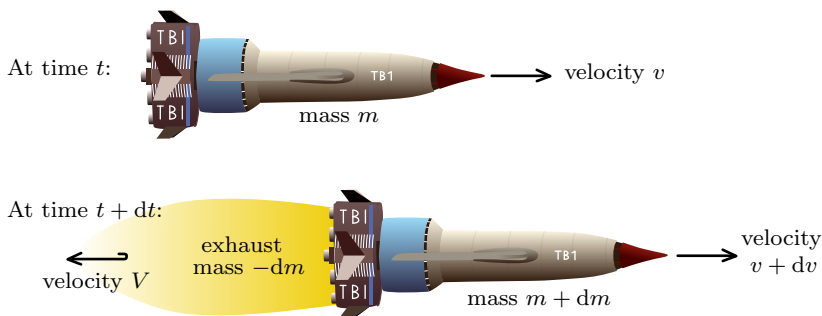


Fig. 1.9 The “rocket equation” that governs a rocket’s dynamics begins with an infinitesimal evolution in time. If Thunderbird 1’s mass at some moment is m , then after time dt its new mass is $m + dm$, and it has exhausted mass $-dm$. Writing Thunderbird’s mass as $m - dm$ with an exhausted mass of dm is incorrect: it will only get you into sign difficulties later. (The meaning of the exhaust’s velocity vector with the curved tail is explained in the text.)

cylinder. In Section 3.4.1, we’ll show that the work we do on the gas equals the pressure P that we apply (always positive) times the loss in volume $-dV$ (again positive); hence, the energy E of the gas increases by this amount: $dE = -PdV$. As the piston moves and the pressure changes while the volume decreases, we cannot simply write the total work that we do on the gas as “ $\Delta E = -P\Delta V$ ”. Instead, the total work we do is

$$\Delta E = \int dE = \int_{V_{\text{initial}}}^{V_{\text{final}}} -PdV. \quad (1.139)$$

This expression has precisely the same content as the infinitesimal version $dE = -PdV$.

Infinitesimals and the Rocket Equation

An example of the correct and incorrect uses of infinitesimals can be found in setting up the scenario for deriving the *rocket equation*. This is the central equation describing rocket motion, and is derived by applying Newton’s laws to a rocket that exhausts its burnt fuel. Figure 1.9 shows the scenario. At some initial time t , the rocket has mass m and velocity v . At a time dt later, the rocket’s mass *must* be written as $m + dm$, and its velocity *must* be written as $v + dv$. Conservation of mass then says that the mass of the exhaust is $-dm$ (which is certainly positive). The exhaust velocity is V in the frame of the figure, in which the rocket is moving. This exhaust velocity is represented by an arrow with a curved end, and the reason for this curve is as follows. The tails of all the velocity arrows are drawn with a pen that starts moving to the right, which indicates that each velocity is taken as positive to the

right: like any scenario in classical mechanics, we use a single convention of a positive direction for displacement (which is then inherited by velocity and acceleration). The head of each velocity arrow shows the actual direction of motion of the object: to the right for the rocket, and to the left for the exhaust. (The initial direction of the exhaust vector's tail shows that V is positive for right motion. It follows that in our scenario, V is negative. If you are not convinced of the value of the curved tail on the exhaust velocity vector, ask yourself: if that velocity vector were a simple arrow " $\leftarrow$ " that pointed left and was labelled V , would you say that V was positive, or negative?)

The rocket accelerates because the right-moving particles of the expanding hot gas produced by burning fuel in the combustion chamber push the rocket to the right, while the left-moving particles of the expanding hot gas have nothing to push on, since the left end of the combustion chamber is open. This left-moving gas becomes the exhaust that is left behind. No force acts on the entire system of rocket plus exhaust, and so the rocket's momentum before the expulsion of the exhaust must equal the total momentum of rocket plus exhaust after the expulsion:

$$mv = (m + dm)(v + dv) + -dm \times V. \quad (1.140)$$

We can expand this, remembering from the discussion just after (1.133) that we need only write the leading-order terms in the infinitesimals:

$$mv = mv + m dv + v dm - V dm. \quad (1.141)$$

This simplifies to

$$m dv = (V - v) dm. \quad (1.142)$$

The speed of the exhaust *relative to the rocket* is determined by the fuel used, and is a (positive) number supplied by engineers: call this v_{ex} . Note that the *velocity* of the exhaust relative to the rocket equals the velocity of the exhaust in the frame of the figure (V) minus the velocity of the rocket in the frame of the figure (v): hence the exhaust velocity relative to the rocket equals $V - v$. This is negative (remember that V is negative), and so the exhaust *speed* equals minus this: $v_{\text{ex}} = v - V$. Equation (1.142) becomes

$$m dv = -v_{\text{ex}} dm. \quad (1.143)$$

This differential equation is easily integrated to give the well-known *rocket equation*, which tells us the expected velocity boost $v(t) - v(0)$ resulting from burning fuel whose mass is, of course, the starting mass $m(0)$ minus the current mass $m(t)$:

$$v(t) - v(0) = v_{\text{ex}} \ln \frac{m(0)}{m(t)}. \quad (1.144)$$

In contrast to the above analysis, most expositions of the rocket equation say that from time t to $t + dt$, the rocket exhausts a mass dm . Then, they

either say that the rocket's mass drops from m to $m - dm$ in this time dt , or that it drops from $m + dm$ to m . These assignments break the single, simple rule for infinitesimals, which says that from time t to $t + dt$, any quantity x evolves to become $x + dx$. Hence, the rocket's mass must be m at time t : it cannot be $m + dm$; and neither can it drop from m to $m - dm$. Recall the discussion just after (1.125): writing " $m - dm$ " is a sure flag that the subsequent analysis will either require a fix-up minus sign, or just be completely inconsistent. If, and only if, you assign the various quantities as in Figure 1.9, will the resulting analysis be self consistent. Naturally, expositions that don't assign the infinitesimals correctly are always contrived to produce the correct rocket equation; but they have no value as exercises in mathematics. It's straightforward to assign your infinitesimals correctly from the outset, and doing so will always create a firm foundation for your analyses.

1.6.1 Basis Vectors

Differential geometry has always used the time-honoured infinitesimals as we have used them above, and they form part of a wider language involving vectors.¹¹ In particular, one application of infinitesimals that is useful in many areas is the construction of the *basis vectors* that represent some set of coordinates. We will have occasion to use basis vectors in the coming chapters. The most well-known basis vectors are the cartesian ones, often called $\mathbf{i}, \mathbf{j}, \mathbf{k}$, but more usefully called $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$ —or $\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z$ when we want to emphasise their property of having unit length. (The $\mathbf{i}, \mathbf{j}, \mathbf{k}$ notation actually predates vectors; it springs from Hamilton's invention of *quaternions*, which were designed to be an extension of complex numbers.) These cartesian basis vectors are identical no matter where they are drawn: $\mathbf{e}_x$ always points in the positive- x direction and has unit length, and similarly for $\mathbf{e}_y$ and $\mathbf{e}_z$. The $\mathbf{e}_x$ and $\mathbf{e}_y$ vectors are shown in Figure 1.10.

More generally, basis vectors $\mathbf{e}_{q_1}, \mathbf{e}_{q_2}, \dots$ can be defined at any point P for any set of coordinates $q_1, q_2, \dots$. To draw the basis vector $\mathbf{e}_q$ (corresponding to coordinate q) at point P , refer again to Figure 1.10. First, draw the position vector $\mathbf{s}(q)$ from the origin out to P . Then, draw the nearby position vector $\mathbf{s} + d\mathbf{s} \equiv \mathbf{s}(q + dq)$, which corresponds to increasing q to $q + dq$. Now define

¹¹ Differential geometry also has a modern, abstract branch that defines quantities called *forms*. The simplest of these, *one-forms*, are on a par with vectors, and obey linear rules similar to those governing vectors and infinitesimals. The creation of one-forms has led to an idea that they supply a needed rigor to infinitesimals. I think that idea is mistaken; you will find all the necessary and sufficient rigor of infinitesimals in (1.131)–(1.133), without needing one-forms. Even on their home soil of general relativity, I have never found an application where forms are necessary, simplifying, or elegant. All of their advertised uses in physics that I have ever seen can be handled more simply with vectors only, in shorter time, with far more elegance, and with less mathematical effort and manipulation.

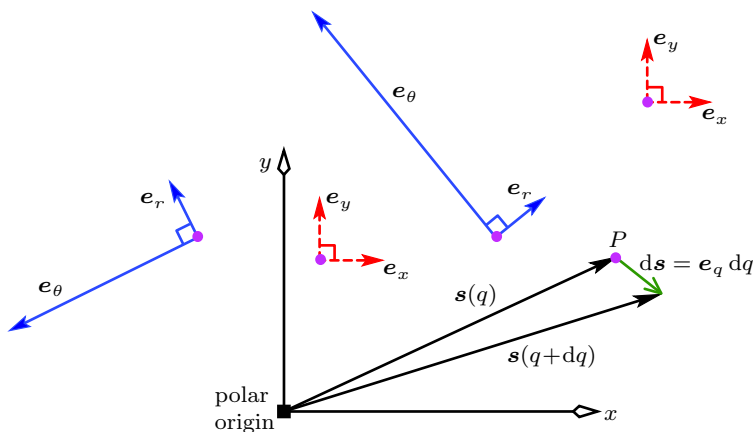


Fig. 1.10 Examples of basis vectors in two dimensions. Each pair is sited at the relevant black dot. The dashed vectors are the cartesian set e_x, e_y , which look the same everywhere. The full vectors are the polar set e_r, e_θ , whose length or direction depend on where they are. The construction of e_q at point P is also shown, for a general coordinate q

the basis vector as¹²

$$e_q \equiv \frac{ds}{dq} = \frac{s(q + dq) - s(q)}{dq}. \quad (1.145)$$

You can easily check that this prescription gives the usual cartesian basis vectors e_x, e_y, e_z . Figure 1.10 also shows polar basis vectors e_r, e_θ . These depend on the location (r, θ) of P : the radial basis vector e_r always points radially from the polar origin and has unit length, and the transverse basis vector e_θ always points at right angles to e_r (right-handed around the z axis), and has length r . Basis vectors can be normalised to useful effect, with various notations such as:

$$u_q \equiv \hat{e}_q \equiv e_{\hat{q}} \equiv e_q / |e_q|. \quad (1.146)$$

When basis vectors are used in this book with the notations “ $e_{\text{coordinate}}$ ” and “ $u_{\text{coordinate}}$ ” such as in (1.173) and Section 9.6, they have been defined in the standard way using (1.145) and (1.146).

¹² Some differential geometers omit the s in (1.145) and write e_q as “ $\partial/\partial q$ ”, and then redefine a basis vector to be this partial-derivative *operator*. Redefining a vector in this way from being an arrow (a very geometrical object) to becoming a decidedly non-geometrical object does not strike me as keeping the “geometry” in “differential geometry”!

1.6.2 The Probability Density

Earlier, we approximated the binomial probability distribution (1.66) by the gaussian distribution (1.72), because the gaussian was simpler to evaluate. But the same idea of replacing discrete with continuous is common throughout statistical mechanics, not so much because of any numerical advantage, but instead because the continuous case is usually easier to analyse.

Just how to define the probability of a continuous random variable is a refined concept, yet one that we are already familiar with when dealing with physical properties of everyday objects. In particular, suppose you have an “ideal” analogue 12-hour clock with a single hand: this hand moves not in tiny discrete steps, but instead completely smoothly, so that it points to some time t whose value is part of a continuum. If you observe this clock at some random moment, what is the probability that the time indicated is *exactly* “2:00”? The answer is zero, because the clock has an infinite number of possible times that it can display: since the probability is spread uniformly over all these readings, its value must be less than any number that we can nominate. If this seems nonintuitive, imagine that the clock display was generated by a computer using a random-number generator. The reading would have to be specified digitally, using, say, base 10. What is the chance that the computer chooses exactly “2.000...”? Aside from the fact that no computer is able to print out an infinite number of decimal places, and that we don’t have the infinite amount of time needed to check each of the digits, consider that the computer must generate each digit in turn with an algorithm that returns any digit with a probability of $1/10$. The chance of returning “2” is $1/10$. The chance of returning “2.0” is $(1/10)^2$. The chance of returning “2.00” is $(1/10)^3$, and so on. The chance of returning ever-finer approximations to exactly 2 tends toward zero in the limit of the infinitely long decimal representation required for the time shown on the clock.

This example shows that continuous random variables cannot be treated in the same way as discrete ones. The situation is just like the case of an idealised ruler whose mass is a truly continuous function of its length. (We will discuss real rulers again in Section 2.5.) The ruler’s mass from its zero-notch to a distance x from that end is some $M(x)$. What is the mass “at” $x = 2$? We know that the mass in between x and $x + \Delta x$ gets smaller as $\Delta x \rightarrow 0$, and so we can only say that the mass “at” any point is exactly zero. Where, then, is the mass, if it’s not at any point we’re able to name?

Here, we must use the idea of a mass *density*: the *linear mass density* $\lambda(x)$ at x is the mass *per unit length* at x , a well-defined quantity found by dividing the mass of some length in the vicinity of x by that length, and taking the limit as that length goes to zero. In the language of infinitesimals,

$$\lambda(x) = \frac{M(x + dx) - M(x)}{dx} = M'(x). \quad (1.147)$$

We can use this density to find approximately how much mass is located in some small length Δx of the ruler near x : this amount is $\Delta M \simeq \lambda(x) \Delta x$.

The spread of probability of a continuous random variable is analogous to the spread of mass of the above ruler: the chance of the clock displaying a time of “ $t = 2$ ” is zero, just as the amount of mass at $x = 2$ is zero. So, we must appeal to the idea of a *probability density*. For the clock, write the chance of finding a displayed time anywhere from 0 to some t as $P(t)$. This is a *cumulative probability*, which is well defined at any value of t (that the clock can show, of course). By definition of the cumulative probability, the chance of finding a displayed value anywhere from t to $t + \Delta t$ is $P(t + \Delta t) - P(t)$. The probability density of the displayed value being t is then defined as

$$\text{probability density } p(t) \equiv \lim_{\Delta t \rightarrow 0} \frac{P(t + \Delta t) - P(t)}{\Delta t} = P'(t). \quad (1.148)$$

Probability density is usually written as a lower-case “ p ”—so $p(t)$ in this case—but we must be aware that this is a *density*: in this case, it has dimensions of “1/time”.

The clock’s cumulative probability is clearly $P(t) = t/12$. The probability density is thus

$$p(t) = P'(t) = 1/12. \quad (1.149)$$

This is non-zero at any t . Contrast it with the probability of finding a reading of t , which is trivially

$$\text{chance of displaying “}t\text{”} = \lim_{\Delta t \rightarrow 0} P(t + \Delta t) - P(t) = P(t) - P(t) = 0. \quad (1.150)$$

To speak meaningfully of a continuous random variable, we must use probability *density*.

What is the expected value $\langle t \rangle$ of the clock display? If the display were confined to *discrete* values of t as $0, 1, \dots, 11$, then each value would have probability $1/12$, and its expected value would be, from (1.38),¹³

$$\langle t \rangle = 0 \times \frac{1}{12} + 1 \times \frac{1}{12} + \dots + 11 \times \frac{1}{12} = 5.5. \quad (1.151)$$

Or perhaps we might replace the top hour’s value $t = 0$ with $t = 12$: if the reading were confined to t values of $1, 2, \dots, 12$, then

$$\langle t \rangle = 1 \times \frac{1}{12} + 2 \times \frac{1}{12} + \dots + 12 \times \frac{1}{12} = 6.5. \quad (1.152)$$

Now, what about the continuous case? Imagine “coarse graining” the continuous case to a discrete approximation: envisage displayed values “ t ” that are spaced by some Δt . If we allow $t = 0$ and not $t = 12$, these readings are

¹³ I have applied the easily proved identity $1 + 2 + \dots + n = n(n + 1)/2$ in these equations.

$$\text{possible readings } t = 0, \Delta t, 2\Delta t, 3\Delta t, \dots, 12 - \Delta t. \quad (1.153)$$

The chance of displaying any given t in this list is spread uniformly, and so is “one divided by the number of values in the list”. How many values are there? Count them by ignoring the factor of Δt : so, divide all values by Δt to produce a corresponding list:

$$0, 1, 2, 3, \dots, 12/\Delta t - 1. \quad (1.154)$$

This clearly has $12/\Delta t$ elements, and so the chance of our finding a displayed time of any given t in the list (1.153) is one divided by $12/\Delta t$, or $\Delta t/12$. The expected value of the possible readings (1.153) is then

$$\begin{aligned} \langle t \rangle &= 0 \times \frac{\Delta t}{12} + \Delta t \times \frac{\Delta t}{12} + 2\Delta t \times \frac{\Delta t}{12} + \dots + (12 - \Delta t) \times \frac{\Delta t}{12} \\ &= \frac{\Delta t^2}{12} \left[1 + 2 + 3 + \dots + \frac{12}{\Delta t} - 1 \right] = \frac{\Delta t^2}{12 \times 2} \left(\frac{12}{\Delta t} - 1 \right) \frac{12}{\Delta t} \\ &= 6 - \Delta t/2. \end{aligned} \quad (1.155)$$

Note, in passing, that if we had replaced $t = 0$ with $t = 12$, almost the same calculation would produce $\langle t \rangle = 6 + \Delta t/2$. This means that for the two choices of the top hour being displayed as 0 or 12, both of these expected values tend toward 6 as the set of possible readings becomes ever more finely grained.¹⁴

But the same result for the limit of a continuous probability can still be found much more simply by applying (1.38). The probability of finding any particular t displayed equals the probability density times an infinitesimal displayed-time interval dt . Thus, in (1.38), we replace p_i with $p(t) dt$. The sum in that equation becomes an integral:

$$\langle t \rangle = \int_0^{12} t p(t) dt \stackrel{(1.149)}{=} \int_0^{12} \frac{t}{12} dt = \left[\frac{t^2}{24} \right]_0^{12} = 6, \quad (1.156)$$

as expected. Memorise the continuous version of (1.38):

$$\langle x \rangle = \int x p(x) dx, \quad (1.157)$$

where $p(x)$ is a probability *density*; also, replace a discrete probability with a density times an infinitesimal in the continuous case. Hence, the probability of finding exactly some given x “is zero” when x is continuous, but “is zero”

¹⁴ Is it paradoxical that in the continuum limit, $\langle t \rangle$ doesn’t depend on whether the top hour is 0 or 12? Don’t forget that in the limit of a continuum of displayable times, the chance of finding either exactly 0 or exactly 12 is *zero*, and hence the precise value of the top hour doesn’t contribute to $\langle t \rangle$!

really means “ $p(x)dx$ ”. And $p(x)dx$ is the appropriate probability to use when aggregating an infinite number of values of x , such as when integrating over x .

1.7 Exercising Care with Partial Derivatives

The language of statistical mechanics uses partial derivatives liberally, so if you spend some time exploring partial-derivative notation, later analyses will become more transparent.

We can begin such exploring by taking a function $f(x)$, and applying Taylor’s theorem to it:

$$f(x + \Delta x) = f(x) + f'(x) \Delta x + f''(x) \Delta x^2/2! + \dots \quad (1.158)$$

How do we Taylor-expand a function of *two* variables, $f(x, y)$? We can apply a Taylor expansion for each variable in turn, by first treating $f(x, y)$ as a function of x with y fixed, and then treating each resulting expression as a function of y with x fixed. First, define the partial-derivative notation

$$\begin{aligned} f_x &\equiv \frac{\partial f}{\partial x} \equiv \frac{df}{dx} \text{ when } y \text{ is held fixed,} \\ f_{xx} &\equiv \frac{\partial^2 f}{\partial x^2} \equiv \frac{d^2 f}{dx^2} \text{ when } y \text{ is held fixed,} \\ f_{xy} &\equiv \frac{\partial^2 f}{\partial y \partial x} \equiv \frac{d^2 f_x}{dy} \text{ when } x \text{ is held fixed, and so on.} \end{aligned} \quad (1.159)$$

In contrast to the fact that df/dx *can* be treated as the fraction “ df divided by dx ”, the same is not true for $\partial f/\partial x$: the symbols ∂f and ∂x are not defined in isolation. On rare occasions, you might encounter an expression such as “ $x + \partial x$ ”; but this has no meaning: whoever wrote it almost certainly meant to write $x + \Delta x$.

Voicing Derivatives and Some Other Things

Textbooks seldom give direction on how to pronounce mathematical language. As a result, you will encounter various ways of saying df/dx , d^2f/dx^2 , $\partial f/\partial x$, and $\partial^2 f/\partial x^2$. Because I treat an ordinary derivative as a fraction, df/dx becomes “d f on d x”. The second derivative d^2f/dx^2 is “d two f on d x squared” (not “d squared f on d x squared”). This leaves room for “d f d x” to denote $df dx$ (thus, “d x d y” is the area element $dx dy$). In contrast, I pronounce the partial derivative $\partial f/\partial x$ as simply “del f del x”, because it is *not* a fraction “del f divided by del x”, since—as

stated above— ∂f and ∂x are not defined in isolation. $\partial^2 f / \partial x^2$ becomes “del two f del x squared”.

You will frequently hear the vector partial-derivative operator ∇ pronounced as “del”. But this operator has its own name, “nabla”, when used in isolation, and so it makes sense to let “del” refer only to ∂ , which is a sort of modified delta. When nabla (∇) operates on a scalar function f , or on a vector function $\mathbf{v}$, the expressions ∇f , $\nabla \cdot \mathbf{v}$, and $\nabla \times \mathbf{v}$ are usually pronounced “grad f”, “div v”, and “curl v”, respectively.

And on a note of enunciative elegance, try pronouncing e^x not as “e to the x” or “e to the power of x”, but instead as “e raised to x”.

For any well-behaved function f , the order of partial differentiation is immaterial: $f_{xy} = f_{yx}$. Now Taylor-expand $f(x, y)$ in x , holding y fixed:

$$\begin{aligned} f(x + \Delta x, y + \Delta y) &= f(x, y + \Delta y) + f_x(x, y + \Delta y) \Delta x \\ &\quad + f_{xx}(x, y + \Delta y) \Delta x^2 / 2! + \dots \end{aligned} \quad (1.160)$$

Now expand each of these terms in y , holding x fixed. We will write all terms up to second order, and will drop all mention of “ (x, y) ” on the partial derivatives:

$$\begin{aligned} f(x + \Delta x, y + \Delta y) &= f(x, y) + f_y \Delta y + f_{yy} \Delta y^2 / 2! + \dots \\ &\quad + f_x \Delta x + f_{xy} \Delta y \Delta x + \dots \\ &\quad + f_{xx} \Delta x^2 / 2! + \dots \\ &= f(x, y) + f_x \Delta x + f_y \Delta y \\ &\quad + f_{xx} \Delta x^2 / 2! + f_{xy} \Delta x \Delta y + f_{yy} \Delta y^2 / 2! + \dots \end{aligned} \quad (1.161)$$

We can retain terms to first order only, writing either the approximation

$$f(x + \Delta x, y + \Delta y) \simeq f(x, y) + f_x \Delta x + f_y \Delta y, \quad (1.162)$$

or the *exact* expression involving infinitesimals:

$$f + df = f(x + dx, y + dy) = f(x, y) + f_x dx + f_y dy. \quad (1.163)$$

This last expression shortens to

$$df = f_x dx + f_y dy. \quad (1.164)$$

If f is a function of three variables x, y, z , then a similar analysis shows that

$$df = f_x dx + f_y dy + f_z dz, \quad (1.165)$$

and so on for any other number of variables. It follows that an infinitesimal expression such as

$$dF = A dX + B dY + C dZ \quad (1.166)$$

is equivalent to writing

$$\frac{\partial F}{\partial X} = A, \quad \frac{\partial F}{\partial Y} = B, \quad \frac{\partial F}{\partial Z} = C. \quad (1.167)$$

We'll make extensive use of this idea in the pages to come.

The Nabla Operator ∇

On a point of notation, return briefly to (1.164) and (1.165). Suppose we define a vector

$$\begin{aligned} \nabla f &= (f_x, f_y) && \text{in two dimensions, and} \\ \nabla f &= (f_x, f_y, f_z) && \text{in three dimensions.} \end{aligned} \quad (1.168)$$

We also denote

$$\begin{aligned} d\mathbf{x} &= (dx, dy) && \text{in two dimensions, and} \\ d\mathbf{x} &= (dx, dy, dz) && \text{in three dimensions.} \end{aligned} \quad (1.169)$$

Then, both (1.164) and (1.165) can be written as

$$df = \nabla f \cdot d\mathbf{x}. \quad (1.170)$$

The generic name of ∇ is “nabla”. The vector ∇f is called the *gradient* of f , or “grad f ”. What is the meaning of ∇f ? To gain a feel for it, set $f(\mathbf{x})$ to be the temperature $T(\mathbf{x})$ in a room. As we make an infinitesimal step $d\mathbf{x}$ anywhere in space, the increase in temperature dT that we experience along the step is given by $dT = \nabla T \cdot d\mathbf{x}$. Now, recalling the relationship of a dot product to the cosine, and writing the angle between ∇T and $d\mathbf{x}$ as “ $(\nabla T, d\mathbf{x})$ ”, we observe that

$$dT = \nabla T \cdot d\mathbf{x} = |\nabla T| \times \text{step length} |\mathbf{dx}| \times \cos(\nabla T, d\mathbf{x}). \quad (1.171)$$

Suppose we start from some point $\mathbf{x}$ and take “probing steps” $d\mathbf{x}$ in various directions, always returning to $\mathbf{x}$ before taking a new probing step. Then, (1.171) says that dT is maximised when $\cos(\nabla T, d\mathbf{x})$ is maximised, which occurs when the cosine equals one, meaning the angle $(\nabla T, d\mathbf{x})$ equals zero—which happens when we take our step $d\mathbf{x}$ in the direction of ∇T . In other words, ∇T points in the direction in which T increases most rapidly. If a candle were placed somewhere in a cold room, at each point in the room, the vector ∇T would point toward the candle, and its length would be

$$|\nabla T| = dT/|d\mathbf{x}| \quad (1.172)$$

for the temperature increase dT occurring along a step $d\mathbf{x}$ heading directly toward the candle. We see that ∇f is the spatial rate of increase of $f(\mathbf{x})$, and it points in the direction in which $f(\mathbf{x})$ increases most rapidly. We'll use this property of ∇T to study heat flow in Section 4.1.2.

On an advanced note (something we won't use in this book), what is the generalisation of (1.168) to *any* coordinates that need not necessarily be orthogonal? It can be shown that for completely general coordinates,

$$\nabla = \sum_{\alpha\beta} g^{\alpha\beta} \mathbf{e}_\beta \partial_\alpha, \quad (1.173)$$

where $g^{\alpha\beta}$ is the $\alpha\beta^{\text{th}}$ element of the *inverse* of the matrix whose $\alpha\beta^{\text{th}}$ element is $\mathbf{e}_\alpha \cdot \mathbf{e}_\beta$, and the set $\{\mathbf{e}_\alpha\}$ comprises the basis vectors corresponding to the coordinates α , as defined in (1.145). These basis vectors are generally not orthogonal and not of unit length.

Two alternative notations are commonly used for the partial derivative. Given a function $f(x, y, z)$, these two expressions have the same meaning:

$$\frac{\partial f(x, y, z)}{\partial x} \quad \text{and} \quad \left(\frac{\partial f}{\partial x} \right)_{y,z}. \quad (1.174)$$

They both denote the rate of increase of f with respect to x when y and z are held fixed. That is, when y and z are held fixed, f becomes a function only of x , and this rate of increase is then df/dx . Unless otherwise indicated, the expression $\partial f/\partial x$ with no parentheses is understood to mean $(\partial f/\partial x)_{y,z}$. Of course, this assumes that everyone is aware of the variables y and z . Whenever you encounter an expression such as “ $\partial P/\partial S$ ”, be sure to know precisely what other variables P is a function of, besides S .

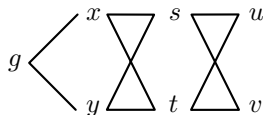
The very standard parenthesis notation of (1.174) might be a little confusing, so it's worth emphasising that the large parentheses at the right in (1.174) do not mean that “ $(\partial f/\partial x)_{y,z}$ ” is some kind of special instance of something called “ $\partial f/\partial x$ ”. The parentheses do not “operate” on $\partial f/\partial x$; they merely reinforce that y and z are held fixed. The notation “ $(df/dx)_{y,z}$ ” would serve just as well as $(\partial f/\partial x)_{y,z}$ and, in fact, is probably better: we should perhaps write

$$\frac{\partial f(x, y, z)}{\partial x} \equiv \left(\frac{df}{dx} \right)_{\text{fixed } y, z} \quad (1.175)$$

The main thing to remember is that whenever you use “ ∂ ”, always be aware of which variables are being held fixed in the process of differentiating. If the full set of variables on which a function f depends is not known, then the expression “ $\partial f/\partial x$ ” has no meaning. And if those variables *are* known, then

the parentheses and subscripted variables are technically superfluous. Their sole use is to remind us as to which variables are being held fixed.

More generally, suppose that a function g depends on x, y , which both depend on s, t , which both depend on u, v . This chain of dependencies is conventionally written as



This picture helps us to keep track of the variables when applying the *chain rule* of partial differentiation. To calculate $\partial g/\partial u$, for example, we start at g in the above network and follow all paths to u , to obtain

$$\frac{\partial g}{\partial u} = \frac{\partial g}{\partial x} \frac{\partial x}{\partial s} \frac{\partial s}{\partial u} + \frac{\partial g}{\partial x} \frac{\partial x}{\partial t} \frac{\partial t}{\partial u} + \frac{\partial g}{\partial y} \frac{\partial y}{\partial s} \frac{\partial s}{\partial u} + \frac{\partial g}{\partial y} \frac{\partial y}{\partial t} \frac{\partial t}{\partial u}. \quad (1.176)$$

It would be tedious to have to include parentheses everywhere in (1.176) to denote the variables that are being held fixed; instead, the presence of these parentheses and the relevant variables is implicitly understood.

An ordinary derivative behaves notationally like a fraction, meaning that $dy/dx = 1/(dx/dy)$. The corresponding idea is more complicated for partial derivatives. The first thing we can say is that, because the expression $(\partial f/\partial x)_{y,z}$ means df/dx when y, z are held fixed, it follows trivially that¹⁵

$$(\partial x/\partial f)_{y,z} = \frac{1}{(\partial f/\partial x)_{y,z}}. \quad (1.177)$$

But usually, when we swap the roles of, say, f and x , the set of variables that are being held fixed actually changes, and so a simple reciprocation of the notation cannot be used. Even so, a more familiar example of relating polar coordinates to cartesians demonstrates just how “ $dy/dx = 1/(dx/dy)$ ” becomes modified for partial derivatives. Begin with

$$x = r \cos \theta, \quad y = r \sin \theta. \quad (1.178)$$

Again, when we write $\partial x/\partial r$, we really mean $(\partial x/\partial r)_\theta$: we are differentiating with respect to one variable (r), while holding fixed all others of its family (θ). The set of partial derivatives of one set of coordinates with respect to the other can be written as the elements of a matrix known as the *jacobian matrix* of the coordinate transform:

¹⁵ Indeed, (1.177) is very clear if we use the notation of (1.175). It's unfortunate that the notation of (1.175) is not commonly found in the subject of partial differentiation.

$$\begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix}. \quad (1.179)$$

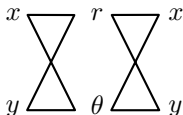
Two jacobian matrices relate cartesian and polar coordinates: one has the partial derivatives of cartesians with respect to polars, and the other has the partial derivatives of polars with respect to cartesians. Now notice what happens when these two jacobian matrices are multiplied to form a matrix A :

$$A = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix} \begin{bmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{bmatrix}. \quad (1.180)$$

The top-left element of A is

$$A_{11} = \frac{\partial x}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial x}{\partial \theta} \frac{\partial \theta}{\partial x}. \quad (1.181)$$

This is the chain rule applied to the following dependencies:



It follows that $A_{11} = \partial x / \partial x = 1$, because $\partial x / \partial x$ means the rate of increase of x with respect to x , holding y fixed. Similarly,

$$A_{12} = \frac{\partial x}{\partial r} \frac{\partial r}{\partial y} + \frac{\partial x}{\partial \theta} \frac{\partial \theta}{\partial y} = \frac{\partial x}{\partial y} = 0, \quad (1.182)$$

because $\partial x / \partial y$ means the rate of increase of x with respect to y , holding y fixed: and if y is fixed, then x cannot change. The other two elements of A follow just as easily, and the result is that A is just the identity matrix. This is a very useful result that holds quite generally: the two jacobian matrices are inverses of each other, making this matrix inversion the extension of “ $dy/dx = 1/(dx/dy)$ ” to partial derivatives. For polar coordinates, we can immediately write

$$\begin{bmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix}^{-1} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\frac{\sin \theta}{r} & \frac{\cos \theta}{r} \end{bmatrix}. \quad (1.183)$$

We see here how to invert partial derivatives when the set of variables being held fixed switches from one set of coordinates to the other. For example, compare the “1,1” elements of (1.179) and (1.183) to infer

$$\left(\frac{\partial x}{\partial r}\right)_\theta \stackrel{(1.179)}{=} \cos \theta, \quad \text{and} \quad \left(\frac{\partial r}{\partial x}\right)_y \stackrel{(1.183)}{=} \cos \theta. \quad (1.184)$$

With the above convention of omitting the fixed variables in mind, (1.184) is normally written as

$$\frac{\partial x}{\partial r} = \frac{\partial r}{\partial x} = \cos \theta. \quad (1.185)$$

This might at first look a little odd—until we realise that each derivative assumes that a different variable is held fixed, and so the simple reciprocation of (1.177) cannot be used to relate them.

Partial-Derivative Gymnastics

To demonstrate the straightforward reciprocation of (1.177), we show that $(\partial x/\partial r)_\theta = 1/(\partial r/\partial x)_\theta$. Start with (1.178) and write $r = x/\cos \theta$. Then, conclude that

$$\left(\frac{\partial r}{\partial x}\right)_\theta = \frac{1}{\cos \theta}. \quad (1.186)$$

But writing $x = r \cos \theta$ makes it clear that $(\partial x/\partial r)_\theta = \cos \theta$; and hence (1.186) becomes

$$\left(\frac{\partial r}{\partial x}\right)_\theta = \frac{1}{\cos \theta} = \frac{1}{(\partial x/\partial r)_\theta}. \quad (\text{QED}) \quad (1.187)$$

How might we calculate an expression such as $(\partial \theta/\partial r)_y$? (Remember that “ $\partial \theta/\partial r$ ” with no enclosing parentheses is understood to mean $(\partial \theta/\partial r)_\theta$, which is zero. But in the expression $(\partial \theta/\partial r)_y$, it is y rather than θ that is being held fixed.) The easiest approach is to differentiate both sides of $y = r \sin \theta$ with respect to r while holding y fixed, to get

$$0 = \sin \theta + r \cos \theta (\partial \theta/\partial r)_y. \quad (1.188)$$

This rearranges to yield the required result:

$$\left(\frac{\partial \theta}{\partial r}\right)_y = \frac{-\sin \theta}{r \cos \theta}. \quad (1.189)$$

For a slight variation, calculate $(\partial r/\partial \theta)_y$ in the same way and find its reciprocal. Here is a third way, which might give you more insight. Draw two infinitesimally separated points at a fixed y , as shown in Figure 1.11. One point has polar coordinates (r, θ) ; the other has $(r + dr, \theta + d\theta)$.

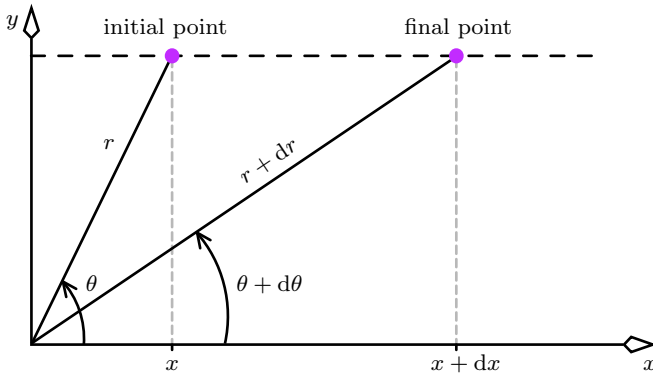


Fig. 1.11 To calculate the rate of increase of r with respect to θ with y held fixed, consider two infinitesimally separated points at fixed y (drawn well separated here for clarity). Then calculate dr and $d\theta$ when moving from the initial to the final point. Which point is chosen as the “initial” one is immaterial

Noting that we need keep only the lowest powers necessary of the infinitesimals, write

$$\begin{aligned} y &= r \sin \theta = (r + dr) \sin(\theta + d\theta) \\ &= (r + dr) (\sin \theta + \cos \theta d\theta) \\ &= r \sin \theta + \sin \theta dr + r \cos \theta d\theta. \end{aligned} \quad (1.190)$$

It follows that the sum of the last two terms on the last line is zero:

$$\sin \theta dr + r \cos \theta d\theta = 0. \quad (1.191)$$

Then, since this was all done at fixed y ,

$$\left(\frac{\partial \theta}{\partial r} \right)_y = \frac{d\theta}{dr} \text{ from (1.191)} = \frac{-\sin \theta}{r \cos \theta}. \quad (1.192)$$

Note that although we seemed to work to first order only, the result is *exact*. If this seems strange, revisit (1.131)–(1.133) in the more familiar language of $f(x)$ in one dimension.

1.8 Exact and Inexact Differentials

The infinitesimals of Section 1.6, such as dV in (1.138), are the result of an implied limit procedure in which the corresponding non-infinitesimal quantity

(ΔV in that case) is understood to tend toward zero. These infinitesimals relate to a variable that describes the state of a system: in this case, the system has some volume V . They are often called *exact differentials*. (The nouns “infinitesimal” and “differential” are interchangeable; but “exact differential” is commonly preferred over “exact infinitesimal”.)

The Exact Differential

An infinitesimal amount of some quantity x is called an *exact differential*, denoted dx , if every state of a system corresponds to a single value of x . The quantity x is called a *state variable*.

Exact differentials are the infinitesimals that we have described up until now. In contrast, a less formal type of infinitesimal is the *inexact differential*, written here with a “ $\mathfrak{d}$ ” and meaning “a small amount of the following quantity that is not a state variable”. For example, no system has a state variable called work W , and yet we *can* perform an infinitesimal amount of work $\mathfrak{d}W$ on a system.

A well-known example of an inexact differential is the small area element used to discuss Gauss’s theorem in electromagnetism. This is usually called dA , but that “ d ” is not meant to imply that we are considering an increase in the area A of the relevant surface, because no “area variable” A is defined at each point on the surface. Hence, a small surface-area element is an inexact differential, and is better written as $\mathfrak{d}A$. Likewise, the force due to air pressure that acts on a small surface of area $\mathfrak{d}A$ can be written as $\mathfrak{d}F$, despite more commonly being written as dF . It’s rather odd that the use of $\mathfrak{d}$ seems to be confined to statistical mechanics, when it could be employed far more widely.

For a more involved example of exact and inexact differentials, suppose that two hikers “1” and “2” walk from Adelaide to Melbourne. They follow different paths and meet at some point en route. This point has a particular height h above sea level. The position of each hiker always has a unique value of h associated with it; hence h is a state variable (here, “state” denotes the current position of a hiker), and dh is then an exact differential. We can write the ground’s height $h(\lambda, \phi)$ as a function of latitude λ and longitude ϕ . Each hiker’s total increase in height above sea level Δh in walking from Adelaide to the point (λ, ϕ) is independent of the path that each followed to arrive at their meeting point:

$$\Delta h = \int_{\text{Adelaide}}^{(\lambda, \phi)} dh = \text{height at } (\lambda, \phi) - \text{height at Adelaide.} \quad (1.193)$$

The hikers’ meeting point has a unique latitude λ and longitude ϕ , and so latitude and longitude are also state variables; and that means $d\lambda$ and $d\phi$ are exact differentials.

Contrast the hikers' common height h at their meeting point with the distances s_1 and s_2 that they each have walked to a given point (λ, ϕ) . Distance s_1 depends on the path taken by hiker 1, and similarly for distance s_2 . Each of these distances is defined on, and only on, the relevant hiker's chosen path. We can then certainly define distances Δs_1 and Δs_2 covered between two points A and B :

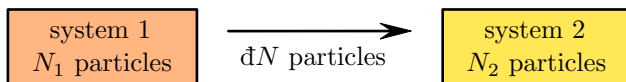
$$\Delta s_1 = \int_{A, \text{ path 1}}^B ds_1 \quad \text{and} \quad \Delta s_2 = \int_{A, \text{ path 2}}^B ds_2, \quad (1.194)$$

because “distance s_1 from Adelaide” is a function of position along the path that hiker 1 takes, and similarly for hiker 2. But being path dependent, that distance is not a function solely of any general position in the country. It follows that the hikers' common position at their meeting point cannot be associated with any single-valued variable called “distance walked s ”. But the distance traversed in an infinitesimal step taken by a hiker *can* be written as an inexact differential $\text{d}s$. The total distance that any generic hiker covers from point A to point B can be written as

$$s = \int_A^B \text{d}s, \quad (1.195)$$

where the value of s is hiker dependent. Compare this with (1.194): notice that we don't call this total distance Δs . That is because s is not a state variable (where a hiker's “state” is his current position): whereas we do have a concept of an “initial s ” (which is zero for all hikers), we have no concept of a *final* s that is independent of any hiker and depends only on the meeting point (λ, ϕ) . Hence, Δs —meaning “final s minus initial s —has no place here.

An inexact differential can often be treated mathematically by converting it to an exact differential. A very simple example of this occurs in a scenario where some small number of particles is transferred from system 1 to system 2, which themselves have total particle numbers N_1 and N_2 , respectively. There is no state variable called N , so we start out by saying “An infinitesimal number of particles $\text{d}N$ is transferred from system 1 to system 2”, and then draw a picture:



System 1 loses $\text{d}N$ particles and system 2 gains $\text{d}N$ particles. It follows that the exact differentials $\text{d}N_1$, $\text{d}N_2$ are related to the inexact differential $\text{d}N$ by

$$\begin{aligned} -\text{d}N_1 &= \text{d}N : \text{ the loss in } N_1 \text{ equals } \text{d}N, \\ \text{d}N_2 &= \text{d}N : \text{ the gain in } N_2 \text{ equals } \text{d}N. \end{aligned} \quad (1.196)$$

Subsequent calculations can now dispense with $\mathfrak{d}N$ by replacing it with dN_1 and dN_2 . These latter two quantities might both, for example, then be integrated to give information on how the individual particle numbers N_1 and N_2 change.

A More Abstract Example of an Inexact-to-Exact Conversion

Here is another example of converting an inexact differential to an exact differential. Consider the differential $y dx + 2x dy$. We can show that this is inexact with a proof by contradiction. Suppose it was exact; then there would exist a function $f(x, y)$ such that df equalled $y dx + 2x dy$. But recalling (1.164), we know that

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy. \quad (1.197)$$

Comparing this with “ $df = y dx + 2x dy$ ”, we would conclude that

$$\frac{\partial f}{\partial x} = y, \quad \text{and} \quad \frac{\partial f}{\partial y} = 2x. \quad (1.198)$$

That means the mixed second partial derivatives of $f(x, y)$ would have to be

$$\begin{aligned} \frac{\partial^2 f}{\partial y \partial x} &\equiv \frac{\partial}{\partial y} \frac{\partial f}{\partial x} = \frac{\partial}{\partial y} y = 1, \\ \frac{\partial^2 f}{\partial x \partial y} &\equiv \frac{\partial}{\partial x} \frac{\partial f}{\partial y} = \frac{\partial}{\partial x} 2x = 2. \end{aligned} \quad (1.199)$$

But this is a contradiction, since these two mixed second partial derivatives of any well-behaved function must always be equal. We conclude that the function $f(x, y)$ does not exist. Nonetheless, we can certainly define the infinitesimal “ $\mathfrak{d}f$ ” such that

$$\mathfrak{d}f \equiv y dx + 2x dy. \quad (1.200)$$

Next, introduce the function

$$g(x, y) = xy^2. \quad (1.201)$$

The two mixed second partial derivatives of $g(x, y)$ must certainly be equal [because $g(x, y)$ is a well-behaved function]; but we can check anyway:

$$\frac{\partial^2 g}{\partial y \partial x} \equiv \frac{\partial}{\partial y} \frac{\partial g}{\partial x} = \frac{\partial}{\partial y} y^2 = 2y,$$

$$\frac{\partial^2 g}{\partial x \partial y} \equiv \frac{\partial}{\partial x} \frac{\partial g}{\partial y} = \frac{\partial}{\partial x} 2xy = 2y. \quad (1.202)$$

QED: these mixed second partials are equal, as expected. Now, because $g(x, y)$ is a well-behaved function, the exact differential dg exists:

$$\begin{aligned} dg &= \frac{\partial g}{\partial x} dx + \frac{\partial g}{\partial y} dy \\ &= y^2 dx + 2xy dy = y(y dx + 2x dy) \\ &= y df. \end{aligned} \quad (1.203)$$

Put another way, what we have done is multiply the inexact differential df by y , to produce an exact differential dg corresponding to a new function $g(x, y)$. Hopefully, studying $g(x, y)$ will shed light on whatever circumstances gave rise to the original expression df . This sort of “inexact-to-exact” conversion that multiplies the inexact differential by a well-chosen quantity will enable us to make analytical sense of the First Law of Thermodynamics, in Chapter 3.

1.9 Numerical Notation, Units, and Dimensions

In the chapters to come, we will frequently substitute numbers into long equations, and will use an abbreviated but efficient notation to do so. The first point of this notation is a simple device used when powers of ten appear:

Notational device: within a long equation, write $a \stackrel{b}{=}$ for $a \times 10^b$. But use the full form $a \times 10^b$ when writing the final answer.

This notation keeps numbers together that belong together, without using extraneous multiplication signs. For example, rather than write the product of Planck’s constant h and the speed of light c as

$$hc = 6.626 \times 10^{-34} \times 2.998 \times 10^8 \text{ SI units}, \quad (1.204)$$

we write

$$hc = 6.626 \stackrel{-34}{=} \times 2.998 \stackrel{8}{=} \text{ SI units}. \quad (1.205)$$

This is effectively the same concise notation found in many programming languages, that write

$$\text{hc} = 6.626\text{e-}34 * 2.998\text{e}8$$

The exponent stays with the number; thus, for example,

$$\begin{aligned} h/c &= 6.626 \overset{-34}{=} / 2.998 \overset{8}{=} \text{ SI units} \\ &\equiv 6.626 \overset{-34}{=} / (2.998 \overset{8}{=}) \text{ SI units.} \end{aligned} \quad (1.206)$$

The second part of our abbreviated notation concerns a choice of units. Typically, we perform all numerical calculations by stating that SI units are being used, and then writing the end-product SI unit only where it is necessary—thus writing all numbers *within* an expression without their units. Each statement of equality in the resulting mathematics remains strictly correct, and the calculation is unencumbered by internal units that would only cancel each other out anyway. The tidiness and minimalism that results is, after all, the whole point of using a consistent set of units.

To demonstrate, consider numerically evaluating an area given by the following equation:

$$A = \frac{h^2}{2\pi mkT}, \quad (1.207)$$

where m is the mass of a prototypical air molecule, k is Boltzmann's constant, T is room temperature, and h is Planck's constant. If we choose SI units, the standard format is

$$\begin{aligned} A &\simeq \frac{(6.626 \times 10^{-34} \text{ J s})^2}{2\pi \times 4.8 \times 10^{-26} \text{ kg} \times 1.381 \times 10^{-23} \text{ J/K} \times 298 \text{ K}} \\ &\simeq 3.5 \times 10^{-22} \text{ m}^2. \end{aligned} \quad (1.208)$$

This format is dominated by a long string of numbers that are connected by multiplication signs, and various SI units that we just don't need to know; after all, A is an area, so, if we use SI units, the result must have units of square metres. Here is the alternative format that we will use in this book: using SI units, we have

$$A \simeq \frac{(6.626 \overset{-34}{=})^2}{2\pi \times 4.8 \overset{-26}{=} \times 1.381 \overset{-23}{=} \times 298} \text{ m}^2 \simeq 3.5 \overset{-22}{=} \text{ m}^2. \quad (1.209)$$

The whole point of using a consistent set of units is that we don't need to know the individual units in the calculation; if we are using SI, then we can state with confidence that the final unit is simply “m²”. Also, every part of (1.209) is strictly correct as it stands, even if we don't state that we are going to use SI units to evaluate it. After all, even without mentioning any use of SI, both statements of (1.209) are true as they stand: the area A really does approximately equal $(6.626 \overset{-34}{=})^2$ [etc.] m², and that latter expression really does approximately equal 3.5×10^{-22} m².

Omit Units with Care!

Leaving intermediate units out of a calculation as we have done in (1.209) means that we are no longer “doing a dimensional analysis” each time we put numbers into a formula, and so is akin to removing the safety barriers from a drive around the cliffs. We must have confidence in the formula: confidence that we understand what each symbol really represents.

A classic situation when care is needed occurs in relativity. There, it is very conventional to define a “time” variable t and a “velocity” variable v as follows, where c is the inertial-frame speed of light through a vacuum:

$$t \equiv c \times \text{time}, \quad v \equiv \text{velocity}/c. \quad (1.210)$$

The variable t is conventionally called time, but its dimension is distance; v is conventionally called velocity, but it is dimensionless. This convention has several advantages for the subject:

- It makes the equations symmetrical regarding the appearance of “time” and space, which makes them easier to remember.
- These symmetrical equations—along with “time” having the same dimension (of distance) as space—suggest the possibility that “time” and space might be considered as the two sides of a single coin called “spacetime”. This is only a *possibility* from the outset, since the real thing that is needed for time and space to be joined into one entity is the idea of a *metric*. But that is another story, for relativity books to tell.
- The many occurrences of c that would otherwise clutter the equations are eliminated.

But a big potential problem with writing (1.210) is that we might forget that the “time” t is not really time as we know it. For example, a standard relativistic analysis produces a particular “time” difference of vL , where v is the “velocity” of an object, and L is its rest length (its length in its own frame). If we settle on, say, using SI units, and consider an object moving at 20 m/s with a rest length of 50 m, we must not fall into the trap of writing “The time difference is $vL = 20 \times 50$ seconds”. Instead, including units makes it apparent that the time difference is really $vL = 20 \times 50 \text{ m}^2/\text{s}$. We must restore this expression to conventional dimensions by dividing by c^2 . Doing so yields a time difference of $20 \times 50 / (3 \times 10^8)^2$ seconds, or about 10^{-14} seconds: a number that is vastly smaller than the 1000 seconds that we might at first have written.

The moral of the story is that if you remove the safety barriers to your calculation by not explicitly writing all internal units, then you must

always be careful to ascertain that what you think you are calculating is what you really *are* calculating.

1.9.1 Units versus Dimensions

Whereas it's convenient to choose a standard set of units such as SI, the choice has sometimes been made for us, typically when writing computer code to an established specification. Switching between units can be hard work for some scientists, who might “wing it” by picturing how the relevant numbers grow bigger or smaller for various choices of unit, and then inserting the appropriate compensating factors manually, followed by some ad-hoc testing to see if everything works. This is fine in simple cases; but with more complex mixes of units, employing a structured approach is far less stressful. We discuss such an approach next: one that we can walk away from mid-calculation and come back to later, without having to remember what factors we were adding based on intuitive ideas of what is large or small.

Begin with the following typical sentence that you might encounter in an engineering textbook:

“The speed of a wave is given by $v = f\lambda$, where v is in metres/second, f is in hertz, and λ is in metres.”

Although it's true that engineers must pay careful attention to the units they use, the above engineering sentence is somewhat misleading, because it implies that the choice of units is crucial to the mathematical form of the expression. Sometimes, the choice of units *does* change the form of the expression, but not in the above case! (If the choice of units was relevant in a particular equation, those units would be included in that equation.) The speed of a wave depends on its frequency and wavelength via $v = f\lambda$ *irrespective* of the units chosen, provided only that those units have the correct *dimensions*. This is an important point: the length of a ruler has a single *dimension*: “length”, but it can be expressed in any of an infinite choice of *units*: metres, feet, furlongs, angstroms, etc. A quantity's definition fixes its dimensions, but we are always free to express that quantity in whatever units we prefer, as long as those units have the correct dimensions. In a similar vein, it is incorrect to state that the density of a substance is defined to be its mass per cubic metre. No; the density of a substance is defined to be its mass per unit *volume*. This unit of volume is set by our choice of length units. It will be a cubic metre if we use SI, but it can be something else if we choose to work in a different unit system.

But consider that in the numerical manipulations of a computer, the variables used in a computer programme really have no units; they are simply

dimensionless numbers. If we are required to obey a convention that v is measured in, say, miles per hour, f in cycles per fortnight, and λ in kilometres, then how should we write the expression “ $v = f\lambda$ ” in computer code, which always takes only dimensionless numbers as input? The answer calls on a simple rule that will always serve to disentangle any problem involving multiple systems of units. Observe that “6/2” means “the number of 2s in 6” (that is, 3). In exactly the same way, the expression “ $x/(1\text{ km})$ ” means “the number of kilometres in x ”, which is another way of saying “ x expressed in kilometres”. The number of kilometres in a given distance is simply a number: it is dimensionless. Hence, in computer code, when we use a variable that represents a distance x that we wish to express in kilometres, we do not and cannot use x , since x has dimensions (of length), and computers can only process dimensionless numbers. Instead, we must use a dimensionless quantity such as $x/(1\text{ km})$. It is good programming practice to name this variable something like “**x_km**” to emphasise that it is a number of kilometres; **x_km** is (and can only be) a dimensionless number: it equals $x/(1\text{ km})$, and not x .

Return to the original task of writing $v = f\lambda$ in a form that uses the units for its three quantities mentioned above (v in miles per hour, f in cycles per fortnight, λ in kilometres). We note that:

- Expressing v in miles per hour means that only the *dimensionless* quantity $v/(1\text{ mile hour}^{-1})$ can appear.
- With f in cycles per fortnight, only the dimensionless $f/(1\text{ fortnight}^{-1})$ can appear; “cycles” needn’t be written, because it isn’t a dimension.
- λ in kilometres stipulates the dimensionless term $\lambda/(1\text{ km})$.

To write $v = f\lambda$ in the required way, begin by dividing and multiplying by each required factor in turn, which, of course, cannot change the original equation:

$$\frac{v}{1\text{ mile hour}^{-1}} \frac{1\text{ mile}}{1\text{ hour}} = \frac{f}{1\text{ fortnight}^{-1}} \frac{1}{\text{fortnight}} \frac{\lambda}{1\text{ km}} 1\text{ km}. \quad (1.211)$$

Now collect the factors containing only the units into one factor on one side of the equation:

$$\frac{v}{1\text{ mile hour}^{-1}} = \frac{f}{1\text{ fortnight}^{-1}} \frac{\lambda}{1\text{ km}} \underbrace{\frac{1}{\text{fortnight}} 1\text{ km} \frac{1\text{ hour}}{1\text{ mile}}}_{\text{call this } a}. \quad (1.212)$$

What is a ? Convert all of its parts into any convenient system of units, using a single unit for each dimension; that way, as many units as possible will cancel each other out. Let’s use SI units, along with 5280 feet in a mile, and 3.28084 feet in a metre:

$$\begin{aligned}
 a &= \frac{1 \text{ km} \times 1 \text{ hour}}{1 \text{ fortnight} \times 1 \text{ mile}} = \frac{1000 \cancel{\mu} \times \cancel{3600 \text{ s}}}{14 \times 24 \times \cancel{3600 \text{ s}} \times \frac{5280}{3.28084} \cancel{\mu}} \\
 &= \frac{1000 \times 3.28084}{14 \times 24 \times 5280} \simeq 1.85 \times 10^{-3} \quad (\text{dimensionless!}).
 \end{aligned}$$

Equation (1.212) is now written as follows, where we ignore that our value of a is only approximate:

$$\frac{v}{1 \text{ mile hour}^{-1}} = \frac{f}{1 \text{ fortnight}^{-1}} \frac{\lambda}{1 \text{ km}} \times 1.85 \times 10^{-3}. \quad (1.213)$$

That is, v expressed in miles per hour equals f expressed in cycles per fortnight times λ expressed in kilometres times 1.85×10^{-3} . But take note: (1.213) does not define v as a number of miles per hour (and similarly for the other two variables). Rather, the expression “ $v/(1 \text{ mile hour}^{-1})$ ” means “ v expressed in miles per hour”. In other words (and consider this carefully), just as “ $v = f\lambda$ ” is true for any choice of units of v, f, λ , so too (1.213) is true for any choice of units of v, f, λ . For example, if the wavelength λ is ten kilometres, then we might express this as “ $\lambda = 10 \text{ km}$ ”, or we might choose to write “ $\lambda = 32,808.4 \text{ feet}$ ”, or maybe something else. In either case, we’ll conclude that $\lambda/(1 \text{ km}) = 10$:

$$\frac{\lambda}{1 \text{ km}} = \frac{10 \cancel{\text{ km}}}{1 \cancel{\text{ km}}} = 10, \quad \text{or} \quad \frac{\lambda}{1 \text{ km}} = \frac{32,808.4 \cancel{\text{ feet}}}{3280.84 \cancel{\text{ feet}}} = 10. \quad (1.214)$$

Take note of the economy of such expressions. If, for example, L is some length, then $L/(1 \text{ cm})$ is a genuine fraction that obeys all the rules of algebra while also expressing that length in centimetres; $L/(1 \text{ in})$ does likewise while expressing that length in inches, and so on. There is a single length here, L , and it is a true dimensioned length, not just a number. In contrast, you will sometimes see recipes for converting between units that takes, say, L_{cm} to be a dimensionless number—the length in centimetres—and then another symbol, say, L_{in} , is created—again, a dimensionless number that expresses the length in inches, so that $L_{\text{cm}} = 2.54 L_{\text{in}}$. This procedure not only defines a possible zoo of symbols, but each symbol is a dimensionless number instead of a real length, which runs contrary to the usual use of symbols in physics. Even so, when writing a computer programme, we *do* need to create separate variables for different unit choices, because computers can only deal with dimensionless numbers: thus, $\text{L\_cm} \equiv L/(1 \text{ cm})$, $\text{L\_in} \equiv L/(1 \text{ in})$, and typical computer code to convert inches to centimetres is

$$\text{L\_cm} = 2.54 * \text{L\_in}$$

Another example of manipulating units can be found with the definition of an *electron volt* “eV”. This is a non-SI unit of energy that is convenient for the scale of atomic energies. Although it can be defined in terms of an

electron, it is more easily defined using a proton, which we do in the following grey box.

The Electron Volt

An electron volt equals the kinetic energy gained by a proton that falls through a drop in electric potential of one volt, where one volt is one joule of electric potential energy per coulomb of charge.

(Analogously, the kinetic energy gained by a mass that falls through a drop in gravitational potential equals its mass times the potential drop.) Hence, the electron volt equals the proton's charge (conventionally¹⁶ denoted $e \equiv 1.602 \times 10^{-19}$ C) times the potential drop of 1 joule/coulomb. If the kinetic energy gained by the proton is E , then E expressed in electron volts is

$$\begin{aligned} \frac{E}{1 \text{ eV}} &= \frac{E}{e \times 1 \text{ J/C}} = \frac{E}{e/(1 \text{ C}) \times 1 \text{ J}} = \frac{E/(1 \text{ J})}{e/(1 \text{ C})} \\ &= \frac{E \text{ in joules}}{e \text{ in coulombs}} = \frac{E \text{ in joules}}{1.602 \times 10^{-19}}. \end{aligned} \quad (1.215)$$

To make an easy mnemonic for the conversion (which we will use on occasion), note that if—just for this mnemonic—we set e to mean the pure number 1.602×10^{-19} , then

$$1 \text{ J} = \frac{1}{e} \text{ eV}, \quad \text{and } 1 \text{ eV} = e \text{ J}. \quad (1.216)$$

Try using these mnemonics when you next need to do the conversion.

On the subject of working with units, you might occasionally hear it said that to calculate, say, a ratio of distances x_1/x_2 , “You can choose any units you like, as long as they are the same for numerator and denominator”. Not at all; we can validly express the first distance in metres, and the second in feet, with, as usual, $1 \text{ m} \simeq 3.28084 \text{ ft}$. For example, if $x_1 = 200 \text{ m}$ and $x_2 = 328.084 \text{ ft}$, then their ratio is

$$\frac{x_1}{x_2} = \frac{200 \text{ m}}{328.084 \text{ ft}} = \frac{200}{328.084} \text{ m/ft}. \quad (1.217)$$

This is fully correct; but the right-hand side of (1.217) is dimensionless—a pure number—and so we can do better by eliminating the units entirely, by

¹⁶ It's a matter of historical inelegance that the electron volt is more easily described with reference to a proton—whose charge is called “ e ”. The electron has charge $-e$, and this minus sign just makes for awkwardness if we describe the electron volt by referring to electrons. But if you want to do that, then begin with the electron volt being the kinetic energy gained by an electron that falls through an *increase* in electric potential of one volt.

converting either “m” or “ft” into the other. The choice is arbitrary, so suppose we convert metres to feet:

$$\frac{x_1}{x_2} = \frac{200 \text{ m}}{328.084 \text{ ft}} = \frac{200 \times 3.28084 \cancel{\text{ft}}}{328.084 \cancel{\text{ft}}} = \frac{200 \times \cancel{3.28084}}{100 \times \cancel{3.28084}} = 2. \quad (1.218)$$

This is exactly equivalent to (1.217), but obviously simpler.

On the same note, we can add quantities expressed in different *units*, as long as they have the same *dimension*. It’s completely correct to say

$$1 \text{ metre} + 1 \text{ centimetre} = 101 \text{ centimetres} = 1.01 \text{ metres}. \quad (1.219)$$

There is nothing mysterious here; we need only appreciate the distinction between a unit and a dimension.

Dimensionless Units

This discussion of units versus dimensions sheds light on *dimensionless units* such as radians and moles.

The Radian

Recall that the size θ of an angle is defined as being the length s of the circular arc that the angle scribes at a given radius r , divided by r :

$$\theta \equiv s/r. \quad (1.220)$$

Clearly, θ is dimensionless: one might say it is equivalent to a metre per metre, an inch per inch, or a light-year per light-year. All of these are the same:

$$2 \text{ radians} = 2 \text{ metres/metre} = 2 \text{ inches/inch} = \dots \quad (1.221)$$

The radian is a dimensionless unit: a sort of “comfort unit” that we voice purely to avoid an awkward-sounding (but correct) phrase such as “The angle is 2”. We are so used to specifying angles in degrees that we expect to hear a unit said after the “2”, in which case “The angle is 2” sounds incomplete. To avoid what sounds like linguistic clumsiness, we say “The angle is 2 radians” instead. In any equation that uses radians, we can choose either to write “radian” or omit it; nothing will be upset in the equation either way, but including the word might well make everything more transparent.

The Mole

The same idea of a “comfort unit”, a dimensionless unit, applies to the *mole*. The mole (a linguistic derivative of “molecule”) relates to *Avogadro’s number*:

$$\begin{aligned} \text{Avogadro's number } N_A &\equiv \text{the number of atoms in 12 grams of carbon-12} \\ &\simeq 6.022 \times 10^{23}. \end{aligned} \tag{1.222}$$

A mole of carbon-12 is defined to be exactly 12 grams of carbon-12. It follows that a mole of carbon-12 consists of exactly N_A atoms of carbon-12. A mole of any substance is then defined to be the amount of that substance contained in N_A of its specified basic entities. An *atom* of carbon-12 is specified in the definition of N_A ; but when atoms can form molecules, we must be more explicit. “A mole of oxygen-16” has no meaning. Rather, a mole of oxygen-16 *atoms* (N_A atoms) has a mass of 16 grams, whereas a mole of oxygen-16 *molecules* (N_A molecules) has a mass of 32 grams, and so on. When we speak of a mole of something, the basic unit of that something must either be stated explicitly or its meaning understood.

Is a Mole an Amount or a Number?

The SI definition of a mole is an *amount* and not a *number*; but in practice, the distinction between these two things becomes little more than semantic. Consider that you might understand the word “dozen” to be the *amount* of a substance contained in 12 of that substance’s basic units. But you might just as well understand the word “dozen” to be simply the *number* 12, provided it refers to an object; after all, we never denote by “dozen” the point on the number line that is halfway between 11 and 13.

In everyday life, we don’t distinguish between these two apparently different meanings of a “dozen”, as being an *amount* or a *pure number*. We can equally well speak of the *amount* (“A dozen eggs is the amount of eggs in 12 eggs”) and the *number* (“A dozen eggs is 12 eggs”). When using the word “dozen”, we always specify the basic unit. For example, a dozen oxygen atoms has half the mass of a dozen oxygen molecules; but in both cases, we are considering 12 instances of the specified unit: atom or molecule.

By the same token, a mole of some object is N_A instances of that object; it can be viewed as a word, like “dozen” (12 objects), “score” (20 objects), and “gross” (144 objects).

Being a dimensionless unit, a mole need never be written as a unit in an equation. But just as with the radian, writing “mole” often renders the meaning more transparent. For example, what is the *molar mass* M_{mol} of water? This is the mass of one mole of water, and its value is approximately 18 grams: “ $M_{\text{mol}} = 18 \text{ g}$ ”. It is quite normal to write “ $M_{\text{mol}} = 18 \text{ g/mol}$ ”, which reinforces that we are talking about a mole; but a “gram per mole” has dimensions of mass only. The unit “mole” is superfluous here, because M_{mol} is *defined* in terms of a mole. (You can certainly ignore the modern pseudo-distinction

that some like to make between Avogadro's number, N_A , and the “Avogadro constant”, which is said to be “ N_A per mole”.)

If you are in doubt about how “mole” should be used, try replacing it with “dozen” and re-ask any question you might have. For example, corresponding to the molar mass might be something called a “dozen price”, which is the price of a dozen of some quantity, say, eggs. If each egg costs 50 cents, then the dozen price of eggs is 6 dollars. We would probably choose to say “Eggs have a dozen price of 6 dollars” rather than “Eggs have a dozen price of 6 dollars per dozen”, because the notion of “dozen” is already built in to the definition of a “dozen price”. Compare this everyday example to the phrases “The molar mass of water is 18 grams” and “The molar mass of water is 18 grams per mole”, and you will see the same idea applies—and yet, in the case of molar mass, it's the *second* phrase that is more widely used!

Another example of the mole's use occurs with the molar heat capacity C^{mol} defined in (4.15) ahead. The molar heat capacity of a given substance is the heat capacity C (units of joules per kelvin, J/K) of one mole of that substance. Thus, the molar heat capacity is a very particular example of a heat capacity, but it's a heat capacity nonetheless; and hence C^{mol} has dimensions of energy over temperature, meaning SI units of J/K. But it is normal to distinguish C from C^{mol} if possible, and so C^{mol} is almost universally given units of “joules per kelvin per mole: J/(K mol), or $\text{J K}^{-1} \text{ mol}^{-1}$. When equation (4.15) defines the molar heat capacity, it refers to n moles of the given substance, and so n is a *pure number*. If we have 2 moles of substance, then $n = 2$ (meaning we have $2N_A$ particles of the substance); but, in practice, we can write “ $n = 2$ moles”, because the unit “mole” is dimensionless. Here, “2 moles” is treated as a pure number, equal to $2 \times 6.022 \times 10^{23}$ —just as “2 dozen” is read as 2×12 .

The Particle, Car, etc.

Besides the radian and mole, another common dimensionless unit is the name of any object, such as “particle”. If we write “Let N be the number of particles in the box”, then N is a pure number: 12. We might write “ $N = 12$ particles” just to show that we know what N is—and also to remind an audience what N is; but it is sufficient to write simply $N = 12$. And if required to calculate N^{1000} , we write the answer as 12^{1000} ; we do not write “ 12^{1000} particle¹⁰⁰⁰” because that expression takes the idea of a comfort unit too far, even though it is harmless. At the start of this chapter, we even wrote expressions such as (1.3) that had a number of particles as the *exponent*; again, that was valid, because that exponent was dimensionless.

Similarly, when we analyse the logistics of a factory that assembles cars, we might define a quantity “ C ” to be the rate at which cars are built. The dimension of C is time^{-1} ; for example, if we build 50 cars per day, then $C = 50/\text{day}$ (or $C = 50 \text{ day}^{-1}$). We might write “ $C = 50 \text{ cars/day}$ ”, but “car” is just a dimensionless unit that can be included or omitted as we see fit. If we must

square this rate as part of some analysis, we should write $C^2 = 2500/\text{day}^2$ (or $C^2 = 2500 \text{ day}^{-2}$). Writing $C^2 = 2500 \text{ cars}^2/\text{day}^2$ takes the idea of a dimensionless unit a little too far.¹⁷

The “Man” in Man-Hour

A somewhat different dimensionless unit is the “man” in “man-hour”. If a team of 3 men work together to mow a lawn in 2 hours, what can be said about how long it takes, generically, to mow the lawn? To say “This lawn takes 2 hours to mow” would be misleading if the size of the team was apt to vary. Instead, we pose the standard question: “How many hours would this lawn take one man to mow?”. The answer is 6 hours. But we cannot just say “This lawn will take 6 hours to mow”, because the job clearly can be done in 2 hours. We signal that we are answering the standard question posed, by saying “This lawn takes 6 man-hours to mow”. It follows that a “man-hour” is simply an hour: its dimension is time. On the other hand, it would be wrong to say “The length of time between 2 p.m. and 3 p.m. is one man-hour”, because now, there is no suggestion of a job being done; we are no longer answering the standard question, “How long would this job take one man to do?”.

“Per Unit of Some Quantity”

Another use of units can be confusing if encountered under the wrong circumstances. The expression “per unit of some quantity” always requires a *proportionality* to that quantity in order to make good sense. That is, if eggs cost “6 dollars per dozen”, we infer correctly that n dozen eggs will cost $6n$ dollars. Ignoring this proportionality requirement can land us in trouble. For example, the resistance of a 1-kilometre-long copper wire of cross-sectional area 1 square metre is about $17 \mu\Omega$; but it would be misleading to say “1-kilometre-long copper wire has a resistance of $17 \mu\Omega$ per square metre of cross-sectional area”. The reason is that the resistance of a wire is *not* proportional to its cross-sectional area. Instead, the resistance is *inversely* proportional to the

¹⁷ I once corresponded briefly with a rather senior economics academic who was convinced that much economics theory is wrong because, it seemed to him, it uses units inconsistently. Whether economics is logically coherent outside of its use of units is a separate question, but it was clear to me that his reasoning on units suffered from the misconceptions that I have outlined above; for example, he was unaware that “car” is a dimensionless unit. He certainly felt that physics uses units correctly, and I thought he was justified in being baffled by the criticism of a referee who had rejected one of his papers, and who had commented, in essence, “Sure, economics doesn’t use units correctly. But neither does physics, so we economists must be doing okay!”. That incorrect presumption about both economics and physics indicates what is probably a widespread lack of knowledge of the correct use of units.

cross-sectional area, and so the resistance of a 1-kilometre-long copper wire with a 2 m^2 cross-sectional area is not $2 \times 17 \mu\Omega$, but rather $1/2 \times 17 \mu\Omega$.

Conductance is defined as the reciprocal of resistance, and so the conductance of a wire *is* proportional to its cross-sectional area; so, we are well entitled to say “1-kilometre-long copper wire has a conductance of $1/17 \mu\Omega^{-1}$ per square metre of cross-sectional area”. It will follow that the conductance of a 1-kilometre-long copper wire with a 2 m^2 cross-sectional area is $2/17 \mu\Omega^{-1}$. This is why it is sometimes better to speak in terms of conductance than in terms of resistance, even though the two terms are apparently so trivially related.

1.9.2 Function Arguments Must Be Dimensionless

Consider that, by *definition*, a function such as the exponential “exp” is not designed to act on dimensions or units: it incorporates no machinery to do so. It can be taken only of a dimensionless quantity; and thus the exponential of, say, a volume, is not defined. Any algorithm that returns the exponential can operate only on a dimensionless quantity; it cannot “know” anything about dimensions or units.

You will sometimes encounter the following argument: “ $e^{2 \text{ metres}}$ can make no sense, because this expression must be writable as the exponential series

$$e^{2 \text{ m}} = 1 + 2 \text{ m} + \frac{2^2 \text{ m}^2}{2!} + \frac{2^3 \text{ m}^3}{3!} + \dots, \quad (1.223)$$

and since we cannot add 1 to a metre to a square metre and so on, the series must be undefined, and hence the expression $e^{2 \text{ metres}}$ is undefined.” This might look reasonable; but, in fact, things are not so straightforward. Consider that if x has dimension of length and the exponential is defined as the usual series, then

$$\begin{aligned} \exp x &= \exp 0 + \exp'(0) x + \frac{\exp''(0) x^2}{2!} + \dots \\ &= \exp 0 + \left. \frac{d \exp x}{dx} \right|_{x=0} x + \left. \frac{d^2 \exp x}{dx^2} \right|_{x=0} \frac{x^2}{2!} + \dots \end{aligned} \quad (1.224)$$

Clearly, each term in the second line above *does* have the same dimension, because whatever dimension x has, this dimension always appears raised to the same power in the numerator and denominator of each term, and so cancels itself out. But this means that if we insist on including dimensions everywhere, then we can no longer say “ $\exp'(x) = \exp x$ ”, because when x has dimension of length, the dimension of $\exp'(x)$ must be the dimension of $\exp x$ divided by length. So, if we insist on allowing “exp” to be taken of

dimensioned quantities, the calculus that we are so familiar with crumbles into a heap of details swirling around the dimensions currently being used. Mathematics simply was not created for such analyses.

With this idea in mind that a function cannot be taken of a dimensioned quantity, consider that, after writing the valid expression

$$A = \log \frac{\text{volume}}{\text{length} \times \text{area}} , \quad (1.225)$$

we cannot then write

$$“ A = \log \text{volume} - \log \text{length} - \log \text{area} ” \quad (\text{Wrong!}). \quad (1.226)$$

Instead, we must choose a set of units such as SI, and write [by dividing both the numerator and denominator of (1.225) by 1 m^3]

$$A = \log \frac{\text{volume}/(1 \text{ m}^3)}{[\text{length}/(1 \text{ m})] \times [\text{area}/(1 \text{ m}^2)]} . \quad (1.227)$$

This now contains only dimensionless quantities, and thus can be written as

$$A = \log \frac{\text{volume}}{1 \text{ m}^3} - \log \frac{\text{length}}{1 \text{ m}} - \log \frac{\text{area}}{1 \text{ m}^2} . \quad (1.228)$$

Just as with the discussion of $v = f\lambda$ in Section 1.9.1, we can use any *mixture* of units in (1.227), as long as they have the correct dimensions and we include any appropriate unit-conversion factors, such as the “ a ” of (1.212). For the sake of discussion, we’ll keep all units here the same, such as SI—which means we don’t need an extra unit-conversion factor [such as the “ a ” of (1.212)]. Thus, if we agree to use the same unit throughout the calculation, we can omit it and simply write

$$A = \log \text{volume} - \log \text{length} - \log \text{area} . \quad (1.229)$$

This expression will always return the correct value for A regardless of our choice of units, provided simply that we use the same unit for each dimension that appears. This implied use of consistent units tidies and simplifies such expressions, so that we *can* write expressions such as “log volume”. If you happen upon an instance of “log volume”, expect to see balancing expressions such as “log distance” or “log area” somewhere close by.

A slightly more involved example of this is pertinent to what will follow in the coming chapters. Consider a dimensionless quantity X :

$$X = V \left(\frac{ME}{h^2} \right)^{3/2} , \quad (1.230)$$

where V is volume, M mass, E energy, and h is Planck’s constant, which has dimensions of “energy $\times$ time”. We wish to take the logarithm of X while

keeping volume separate from the other parameters. We might write

$$\log X = \log V + {}^{3/2} \log \frac{ME}{h^2}, \quad (1.231)$$

but now, we are back to taking logarithms of dimensioned quantities. Just as in (1.227), we fix this by writing (1.230) in the entirely equivalent form of

$$X = \frac{V}{1 \text{ m}^3} \left(\frac{1 \text{ m}^2 \times ME}{h^2} \right)^{3/2}. \quad (1.232)$$

Now take a logarithm:

$$\log X = \log \frac{V}{1 \text{ m}^3} + {}^{3/2} \log \frac{1 \text{ m}^2 \times ME}{h^2}. \quad (1.233)$$

Each term on the right-hand side of (1.233) is now dimensionless, and thus completely well defined. As usual, “ $V/(1 \text{ m}^3)$ ” means “ V expressed in cubic metres”, which is a pure number—it has no units. If $V = 2 \text{ m}^3$, then $V/(1 \text{ m}^3)$ is the pure number 2. Hence, we are content to write (1.231) after all, knowing now that the “ 1 m^3 ” and “ 1 m^2 ” in (1.233) are invisibly present: (1.231) really means (1.233). Then, provided we use a consistent set of units such as SI, we can replace each parameter in (1.231) with its SI value as a pure number, and everything numerical that later follows will be self consistent.

For another example, in (4.192) we’ll encounter the following differential equation to be solved:

$$\frac{dP}{dT} = \frac{aP}{T^2}, \quad (1.234)$$

for pressure P and temperature T , given a known constant a . Our task right now is to solve (1.234), and compare the result with experimental data. Begin by rearranging it into

$$\frac{dP}{P} = \frac{a dT}{T^2}. \quad (1.235)$$

The next step is to integrate this equation, and you will often find the result loosely written as “ $\ln P = -a/T + \text{constant}$ ”. But as we have seen above, “ $\ln P$ ” makes no sense. Instead, recognise that

$$\int \frac{dP}{P} = \ln \frac{P}{P_0}, \quad (1.236)$$

for some constant P_0 with the dimension of pressure. (You should check this expression by differentiating its right-hand side with respect to P .) Equation (1.235) now integrates to

$$\ln \frac{P}{P_0} = \frac{-a}{T} + \text{another constant}. \quad (1.237)$$

Exponentiate both sides of this, to obtain

$$P = b \exp \frac{-a}{T}, \quad (1.238)$$

where b is another constant. When asked to calculate values of a and b from experimental data of P and T , students are apt to write (1.238) in a linear format as

$$“ \ln P = \ln b - a/T ” \quad (\text{Wrong!}), \quad (1.239)$$

and conclude that $\ln P$ must be plotted against $1/T$ to give a straight line of slope $-a$ and “ y -intercept” $\ln b$. But then, how do we calculate $\ln P$: what units must be used for P ? Of course, it makes no sense to calculate $\ln P$, and we should not write (1.239) at all. Instead, choose any system of units to write (1.238) in dimensionless form. For example, dividing by a pascal and a kelvin produces

$$\frac{P}{1 \text{ Pa}} = \frac{b}{1 \text{ Pa}} \exp \frac{-a/(1 \text{ K})}{T/(1 \text{ K})}. \quad (1.240)$$

Now both sides are pure numbers, so it’s perfectly permissible to take their logarithms:

$$\ln \frac{P}{1 \text{ Pa}} = \ln \frac{b}{1 \text{ Pa}} - \frac{a/(1 \text{ K})}{T/(1 \text{ K})}. \quad (1.241)$$

If we now plot $\ln[P/(1 \text{ Pa})]$ (that is, $\ln$ of [pressure expressed in pascals]) versus $1/[T/(1 \text{ K})]$ (the reciprocal of [temperature expressed in kelvins]), we expect the data to lie on a line that has a slope and a “ y -intercept” of

$$\text{slope} = \frac{-a}{1 \text{ K}}, \quad “y\text{-intercept}” = \ln \frac{b}{1 \text{ Pa}}. \quad (1.242)$$

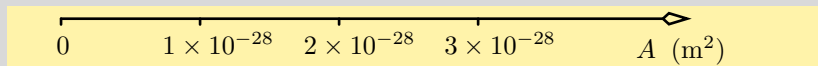
If the measured slope is 5, we conclude that $a = -5 \text{ K}$. And if the measured y -intercept is 2, we conclude that $b = e^2 \text{ Pa}$. Any system of units can be chosen for this exercise: whatever you use, the results will be equivalent.

The role of units in *physics* is logically consistent; but, in practice, *physicists* don’t always get it right. It is not uncommon to see a textbook plot with an axis labelled, say, “log of temperature”, with the accompanying “(log kelvins)” for its unit. But one cannot take the log of a temperature; there is no such thing as the log of a kelvin or of any other unit. One can certainly take the log of a temperature that has been expressed in kelvins, because $T/(1 \text{ K})$ is a pure number, not a temperature: the temperature is that pure number times one kelvin. Such a plot axis should be labelled “log of temperature in kelvins”—say, “ $\log_{10}[T/(1 \text{ K})]$ ”—which has *no units at all*.

Labelling the Axes of a Plot

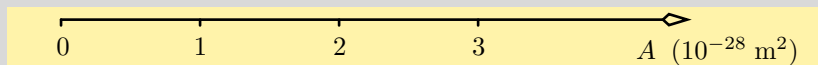
With the above discussion of units in mind, you will appreciate the nuances of labelling plot axes. Suppose we are plotting the results of an

experiment in nuclear scattering, in which a typical target cross-sectional area is around 10^{-28} m^2 . We can label a plot axis that shows a range of these areas A , in the following way:



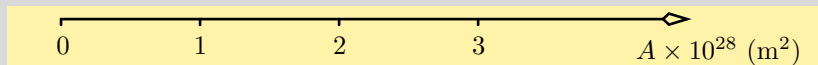
The meaning is clear: for example, we understand that the value of A at the first tick is the label “ 1×10^{-28} ” times the unit in parentheses, m^2 , producing $A = 1 \times 10^{-28} \text{ m}^2$. The value of A at the next tick is the label “ 2×10^{-28} ” times the unit in parentheses, m^2 , and so on. It’s important to note that the standard way of indicating the quantity and its units, “ $A (\text{m}^2)$ ”, does *not* mean that we are plotting $A \times 1 \text{ m}^2$.

But the “ $\times 10^{-28}$ ” is clutter that should be kept away from the ticks. A more concise labelling is



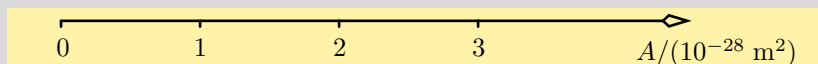
Again, we understand that the value of A at the first tick is the label “1” times the unit in the parentheses, 10^{-28} m^2 , giving us $A = 1 \times 10^{-28} \text{ m}^2$.

Alternatively, we can plot $A \times 10^{28}$:



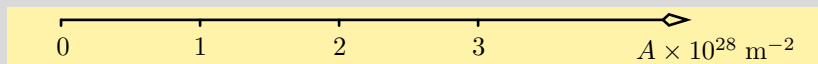
The rule is unchanged: the value of the plotted quantity $A \times 10^{28}$ at the first tick is the label “1” times the unit in the parentheses, m^2 . Then, $A \times 10^{28} = 1 \text{ m}^2$, or $A = 1 \times 10^{-28} \text{ m}^2$.

Or, we can plot the *dimensionless* quantity $A/(1 \text{ m}^2)$ —or better yet, $A/(10^{-28} \text{ m}^2)$, which keeps the 10^{-28} away from the axis ticks:



As always, there is only one rule: the value of the plotted quantity $A/(10^{-28} \text{ m}^2)$ at the first tick is the label “1” times the unit. But now there is no unit, because the parentheses are part of the quantity being plotted; they are not a separate container holding a unit. So, at the first tick, $A/(10^{-28} \text{ m}^2) = 1$, or $A = 1 \times 10^{-28} \text{ m}^2$.

It’s clear that $A/(10^{-28} \text{ m}^2) = A \times 10^{28} \text{ m}^{-2}$. Thus, we can just as well relabel the plot as “ $A \times 10^{28} \text{ m}^{-2}$ ”:



Again, the value of the plotted quantity $A \times 10^{28} \text{ m}^{-2}$ at the first tick is the label “1” times the unit; but there is no unit. So, at the first tick,

$A \times 10^{28} \text{ m}^{-2} = 1$, or $A = 1 \times 10^{-28} \text{ m}^2$. Compare this plot's label with the label of the third plot above: whether or not something is in parentheses makes all the difference!

The five plot labels above all allocate $A = 1 \times 10^{-28} \text{ m}^2$ to the first tick. Which one to use is a matter of taste. Notice the occurrences of both 10^{+28} and 10^{-28} , along with m^2 and m^{-2} , and also note what is inside and outside parentheses. Finally, be aware that you might find (not in this book) labels that read as either $A (\times 10^{-28} \text{ m}^2)$ or $A (\times 10^{28} \text{ m}^2)$. Taken out of context, the meaning of these is not clear. What is being plotted? And what is the unit? The moral of the story is that you should always ascertain from the numbers what is *really* being plotted.

1.9.3 Distinguishing Between an Entity and its Representation

The above discussion of the use of units in physics forms part of a larger subject that distinguishes a quantity of interest from the perhaps arbitrary or conventional way in which we choose to quantify it. This topic is of extreme importance to physics and appears all throughout the subject. Perhaps the simplest case of such a separation can be seen by using pure numbers. Consider the number of sticks here:

|||| | |||| ||

These are often rendered more readably as $\text{||||} \text{||||} \text{||}$. Mankind learned long ago that such a scheme of representing numbers is cumbersome; and thus, in modern mathematics, the base-10 *representation* of this number is universally used. We can write that as

$$[\text{||||} \text{||||} \text{||}]_{\text{base } 10} = \text{"12"}, \quad (1.243)$$

although we would conventionally understand the 12 to be a base-10 number, and so would omit the quotes. The central point here is to distinguish the *number* or *entity* $\text{||||} \text{||||} \text{||}$ from its *representation* "12" in the chosen system of base 10. The brackets in (1.243) make this distinction clear: they indicate a representation of their enclosed contents in the specified number system. For example, we might choose to work in base 2:

$$[\text{||||} \text{||||} \text{||}]_{\text{base } 2} = \text{"1100"}. \quad (1.244)$$

Or perhaps we prefer Egyptian hieroglyphic:

$$[\text{||||} \text{||||} \text{||}]_{\text{Egypt}} = \cap \cap \quad (1.245)$$

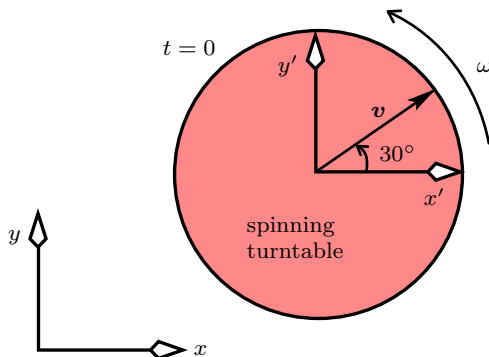


Fig. 1.12 The vector $\mathbf{v}$ is formed by an arrow fixed to a spinning turntable. What are the components of $\mathbf{v}$ in primed and unprimed coordinates?

or maybe base π :

$$[\text{||||} \text{||||} \text{||}]_{\text{base } \pi} = \text{“102.01022 12222 11211...”}. \quad (1.246)$$

The same number $\text{||||} \text{||||} \text{||}$ is always present here (as enclosed in the brackets), but it can be represented in an infinite number of different ways.

This idea of distinguishing between an entity and its representation in an appropriate language also occurs with dimensioned quantities in physics. The length of a table is some unique quantity L :

$$L = 5 \text{ metres} \simeq 16.404 \text{ feet} \simeq 5.285 \times 10^{-16} \text{ light years}, \quad (1.247)$$

and this single quantity can be represented using any choice of units:

$$\begin{aligned} [L]_{\text{SI}} &\equiv \frac{L}{1 \text{ metre}} = 5, \\ [L]_{\text{imperial}} &\equiv \frac{L}{1 \text{ foot}} \simeq 16.404, \\ [L]_{\text{astronomy}} &\equiv \frac{L}{1 \text{ light year}} \simeq 5.285 \times 10^{-16}. \end{aligned} \quad (1.248)$$

There are not three separate variable names here; just one: L , with the brackets indicating a representation of L . The quantity itself, L , is what appears in equations. In contrast, as mentioned earlier, a computer can deal only with dimensionless numbers; and so in a computer programme, we must represent L by a dimensionless variable, such as “ $L_{\text{SI}} = 5$ ”.

This concept of a representation extends to the realm of vectors, but now, along with a choice of length units, we must choose a *coordinate system*. Picture the vector $\mathbf{v}$ —an arrow—shown in Figure 1.12. This vector points from the centre of a turntable that spins with angular velocity ω in a laboratory.

The laboratory has coordinate axes x, y , and the turntable has coordinate axes x', y' , where, at time $t = 0$, the x' axis is parallel to the x axis. At all times, the vector makes the angle 30° to the x' axis, and has a fixed length $v \equiv |\mathbf{v}|$. This describes the vector, but we might have a need to *coordinatise* it, meaning represent it in a chosen coordinate system. It is simplest to choose the primed coordinate system x', y' , for which we write

$$[\mathbf{v}]_{\text{primed}} = \begin{bmatrix} v_{x'} \\ v_{y'} \end{bmatrix} = \begin{bmatrix} v \cos 30^\circ \\ v \sin 30^\circ \end{bmatrix} = v \begin{bmatrix} \sqrt{3}/2 \\ 1/2 \end{bmatrix}. \quad (1.249)$$

Again, the brackets emphasise a *representation* of the unique vector $\mathbf{v}$ in the chosen coordinates: note that $\mathbf{v}$ is a vector (an arrow), whereas $[\mathbf{v}]_{\text{primed}}$ is a 2×1 matrix. In the unprimed coordinate system x, y (the laboratory), the vector spins at ω , and we use the well-known rotation matrix in the xy plane to write $[\mathbf{v}]_{\text{primed}}$ at an arbitrary time t :

$$\begin{aligned} [\mathbf{v}]_{\text{unprimed}}(t) &= \begin{bmatrix} \cos \omega t & -\sin \omega t \\ \sin \omega t & \cos \omega t \end{bmatrix} [\mathbf{v}]_{\text{primed}} = \begin{bmatrix} \cos \omega t & -\sin \omega t \\ \sin \omega t & \cos \omega t \end{bmatrix} v \begin{bmatrix} \sqrt{3}/2 \\ 1/2 \end{bmatrix} \\ &= \frac{v}{2} \begin{bmatrix} \sqrt{3} \cos \omega t - \sin \omega t \\ \sqrt{3} \sin \omega t + \cos \omega t \end{bmatrix}. \end{aligned} \quad (1.250)$$

To reiterate, we have two representations of a single vector $\mathbf{v}$. In a computer programme, we must choose a coordinate system *and* a system of units; so, we will work with a length-2 array of numbers called, say, `v_unprimed_SI`.

A vector is an example of a larger set of entities called *tensors*. As with vectors, a tensor $\mathbf{T}$ is a uniquely defined object, and it has a representation $[\mathbf{T}]_{\text{unprimed}}$ in, say, unprimed coordinates. The components that make up $[\mathbf{T}]_{\text{unprimed}}$ can be written as a matrix for lower-order tensors; but no conventional way exists to write all of the numerical components of higher-order tensors similarly in some kind of tableau. Instead, the components of such higher-order tensors are conventionally written collectively with a compact notation that is not a tableau, and so does not specify the actual numerical values of those components. But that is another subject; the central point to remember here is that you should always distinguish between the uniquely defined quantity—be it a pure number, a dimensioned quantity, a vector, or a tensor—and its representation in some set of units and, if necessary, in some set of coordinates.

This chapter has seen little statistical mechanics, but now, at its end, you will have picked up a major set of tools needed to tackle the subject with ease. We have covered the ideas of counting, Stirling's rule, the mean and standard deviation, the gaussian function and its integrals, the meaning of infinitesimals, exact and inexact differentials, probability density, how to handle partial derivatives, and ideas of units versus dimensions. These concepts appear frequently in the chapters to follow.



Chapter 2

Accessible States and the Fundamental Postulate of Statistical Mechanics

In which we give the fundamental postulate of statistical mechanics, and count the microstates accessible to some basic yet important systems.

2.1 States and Microstates

Statistical mechanics begins with our ability to count the number of configurations that the constituents of a system can occupy. The configuration of a system is more usually called its *state*. When a system is isolated, all of the states available to it must have the same energy. When the details of a system's state are known in microscopic detail—to the extent that we could use them to recreate the system from scratch—then this “microscopically specified state” is called a *microstate*. A microstate is thus a state about which we have complete knowledge. The word “microstate” is used when we wish to remind ourselves or others that—at least in principle—we know everything about that state.

We do not always know all of the information necessary to describe a system, or perhaps we *can* know that information but have no need to; instead, a small set of parameters might be sufficient to represent the system for our purposes. These parameters might be the pressure, volume, and temperature of a gas, or the mass and temperature of a metal. In this broader sense, a possibly limited set of parameters that describes a system to some sufficient extent is sometimes said to describe a *macrostate* of that system. A macrostate of a system is a collective term that encompasses all states (called microstates!) that are consistent with the little that we know or care to know about the system.

For example, consider flipping two numbered coins. The outcomes could be listed as three *macrostates*: “both heads” (probability $1/4$), “both tails” (probability $1/4$), and “one of each” (probability $1/2$). But this system has four *microstates*, all with the same probability of $1/4$: *hh*, *tt*, *ht*, *th*.

Figure 2.1 shows another example: two of the microstates that make up a macrostate of a six-molecule water-ink mixture in one dimension, which we might call “left half: ink, right half: water”. The particles are treated as distinguishable. Particles 1–3 are ink, and particles 4–6 are water. There are

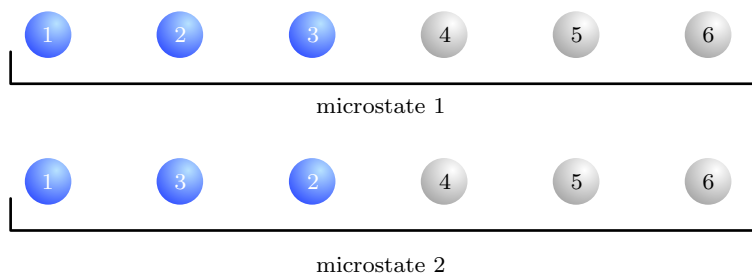


Fig. 2.1 Two *microstates* of a six-molecule water-ink mixture, differing only in the arrangement of their distinguishable particles. There is only one *macrostate* here, described as “left half: ink, right half: water”

$3! \times 3!$ microstates here, but only one macrostate for such a situation that can be described as “left half: ink, right half: water”.

When using the generic word “state”, whether we mean a microstate or a macrostate will come from the context of how much information we have about the system.

States are the bread and butter of statistical mechanics, and yet a precise definition of a state beyond the generic statements above is elusive and dependent on the application.¹ When we speak of isolated systems in great detail, the word “state” will always denote complete knowledge: a microstate. Later, when we study systems that are in contact with an environment, a microstate will refer to the entire (isolated) system-environment combination. But because we are then usually interested in the physics of the system and not its environment, the word “state” will be reserved for the system only.

Each microstate of the bathtub’s water-ink mixture in the previous chapter was a single arrangement of colours for its distinguishable particles, in which each particle’s position and colour were specified. Typically, we know nothing about the current motions of the particles that make up the water-ink mixture, and so macrostates of the mixture are, for example, “left half: ink, right half: water”, “evenly spread blue mixture”, and “almost evenly spread blue mixture with heavier blue spot at top right-hand corner”. A coin lying on a table occupies one of two possible microstates, or states: heads and tails. We’ll soon see that for an ideal gas, a state is defined as a cell in a higher-dimensional space defined by momentum and spatial position, with a cell “volume” determined by the number of internal parameters into which a gas particle can store its energy, and where the entire gas can be described as occupying one of these higher-dimensional cells. In the quantum mechanics of atoms, each atom occupies a single quantum state labelled by a set of quantum numbers. For the three-dimensional oscillators that form Einstein’s model of the heat capacity of a crystal, a state is one dimension of oscillation of a single oscillator. For Debye’s modification of Einstein’s model, a state

¹ A good example is the list of states of ever-increasing complexity in Section 7.6.

is a mode of oscillation of the entire crystal. We'll encounter each of these definitions of a state in turn.

But first, we must define some important terms necessary for the introduction of the fundamental postulate of statistical mechanics. An isolated system is said to be in *equilibrium* when the probabilities that it will be found in each of the states accessible to it are constant over time. The characteristic time needed for a perturbed system to attain equilibrium is called its *relaxation time*. Throughout this book, we'll assume our systems are always at, or arbitrarily close to, equilibrium. This assumption really means that all processes that the system might undergo occur over much longer time scales than its relaxation time. Thus, however the system changes, it always re-attains equilibrium very quickly—or at least very quickly compared to the typical time scale of the changes. Such “relatively slow” processes are called *quasi-static*, because the system can adjust so quickly that it effectively sees these processes as taking a very long time to play out.

Despite this restrictive-sounding definition, quasi-static processes are a very good approximation for a great many interesting systems. For example, consider the burning process happening inside the petrol engine of a car, that is running at R revolutions per minute. Each piston moves through one up–down motion R times per minute, which gives it an average speed of

$$\text{average speed} = \frac{\text{total distance travelled}}{\text{time taken}} = \frac{R \times (\text{up} + \text{down dist.})}{1 \text{ minute}}. \quad (2.1)$$

With a “stroke” distance of, say, 10 cm, the piston's average speed is

$$\frac{R \times 0.20 \text{ m}}{60 \text{ s}} = \frac{R}{300} \text{ m/s}. \quad (2.2)$$

At an engine speed of 2000 revolutions per minute, a piston's average speed is then about 7 m/s. Compare this with the average speed of 500 m/s for the air molecules inside the combustion chamber: from their perspective, the piston moves at a snail's pace! A refinement of this simple example might examine the density of the gas in the combustion chamber, since that affects how quickly the gas can change its configuration, such as when it burns. But the basic idea here is that the motion of the piston can be treated as quasi-static, and so statistical mechanics can be applied to analyse such engines very accurately.

The whole of statistical mechanics is based on the following statement:

The Fundamental Postulate of Statistical Mechanics

At any given moment, an isolated system in equilibrium is equally likely to be found in any of its microstates.

Note that the fundamental postulate does *not* say that an isolated system in equilibrium is equally likely to switch, during the next second, to any other of its microstates. A fly, when released from a corner of a room, will take some time to occupy a faraway position; but if we locate it a couple of minutes after releasing it, it is equally likely to be found anywhere in the room. Similarly, a dispersed ink drop in a bathtub is less likely to drastically alter its appearance during the next few seconds if that requires its molecules to undergo significant motion. But if we examine the ink and water molecules after a long time, then they are equally likely to be found in any possible arrangement. If we wish to examine them again to test the fundamental postulate, then, ideally, we should wait for an amount of time that allows the molecules to wander to any part of the bathtub.

The total number of microstates accessible to an isolated system (which thus has a fixed energy) is called Ω . The fundamental postulate says that the chance of the system being found in any particular microstate is $1/\Omega$. It follows that, to predict the behaviour of a very complex system in a statistical way, we might start by counting the number of available microstates that are consistent with its observed parameters, such as its total energy.

Consider that a system composed of three just-flipped coins can be found in any of $\Omega = 2^3 = 8$ possible microstates, corresponding to all permutations of heads and tails: $\{hhh, hht, hth, htt, thh, tht, tth, ttt\}$. What is the chance that exactly 2 of the coins have landed heads up (the “2-heads” macrostate)? This is another example of the binomial distribution: it corresponds to the last chapter’s example of a room, now with 3 particles, of which 2 are to be found in the front half of the room:

$$P(\text{two heads}) \stackrel{(1.11)}{=} C_2^3 (1/2)^2 (1/2)^1 = 3/8. \quad (2.3)$$

This is, of course, just the fraction of 3 elements in the heads/tails set a few lines up divided by the total number of 8 elements:

$$\{hhh, \underline{hht}, \underline{hth}, htt, \underline{thh}, tht, tth, ttt\}. \quad (2.4)$$

A standard phrasing here is “the 2-heads (macro)state is triply *degenerate*”, meaning that the macrostate with 2 heads up encompasses/corresponds to/is composed of 3 microstates. This number of microstates that are grouped into one macrostate is called the *degeneracy* of the 2-heads state.²

The fundamental postulate of statistical mechanics refers to the different microstates of an *isolated* system, and an isolated system’s energy is fixed. We can distinguish between three types of energy available to the system:

² This perhaps strange-looking word “degenerate” is not so strange when we see that it comes from the Latin “degeneratus”, meaning “no longer of its kind”. The implication is of something lost: in this case, the 3 microstates that each have 2 heads have been grouped together into a single (macro)state and have lost their individuality.

1. *Bulk kinetic energy* is carried by the system as a result of any bulk motion it might have. For example, a can of gas that is being carried at speed by a car has a bulk kinetic energy derived from the car's motion. This energy doesn't affect the statistical mechanics of the gas, and we don't include it when we say "the system has energy E ".
2. *Background potential energy* is inherited by the system from its environment; an example is gravitational potential energy. It is not altered by interactions between particles.
3. *Internal energy* is what can be transferred between the system's particles when they interact. This might be kinetic energy, or it might be chemical energy—which is really just potential energy. An example is when atoms bond to form a diatomic molecule, where their bond is treated as a spring that allows the molecule to oscillate. Both the kinetic and the potential energy of this oscillation are internal energy, because both can be transferred to another molecule in an interaction.

The number of states Ω accessible to a system is a function of its energy E , and so we usually write that number as $\Omega(E)$. This number of states is usually fantastically large, as we'll see soon. It tends to be extremely difficult, if not impossible, to count this number of states in which a system might be found. In fact, it can be conceptually easier to think in terms of $\Omega_{\text{tot}}(E)$, the total number of states available for all system energies up to and including E . We'll see why that's so in Section 2.5.

2.2 Energy Spacing of States

Up until now, the states that we have described have been discrete entities subject to the normal rules of counting and probability. But discrete entities can be difficult to treat mathematically: for example, in pure mathematics, theorems that restrict themselves to dealing with integers can be far more difficult to prove than the corresponding theorems that deal with real numbers. Continuous quantities are often easier to analyse, because for them, the machinery of calculus can be brought to bear. This is nothing new: we are used to describing the mass of an object as spread continuously throughout the object. It isn't spread continuously, of course; it is localised as atoms. But taking a continuous view is useful for many purposes, and, in fact, is quite necessary if we are to make headway in most areas of physics.

This principle that "continuous can be easier than discrete" also often applies to the states of a complex system. It turns out that in many situations, these states are very closely spaced in energy, and so to count them in a statistical treatment, we will profit from treating their size in some appropriate geometrical view as considerably smaller than the characteristic size of the

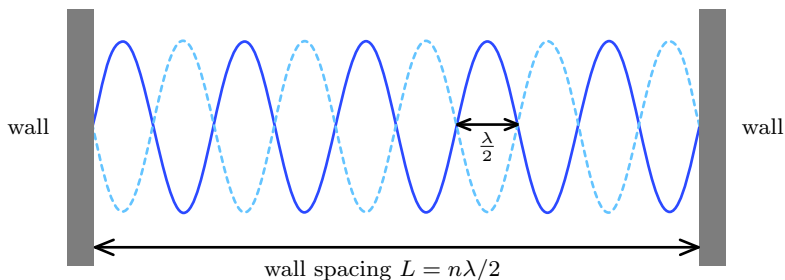


Fig. 2.2 Quantum mechanics prescribes the wavelengths of the basic modes (standing waves) of the wave function of a particle that is confined between two walls. These standing waves go to zero at the walls, and hence a natural number n of their half wavelengths must span the distance between the walls

“whole” geometry. This allows us, for example, to work with the volumes of higher-dimensional ellipsoids in the next sections, without worrying that a microscopic view of their surface would show it not to be smooth at all, but instead to consist of tiny steps.

Picture a gas of point particles in a room that is (without loss of generality) a cube. We assume there is no background potential, in which case the particles’ internal energy is their total energy, which is all kinetic. We will analyse a single one of these particles as a quantum-mechanical “particle in a box”, to calculate a quantum number that represents that total energy. We’ll then increase that quantum number by one and observe that the particle’s energy changes by an incredibly small amount, thus reinforcing the idea that the energies of the various microstates can be treated as a continuum. The maths is rendered simpler by working with a cubic room, but the general idea is unchanged for a room of arbitrary shape.

The task is then reduced to analysing a three-dimensional infinite-potential cubic well of side L . Quantum mechanics interprets the particle’s wave nature as such that in any region, the strength of the “de Broglie” wave associated with the particle—its “wave function”—quantifies the probability that the particle will be found in that region, and this wave function must vary continuously with position. The particle is confined within the box, and thus its wave function must go to zero at the walls. We can Fourier-decompose this wave function into its constituent sinusoids and follow the paradigm of quantum mechanics, which demands that those individual sinusoids must also vanish at the walls. What results is a set of waves with an associated discrete set of energies, any one of which the particle will be found to have when its energy is measured. A sinusoid for one such “energy eigenvalue” is shown in Figure 2.2. A natural number n of half wavelengths $\lambda/2$ of this sinusoid must fit into the box’s side length L ; that is, $n\lambda/2 = L$. Thus $\lambda = 2L/n$, and de Broglie’s relation between the wavelength and momentum of the particle that the wave represents is

$$p = \frac{h}{\lambda} = \frac{hn}{2L}, \quad (2.5)$$

where $h \simeq 6.626 \times 10^{-34}$ Js is *Planck's constant*. The particle's energy E_n is all kinetic:

$$E_n = \frac{p^2}{2m} = \frac{h^2 n^2}{8mL^2}. \quad (2.6)$$

This standard expression can also be produced by solving Schrödinger's equation for this scenario, and appears in any introductory book on quantum mechanics. Note that if the particle is to be in the box at all, n cannot be zero, since that would mean no wave was present; so, the particle must have an energy of at least E_1 . We will assume this value is close enough to zero in the coarse graining used ahead, but will return to it around (2.30).

Now consider each spatial dimension separately; for example, the x contribution to the energy is determined by the quantum number n_x , and similarly for the y and z directions, resulting in a total quantised energy of

$$E_{n_x n_y n_z} \equiv E_{n_x} + E_{n_y} + E_{n_z} \stackrel{(2.6)}{=} \frac{h^2(n_x^2 + n_y^2 + n_z^2)}{8mL^2}. \quad (2.7)$$

The quantum-mechanical state described by the numbers n_x, n_y, n_z is a microstate of the system. For our purpose of calculating the particle's energy increase for one of these numbers being increased by one, it suffices to examine the state whose three quantum numbers are set equal to some n : $n_x = n_y = n_z = n$. We focus on the energy E_{nnn} of the particle in this state:

$$E_{nnn} = \frac{3h^2 n^2}{8mL^2}. \quad (2.8)$$

Suppose now that this particle is merely one of a great number of particles in the box, each with this same energy. For the purpose of this example, we'll borrow from the definition of temperature (with its SI units, kelvins) and the results of Section 3.5 ahead, to say that when this gas has a temperature T , the energy E_{nnn} equals $3/2 kT$, where $k = 1.381 \times 10^{-23}$ J/K is Boltzmann's constant. Now equate these thermodynamic and quantum expressions for energy:

$$\frac{3kT}{2} = \frac{3h^2 n^2}{8mL^2}, \quad (2.9)$$

which yields

$$n = \frac{2L}{h} \sqrt{mkT}. \quad (2.10)$$

For a touch of realism, we'll set the particle to be an air molecule. We need its mass m . Air's molar mass is easily estimated from a basic knowledge of chemistry. Air is about 78% N₂, 21% O₂, and 1% Ar. A typical nitrogen atom has 7 protons and 7 neutrons, leading to a molar mass of 14 g for atomic nitrogen, or 28 g for N₂. Similarly, a typical oxygen atom has 8 protons

and 8 neutrons, giving O_2 a molar mass of 32 g. A typical argon atom has 18 protons and 22 neutrons, giving Ar a molar mass of 40 g. Air's molar mass is then $(0.78 \times 28 + 0.21 \times 32 + 0.01 \times 40)$ g = 29.0 g. Dividing this by Avogadro's number 6.022×10^{23} gives the mass of an air molecule as being about $m = 4.8 \times 10^{-26}$ kg.

Working in SI units with a comfortable temperature of $T = 298$ K in a room of side length $L = 5$ m, equation (2.10) produces

$$n = \frac{2 \times 5}{6.626 \times 10^{-34}} \sqrt{4.8 \times 10^{-26} \times 1.381 \times 10^{-23} \times 298} \simeq 2.1 \times 10^{11}. \quad (2.11)$$

This is a very large number of half wavelengths: the de Broglie wavelength of the particle is minuscule. (This is the quantum-mechanical version of the general observation that if a wave's wavelength is comparable with the size of the system in which it travels, then it will show all the usual properties and behaviour of a wave; but if its wavelength is much smaller, it will act like a particle. Thus, radio waves behave very much as waves, whereas light waves resemble a stream of particles, and so tend to be called light *rays* instead.)

Now, what energy increase ΔE results when we increase one of the three quantum numbers by one? Work in SI units, but recall (1.216) to convert the answer to the more usual electron volts by dividing by 1.602×10^{-19} :

$$\begin{aligned} \Delta E &\equiv E_{n+1,nn} - E_{nnn} = \frac{h^2}{8mL^2} [(n+1)^2 + n^2 + n^2 - 3n^2] \\ &\simeq \frac{h^2 n}{4mL^2} = \frac{(6.626 \times 10^{-34})^2 \times 2.1 \times 10^{11}}{4 \times 4.8 \times 10^{-26} \times 25 \times 1.602 \times 10^{-19}} \text{ eV} \\ &\simeq 1.2 \times 10^{-13} \text{ eV}. \end{aligned} \quad (2.12)$$

How does this compare with the particle's typical energy? Again, we use the value $3/2 kT$ for this, derived in Section 3.5.2:

$$\frac{3kT}{2} \simeq \frac{3 \times 1.381 \times 10^{-23} \times 298}{2 \times 1.602 \times 10^{-19}} \text{ eV} \simeq 0.04 \text{ eV}. \quad (2.13)$$

We see that the energy spacing ΔE is *minuscule* compared with the particle's kinetic energy. In fact, the use of the numbers above has perhaps obscured the simple expression for the ratio of energy spacing to kinetic energy:

$$\frac{\Delta E}{3kT/2} = \frac{h^2 n / (4mL^2)}{3h^2 n^2 / (8mL^2)} = \frac{2}{3n}, \quad (2.14)$$

and recall that $n = 2.1 \times 10^{11}$ here. This tiny energy spacing suggests that we will make only a negligible error if we treat a system's energy as continuous.

This validates the use of calculus in these systems, such as differentiating a quantity with respect to energy.

2.3 Position–Momentum and Phase Space

Advanced classical mechanics studies systems by way of the position and *canonical momentum*³ of each of their constituent particles. In the first instance, these particles are treated as *distinguishable*, meaning that, in principle, they can all be individually labelled and counted.

The classic example of a position–momentum analysis is that of a single particle undergoing *simple harmonic motion*. In such motion, the particle is acted on by a force proportional to its distance from some origin, in the direction of that origin. In one dimension, this motion is the particle’s displacement $x(t)$ from the origin as a function of time, which turns out to be oscillatory. This displacement is

$$x = A \cos(\omega t + \phi), \quad (2.15)$$

where A is the amplitude of the oscillation, ω is the angular frequency of that oscillation, and ϕ is an angle set by the initial conditions. The particle’s velocity is

$$v = dx/dt = -\omega A \sin(\omega t + \phi). \quad (2.16)$$

Momentum plays a stronger role in classical mechanics than velocity.⁴ Our oscillating particle of mass m has momentum

$$p = mv = -m\omega A \sin(\omega t + \phi). \quad (2.17)$$

It’s clear that

$$\frac{x^2}{A^2} + \frac{p^2}{(m\omega A)^2} = 1, \quad (2.18)$$

from which we can see that if the particle’s motion is plotted as p versus x , its motion can be represented by a dot that traces an ellipse clockwise in time. For a single particle, this position–momentum space is called the particle’s *phase space*; the particle’s cyclic trajectory through it appears in the left-hand picture in Figure 2.3. The purple dot marks the point $[x(t), p(t)]$, and traces the ellipse clockwise at a varying speed. Note that the vector to this dot from the origin does *not* generally make an angle $\omega t + \phi$ with the x axis. If we geometrize the motion by scaling p to have the same dimension as x

³ A particle’s canonical momentum is very often just its mass times its velocity, but not always: a notable exception is a charged particle in a magnetic field. The canonical momentum is calculated in a standard way from the system’s lagrangian, which you can find described in books on classical mechanics.

⁴ Again, the two are not simply related for a charged particle in a magnetic field.

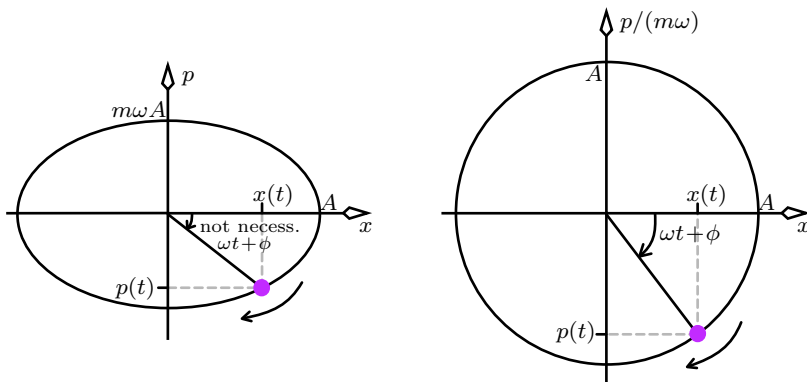


Fig. 2.3 A phase-space portrait of simple harmonic motion. The left-hand picture shows p versus x , but the purple dot doesn't trace the ellipse at constant speed. The right-hand picture replaces p with $p/(m\omega)$, which has dimensions of length. Now the purple dot traces a circle at constant speed

(which is length, of course), a simpler picture results, shown on the right-hand side in Figure 2.3. This is a plot of $p/(m\omega)$ versus x , and the motion of the purple dot is now circular at constant speed. The vector to this dot from the origin now *does* make the angle $\omega t + \phi$ with the x axis, and the motion is clearer geometrically. Nevertheless, the plot of p versus x encodes all of the information about the particle's simple harmonic motion.

Just as the space we live in is always three dimensional in its fullest form (with axes x, y, z), position–momentum space for *point* particles is always really *six* dimensional, with axes x, y, z, p_x, p_y, p_z . When multiple point particles interact, their combined motion can be represented on the six-dimensional space by plotting one point for each particle. So, when two particles interact, we plot two points on the same set of six position–momentum axes, and each of these points follows some trajectory in time.

But classical mechanics goes further than such a picture. Given N point particles, it allocates an *individual* set of six position–momentum variables to each particle, then combines the resulting six axes for each particle into a single set of $6N$ axes, and plots a *single* point that represents the position–momentum of all the particles. This point now follows a trajectory through time that portrays the evolution of the entire system of particles. The set of $6N$ axes now defines the phase space of all the particles. When dealing with point particles, we will distinguish *position–momentum space*—with its three position and three momentum axes—from *phase space*, with its $3N$ position and $3N$ momentum axes. In the case of a system composed of a single point particle such as our simple harmonic oscillator above, the phase space is identical to the position–momentum space.

To demonstrate, consider two point particles. Particle 1 follows a trajectory in position–momentum space; we take a copy of that and relabel the six

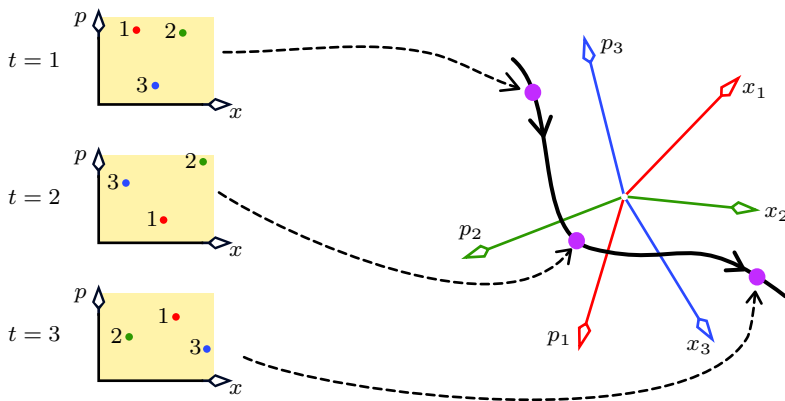


Fig. 2.4 **Left:** Position–momentum space at three closely-spaced instants, carrying three particles that each move in the same single spatial dimension. **Right:** Give each particle its own copy of that position–momentum space (e.g., the “red” particle 1 has the red axes at right). Then combine the three position–momentum spaces into a single six-dimensional space called “phase space”: the six mutually orthogonal axes of this phase space can only be imagined in the picture. The three particles are now represented at any instant by a single point in that phase space, drawn in purple. Their combined motion is described by the trajectory of the purple point through the phase space over time

axes as “ $x_1, y_1, z_1, p_{x1}, p_{y1}, p_{z1}$ ”. (The momentum notation here of, say, p_{x1} , means “ p_x for particle 1”.) Particle 2 follows a trajectory in the *same* position–momentum space (there is only ever a single position–momentum space!), and we take a copy of that space and relabel the six axes as “ $x_2, \dots, p_{z2}$ ”. The motions of both particles are represented by a single point in the 12-dimensional phase space whose axes are $x_1, \dots, p_{z2}$. We have no way of picturing this; it’s not even possible to draw the phase space of the simplest multi-particle case of two particles moving in one spatial dimension, with its four axes x_1, x_2, p_1, p_2 .

Despite such pictorial difficulties, Figure 2.4 is a schematic of three particles moving in the same single spatial dimension. On its left, we see three successive “movie frames” (at times $t = 1, 2, 3$) of the motion of the three particles in position–momentum space. Now give each particle its own copy of the position–momentum space, creating $x_1 p_1$ space for particle 1, $x_2 p_2$ space for particle 2, and $x_3 p_3$ space for particle 3. Combine these three position–momentum spaces into a single six-dimensional phase space (whose mutually orthogonal axes cannot really be depicted). In that phase space, the three particles are represented by a *single point*. The locus of these points over time is a trajectory through the phase space.

When the particles have internal structure, we allow for their rotation and internal oscillation by extending the position–momentum space to more variables. Rotation is represented by three angles that describe a particle’s spatial

orientation, and three angular momenta that say how it is spinning.⁵ Internal oscillation requires three more spatial variables that describe how stretched or compressed the particle is along each of three internal spatial axes, and three linear momenta that describe the particle's internal oscillation.

To begin to discuss the microstates of a system of possibly interacting particles, we introduce the idea of “tiling” position–momentum space into “cells”, whose higher-dimensional volume is given by allocating a factor of Planck's constant h for each pair of position–momentum variables. This idea is a tip of the hat to quantum mechanics, because it invokes Heisenberg's uncertainty principle to acknowledge that each pair of position–momentum variables cannot be considered to encode a particle's position and momentum to an arbitrarily fine accuracy.

Here is an example. For simplicity, treat the air molecules in a cubic room 5 metres on each side at a representative temperature of $T = 298$ K as point particles, and ask the following questions:

- How many particles are in the room?
- How many cells are in the six-dimensional position–momentum space?
- Do we expect any “crowding” of more than one particle in some cells?

Atmospheric pressure is $P = 101,325$ pascals. The number N of particles is given by the ideal-gas law $PV = NkT$ (proved later). Using SI units, we have

$$N = \frac{PV}{kT} = \frac{101,325 \times 5^3}{1.381 \times 10^{-23} \times 298} \simeq 3.08 \times 10^{27}. \quad (2.19)$$

The number of cells in position–momentum space is

$$\text{number of cells} = \frac{L_x L_{p_x} L_y L_{p_y} L_z L_{p_z}}{h^3}, \quad (2.20)$$

where the room has lengths in each spatial dimension of $L_x = L_y = L_z = 5$ m, and the “lengths” of the occupied parts of the momentum space are given by $L_{p_x}, L_{p_y}, L_{p_z}$. What are these “lengths”? At the above temperature, air molecules move with a range of speeds from zero to about 600 m/s. Without being very careful here about the particular form of this speed distribution,⁶ we'll interpret that speed range as a velocity in each spatial direction that can have any value in, say, the range -350 m/s to 350 m/s, since the speed corresponding to a velocity of $(350, 350, 350)$ m/s is about 600 m/s. Air molecules have a mass of about 4.8×10^{-26} kg, and so

⁵ Suppose an object spins with angular velocity ω about an axis described by a unit vector $\mathbf{n}$. Write its angular velocity vector $\boldsymbol{\omega} \equiv \omega \mathbf{n}$ in component form for a set of cartesian axes as $(\omega_x, \omega_y, \omega_z)$. An intriguing and useful result of the theory of rotation then says that the object can be considered to be spinning around each of the x, y, z axes concurrently, with angular velocities $\omega_x, \omega_y, \omega_z$, respectively.

⁶ We'll be more careful in Chapter 6, when studying the Maxwell distribution.

$$L_{p_x} = L_{p_y} = L_{p_z} = (350 - -350) \times 4.8 \overset{-26}{=} \text{kg m/s}^2. \quad (2.21)$$

The number of cells in position–momentum space is then

$$\text{number of cells} = \left(\frac{5 \times 700 \times 4.8 \overset{-26}{=}}{6.626 \overset{-34}{=}} \right)^3 \simeq 1.63 \overset{34}{=}. \quad (2.22)$$

The number of cells per particle is

$$\frac{\text{number of cells}}{\text{number of particles}} = \frac{1.63 \overset{34}{=}}{3.08 \overset{27}{=}} \simeq 5.3 \text{ million}. \quad (2.23)$$

With this vast number of cells available for each particle, there is certainly no crowding in position–momentum space: most cells are not occupied at all, and so the number of cells with more than one particle will be minuscule. We’ll use this result in Section 2.4.2.

2.4 Microstates Are Cells of Phase Space

Each of the three time-slices at the left in Figure 2.4 shows the particles scattered into three of the huge number of cells in position–momentum space—with usually no more than one particle per cell. Each of these time-slices shows a *single* distinct microstate of the system of particles; equivalently, each *cell* of the corresponding phase space at the right of the same figure denotes a distinct microstate of the system. Now we know how to count a system’s microstates: we count the cells in its phase space.

Despite that such a counting of phase-space cells doesn’t seem to resemble the “particle in a box” view of Section 2.2, it does yield the same result for the number of microstates at a given energy, as we’ll soon see. This is a powerful motive for allowing h to set the size of a phase-space cell. In fact, we’ll see later, in Section 3.8.4, that using any multiple of h to set this cell size is actually sufficient for calculations involving increases in *entropy*, a quantity of prime importance introduced in Chapter 3. But despite this, we’ll see in the discussion around (2.28) ahead that allowing just h to set the cell size will give an internal consistency to various calculations.

We have now arrived at a starting point for counting microstates. The idea is that the total number of microstates accessible to a system of energy E equals the higher-dimensional volume of phase space that the system can “explore” divided by the (higher-dimensional) volume of one cell. Suppose that the position–momentum space available to each of a system’s particles—the space coloured yellow in each time slice on the left in Figure 2.4—has D pairs of position–momentum coordinates (and thus $2D$ axes). In that figure, $D = 1$ because a single pair of “ xp ” axes suffices to show the motions of all the

particles in the left-hand sequence of time slices. With N particles present, N “instances” of these D pairs are used to construct the phase space on the right in Figure 2.4; thus, a total of DN pairs of variables are used to build the phase space. Figure 2.4 has $N = 3$ particles, and hence the phase space has $DN = 1 \times 3$ *pairs* of dimensions, as indicated by its six axes at the right in the figure.

The cell volume in phase space is formed by allocating one factor of h for each pair of position–momentum dimensions present; so, this cell volume is h^{DN} . The number of microstates accessible to a system with energy E is

$$\begin{aligned}\Omega(E) &\equiv \text{number of microstates that each “have” energy } E \\ &= \frac{\text{volume of phase space}}{h^{DN}} = \frac{1}{h^{DN}} \int_{\text{energy } E} dx^{DN} dp^{DN},\end{aligned}\quad (2.24)$$

where dx^{DN} and dp^{DN} denote the integral over all DN position variables and all DN momentum variables defined for the system, and we quote the “have” in (2.24) to highlight that some idea of coarse graining over energy is implied, because the energy in the systems we are analysing here is continuous. After all, if the energy were truly continuous, then the number of microstates that each had *exactly* a given value of E would be zero, if the total number of microstates was finite. We introduced that idea for the display on a clock in Section 1.6.2, and we’ll discuss it again in Section 2.5. For now, we’ll sidestep this necessary granularity in E by instead calculating the total number of microstates for all energies *up to* E :

$$\begin{aligned}\Omega_{\text{tot}}(E) &\equiv \left[\begin{array}{l} \text{total number of microstates, each of} \\ \text{which has some energy in 0 to } E \end{array} \right] \\ &= \frac{1}{h^{DN}} \int_{\text{energies 0 to } E} dx^{DN} dp^{DN}.\end{aligned}\quad (2.25)$$

In the remainder of this section, we calculate Ω_{tot} for simple systems of ever-increasing complexity. Each of these systems consists of particles that have non-zero rest mass. We’ll examine massless particles in Section 2.6.

Ω_{tot} for a Free Point Particle in One Dimension

Begin with the simplest case: a single (so $N = 1$) free massive point particle of energy E constrained to move on a line or curve of length L . The particle’s motion is described by a single pair of position–momentum variables x, p ; hence $D = 1$, and the integral (2.25) is simply written over $dx dp$. The space variable x ranges over all values from 0 to L . We are calculating Ω_{tot} rather than Ω , in which case we consider all cases of energy from 0 to E —which corresponds to momentum p anywhere in the range $-\sqrt{2mE}$ to $\sqrt{2mE}$. Be-

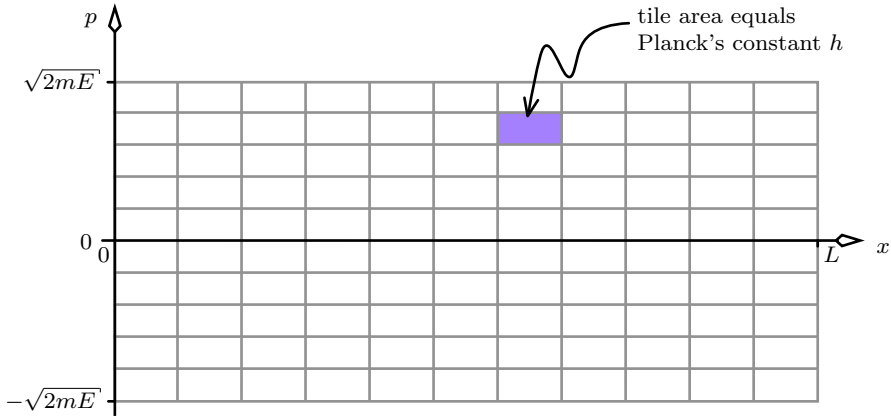


Fig. 2.5 Tiling of the phase space of a single free massive particle confined to a length L in one dimension, and that can have any energy from 0 to E . The ratio of width to height of the tiles—or indeed, even their general shapes—is immaterial: only their area h is fixed

cause x and p are unrelated, the integrals in (2.25) decouple, and integrating over each becomes trivial:

$$\int dx = L, \quad \int dp = 2\sqrt{2mE}. \quad (2.26)$$

Equation (2.25) then yields

$$\Omega_{\text{tot}}(E) = \frac{L2\sqrt{2mE}}{h}. \quad (2.27)$$

Note that by calculating Ω_{tot} instead of Ω , we have avoided the need to focus on the details of the coarse graining referred to just after (2.24) above. Figure 2.5 shows the particle's phase space, with its tiles of area h .

How does this expression (2.27) compare with the single-particle quantum approach of Section 2.2, where we had set the number of microstates for energies in 0 to E to be the quantum number n in (2.6)? The number of microstates obtained from the quantum analysis is

$$\Omega_{\text{tot}}^{\text{quant}}(E) = n \stackrel{(2.6)}{=} \sqrt{\frac{8mL^2E}{h^2}} = \frac{L2\sqrt{2mE}}{h}, \quad (2.28)$$

which exactly matches (2.27). So, the two ways of defining microstates, the “quantum particle in a box” of Section 2.2 and the “phase space with cell size set by h ” of the current section, give the same result for the total number of microstates available to a single particle with energy in the range 0 to E . This agreement happened partly by chance, because in this section's phase-space

picture, we were under no obligation to set the phase-space tile area to be exactly h . But we see now that this choice of tile area was a good one.

What is the number of states $\Omega_{\text{tot}}(E)$ for a single particle in a one-dimensional room of length 5 metres, whose speed is typical of air molecules at room temperature? Use $m = 4.8 \times 10^{-26}$ kg, give the particle a maximum speed of $v = 350$ m/s as per the discussion just after (2.20), and write $E = \frac{1}{2}mv^2$. Equation (2.27) becomes

$$\begin{aligned}\Omega_{\text{tot}} &= \frac{5 \times 2 \sqrt{2 \times 4.8 \overset{-26}{\text{---}} \times \frac{1}{2} \times 4.8 \overset{-26}{\text{---}} \times 350^2}}{6.626 \overset{-34}{\text{---}}} \\ &\simeq 2.5 \times 10^{11}.\end{aligned}\tag{2.29}$$

[This is a similar calculation to what we did previously in (2.11), but there, we used a preliminary notion of temperature to set the particle's energy. Here, we are using a representative maximum speed of the particle to give us that energy: this speed does equate roughly to the room temperature used in (2.11).] The point here is that $\Omega_{\text{tot}}(E)$ is a very large number, and this means that the tiles drawn in Figure 2.5 are far smaller in reality than the schematic size drawn in that figure.

As stated just after (2.6), equation (2.27) cannot hold for arbitrarily small values of energy E . This is so because the de Broglie wavelength h/p of the particle must be no greater than the constraining length L —and preferably a lot less than L . That is,

$$h/p \ll L, \tag{2.30}$$

or $Lp/h \gg 1$, meaning

$$\frac{L\sqrt{2mE}}{h} \gg 1. \tag{2.31}$$

Comparing this with (2.27), we see that (2.31) is equivalent to demanding that $\Omega_{\text{tot}} \gg 2$. This rules out any consideration of taking the limit $E \rightarrow 0$ in (2.27).

Ω_{tot} for a Free Point Particle in Three Dimensions

Here again, the single (i.e., $N = 1$) free massive point particle has energy E , but again, to calculate $\Omega_{\text{tot}}(E)$ rather than $\Omega(E)$, we consider all energies in 0 to E . The particle now resides in three spatial dimensions, so we need $D = 3$ pairs of position–momentum variables to describe its motion: xp_x , yp_y , and zp_z . We will not attempt to draw the six-dimensional version of Figure 2.5. But we *can* draw the space and momentum axes separately, in Figure 2.6. The tiles in Figure 2.5 are now six-dimensional cells of volume $h^{DN} = h^3$, with one factor of h coming from each of the three pairings xp_x , yp_y , and zp_z . Just as in the one-dimensional case, the integrals in (2.25) decouple:

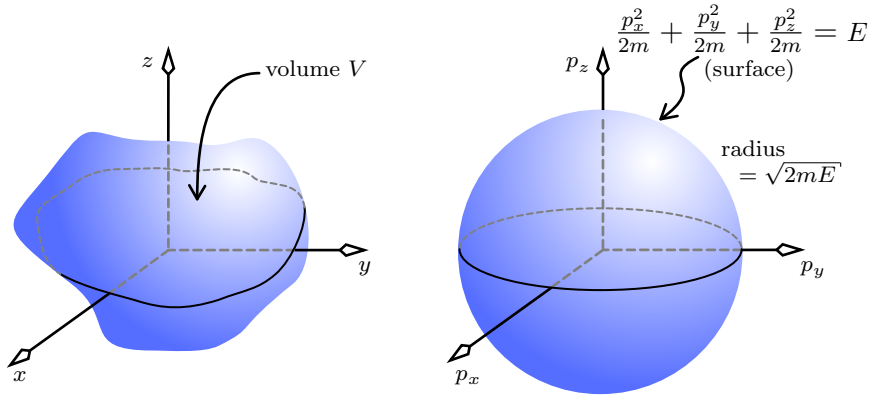


Fig. 2.6 A representation of the space (left) and momentum (right) aspects of what is really the six-dimensional phase space of a single free particle moving in three spatial dimensions with some energy in 0 to E . The tiles in Figure 2.5 have become six-dimensional cells that straddle the position and momentum spaces, and so cannot be drawn here

$$\begin{aligned}
 \int dx^3 &= \int dx dy dz = V, \\
 \int dp^3 &= \int dp_x dp_y dp_z = \text{volume of sphere in momentum space} \\
 &= \frac{4}{3} \pi (2mE)^{3/2}.
 \end{aligned} \tag{2.32}$$

Equation (2.25) then gives

$$\Omega_{\text{tot}}(E) = \frac{V 4\pi (2mE)^{3/2}}{3h^3}. \tag{2.33}$$

Similar to the one-dimensional case in (2.29), let's calculate $\Omega_{\text{tot}}(E)$ for a single particle moving at a typical speed of air molecules, now in a three-dimensional cubic room of side length 5 metres. Use $m = 4.8 \times 10^{-26}$ kg, give the particle a maximum speed of $v = 350$ m/s, and write $E = \frac{1}{2}mv^2$. Equation (2.33) becomes

$$\begin{aligned}
 \Omega_{\text{tot}} &= \frac{5^3 \times 4\pi \left(2 \times 4.8^{\frac{-26}{}} \times \frac{1}{2} \times 4.8^{\frac{-26}{}} \times 350^2 \right)^{3/2}}{3 \times (6.626^{\frac{-34}{}})^3} \\
 &\simeq 1.7 \times 10^{33}.
 \end{aligned} \tag{2.34}$$

As usual, this is a *very* large number. [It is similar to (2.22), but not identical, because (2.22) didn't use a spherical region in momentum space.] The three-dimensional analogy to Figure 2.5's tiling of one-dimensional phase space cannot be drawn, since it requires six dimensions; but we might have supposed

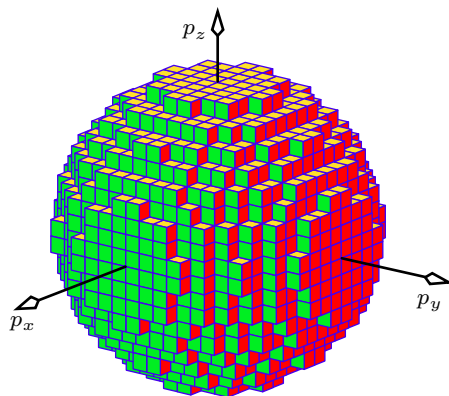


Fig. 2.7 A visual aid for filling three-dimensional momentum space with cells, in analogy to the tile widths along the momentum axis of the one-dimensional picture in Figure 2.5. As we show in (2.34), the cells drawn above, while useful as a mental picture, are far too coarse to represent reality. In an everyday example of particle motion, the cells would be so incredibly tiny that they would fill the blue volumes in Figure 2.6 very completely, with no left-over space worth considering

that if we could isolate the momentum-space part of it and draw a kind of projection of the cells onto that, a picture such as that drawn in Figure 2.7 would emerge. That picture is useful as a visual aid for the idea of dividing the phase space into cells; but we see from the huge size of Ω_{tot} in (2.34) that, in reality, the cells are tiny—so tiny that they fill the position and momentum spaces drawn in Figure 2.6 with no left-over gaps to speak of.

Just as for the free particle, de Broglie would demand that

$$h/p \ll V^{1/3}, \quad (2.35)$$

or $h^3/p^3 \ll V$, or $Vp^3/h^3 \gg 1$, which is

$$\frac{V(2mE)^{3/2}}{h^3} \gg 1. \quad (2.36)$$

Comparing this with (2.33), we see the demand that $\Omega_{\text{tot}} \gg 4$. So again, we cannot consider the small- E limit of (2.33).

Following our comparison with the “quantum particle in a one-dimensional box” analysis of one space dimension around (2.28), compare (2.33) with the result of counting the number of states using the quantum numbers for *three* space dimensions in (2.7). The quantum states are represented by cubes of unit side length in $n_x n_y n_z$ -space; and although n_x, n_y, n_z cannot *all* be zero, any two of them can be zero. Thus, we need only omit the cube allocated to the origin of that space—but, since the space contains an enormous number of cubes, we can even ignore the fact that this cube should be omitted. If each of the cubes is defined to be a microstate in the quantum analysis, the

number of microstates for energies 0 to E in the quantum analysis is thus the volume of one octant of a sphere in $n_x n_y n_z$ -space. The radius of this octant is $\sqrt{n_x^2 + n_y^2 + n_z^2}$, where, for a given energy E , equation (2.7) sets

$$E = \frac{h^2(n_x^2 + n_y^2 + n_z^2)}{8mL^2}. \quad (2.37)$$

The number of microstates obtained from the quantum analysis is then

$$\Omega_{\text{tot}}^{\text{quant}}(E) = \frac{1}{8} \times \frac{4\pi}{3} \left(\sqrt{n_x^2 + n_y^2 + n_z^2} \right)^3. \quad (2.38)$$

The volume of the box is $V = L^3$, and so (2.38) combines with (2.37) to give

$$\Omega_{\text{tot}}^{\text{quant}}(E) = \frac{\pi}{6} \left(\frac{8mL^2 E}{h^2} \right)^{3/2} = \frac{V 4\pi (2mE)^{3/2}}{3h^3} \stackrel{(2.33)}{=} \Omega_{\text{tot}}(E). \quad (2.39)$$

Thus, just as in the one-space-dimensional case, in three space dimensions, the two ways of defining microstates (the “quantum particle in a box” and the “phase space with cell size set by h ”) give the same result for each picture’s definition of the number of microstates available to a single particle with energy in the range 0 to E . Again, we see that the choice of h to set the cell volume was a good one, and so we’ll drop further reference to “ $\Omega_{\text{tot}}^{\text{quant}}$ ”.

So much for one particle. We examine a gas of N non-interacting point particles next.

Ω_{tot} for an Ideal Gas of Point Particles

An *ideal gas* is a set of non-interacting particles that are free to move spatially and whose energy is all kinetic. Whilst they don’t interact with each other, they *do* interact with the walls of their container, and so they can rotate or oscillate if they are not point particles; but no potential energy is associated with their separations.

It’s straightforward to extend the above calculations of Ω_{tot} to an ideal gas of N distinguishable massive *point* particles. The particles have total energy E (purely translational) and occupy a spatial volume V . They can each move in three dimensions; thus $D = 3$, and (2.25) requires

$$\begin{aligned} dx^{3N} &\equiv \underbrace{dx_1 dy_1 dz_1}_{\text{particle 1}} \dots \underbrace{dx_N dy_N dz_N}_{\text{particle } N}, \\ dp^{3N} &\equiv \underbrace{dp_{x1} dp_{y1} dp_{z1}}_{\text{particle 1}} \dots \underbrace{dp_{xN} dp_{yN} dp_{zN}}_{\text{particle } N}. \end{aligned} \quad (2.40)$$

Just as we saw for the free point particle analysed earlier in this section, the momentum of each particle of the ideal gas is not related to its position, and this allows the position and momentum integrals in (2.25) to decouple. Each particle's contribution to the position integral is also independent of the other particles, and so

$$\int dx^{3N} = \int dx_1 dy_1 dz_1 \dots \int dx_N dy_N dz_N = V^N. \quad (2.41)$$

Integrating the momentum requires more thought. As usual, to calculate Ω_{tot} , we allow the gas's total energy to have any value in the range 0 to E . Suppose (without loss of generality) that all particles have the same mass m . The momentum-space integral $\int dp^{3N}$ becomes the volume of the multi-dimensional sphere described by

$$\frac{p_{x1}^2}{2m} + \frac{p_{y1}^2}{2m} + \frac{p_{z1}^2}{2m} + \dots + \frac{p_{xN}^2}{2m} + \frac{p_{yN}^2}{2m} + \frac{p_{zN}^2}{2m} = E. \quad (2.42)$$

This hypersphere is the higher-dimensional generalisation of the sphere at the right in Figure 2.6, residing in a momentum space of $3N$ dimensions (part of the $6N$ -dimensional phase space) and with a “radius” of $R = \sqrt{2mE}$. Determining the volume of such a hypersphere in d dimensions $x_1^2 + \dots + x_d^2 = R^2$ is a standard calculation in advanced calculus; the result is

$$\text{volume} = \frac{\pi^{d/2} R^d}{(d/2)!}. \quad (2.43)$$

Verify this formula for $d = 1, 2, 3$, using $(1/2)! = \sqrt{\pi}/2$. In one dimension ($d = 1$), the “sphere” is a line of half-length R , and its “volume” is its length. In two dimensions ($d = 2$), the “sphere” is a disc of radius R , and its “volume” is its area. In three dimensions, we have a normal sphere of radius R .

Equation (2.43) gives the following volume in momentum space of the hypersphere (2.42), using $d = 3N$ and $R = \sqrt{2mE}$:

$$\int dp^{3N} = \frac{\pi^{3N/2} (\sqrt{2mE})^{3N}}{(3N/2)!} = \frac{(2\pi mE)^{3N/2}}{(3N/2)!}. \quad (2.44)$$

Equation (2.25) now combines (2.41) and (2.44) to give us the total number of states for all energies up to E :

$$\Omega_{\text{tot}}(E) = \frac{1}{h^{3N}} \int_{\text{energies 0 to } E} dx^{3N} dp^{3N} = \frac{V^N (2\pi mE)^{3N/2}}{h^{3N} (3N/2)!}. \quad (2.45)$$

Remember that the “3” in the expression $3N$ above results from each of the particles contributing 3 terms of the form “ $p^2/(2m)$ ” to the gas’s total energy in (2.42); that is, particle i contributes the sum

$$\frac{p_{xi}^2}{2m} + \frac{p_{yi}^2}{2m} + \frac{p_{zi}^2}{2m}. \quad (2.46)$$

And, we have learned by now not to take seriously any small- E limit of (2.45).

It’s an easy exercise to show that setting $N = 1$ in (2.45) reproduces the single-particle result (2.33), as we might expect. Setting $N = 1$ in (2.45) gives [using $(3/2)! = 3\sqrt{\pi}/4$]

$$\Omega_{\text{tot}}(E) = \frac{V(2\pi mE)^{3/2}}{h^3 3\sqrt{\pi}/4} = \frac{V4\pi(2mE)^{3/2}}{3h^3}, \quad (2.47)$$

which is (2.33) again.

To gain a feel for the size of Ω_{tot} , write it first in a more convenient form. For shorthand, set $\gamma \equiv 3N/2$ in the next line only, then note that when N is large, Stirling’s rule approximates $\gamma!$ in (2.45) as

$$\gamma! \approx \gamma^{\gamma+1/2} e^{-\gamma} \sqrt{2\pi} \simeq (\gamma/e)^{\gamma} \sqrt{2\pi}. \quad (2.48)$$

Equation (2.45) then becomes

$$\Omega_{\text{tot}}(E) = \frac{V^N}{\sqrt{2\pi}^N} \left(\frac{4\pi emE}{3Nh^2} \right)^{3N/2} \quad (N \text{ large}). \quad (2.49)$$

Just as we did in Section 2.2, we’ll use the result of Section 3.5 ahead to set $E = N^{3/2} kT$ here, where k is Boltzmann’s constant. Also, the large size of N allows the $\sqrt{2\pi}$ to be omitted—remember that 2π appears in (2.49) as an overall factor of $(2\pi)^{3N/2-1/2}$, and N is typically 10^{27} . Equation (2.49) becomes

$$\Omega_{\text{tot}}(E) \simeq V^N \left(\frac{2\pi emkT}{h^2} \right)^{3N/2} \quad (N \text{ large}). \quad (2.50)$$

To appreciate how truly colossal Ω_{tot} is, calculate it for a room full of monatomic gas particles. The room is a cube of side 5 m with $N = 10^{27}$ distinguishable particles at 298 K, each with mass m equal to the average mass of an air molecule (4.8×10^{-26} kg). (Air molecules are diatomic, not monatomic, and we’ll treat these in a moment. Their diatomicity only increases the magnitude of the result that we are about to calculate.) Taking a logarithm to deal better with the large numbers, working in SI units, and recalling the discussion in Section 1.9.2 that explains the logic behind apparently taking logarithms of dimensioned quantities, (2.50) becomes

$$\begin{aligned}
\log_{10} \Omega_{\text{tot}} &\approx N \log_{10} V + \frac{3N}{2} \log_{10} \frac{2\pi emkT}{h^2} \\
&= 10^{27} \log_{10} 125 + 1.5 \frac{27}{\text{Å}} \log_{10} \frac{2\pi e \times 4.8 \frac{\text{Å}}{\text{Å}} \times 1.381 \frac{\text{Å}}{\text{Å}} \times 298}{(6.626 \frac{\text{Å}}{\text{Å}})^2} \\
&\simeq 3.5 \times 10^{28},
\end{aligned} \tag{2.51}$$

so that

$$\Omega_{\text{tot}} \approx 10^{3.5 \times 10^{28}} \text{ microstates.} \tag{2.52}$$

If we write $10^{3.5 \times 10^{28}}$ as a “1” followed by a string of centimetre-wide zeroes, we’ll have a string of digits whose length is about 37 thousand million light years, or several times the extent of the observable universe.

Ω_{tot} for an Ideal Gas of Rotating Non-Point Particles

Now, we move on to examine an ideal gas of non-point particles, such as molecules. A point particle cannot rotate: being a point, it has no “handles” that can be hooked by another particle or by a field, to apply a torque to it. But now suppose that each of the gas particles has some structure, so that they *can* rotate, as shown in Figure 2.8. All three-dimensional structures of masses carry three *principal axes of rotation* about which they are able to rotate smoothly, like a well-balanced wheel on a car. This well-behaved rotation is easy to describe, because rotation about a principal axis requires no torque to keep the axis pointing in a fixed direction. Structures with symmetry have their principal axes in locations determined by that symmetry; so, for example, the oblong box shown in Figure 2.9 has its principal axes emanating from its centre of symmetry and emerging perpendicular to each face. Surprisingly, these axes are always mutually perpendicular, even when the structure has no symmetry at all. We will allow each particle to have its

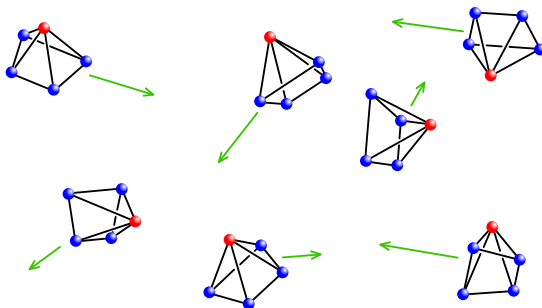


Fig. 2.8 An ideal gas of rotating non-point particles, with their velocity vectors

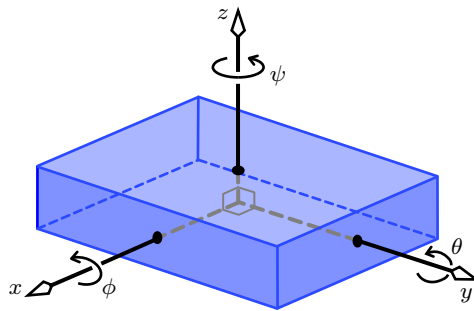


Fig. 2.9 The principal axes of an oblong box are its axes of symmetry. These axes define moments of inertia I_x, I_y, I_z . The box can rotate through angles ϕ, θ, ψ around the axes

own value of mass and its own moments of inertia, and will again calculate Ω_{tot} from (2.25).

The position of a non-point particle can be specified by the 3 spatial displacements x, y, z of its centre of mass relative to some origin. Its spatial orientation can be specified by 3 angular displacements ϕ, θ, ψ : these are rotation angles around, say, the three cartesian axes x, y , and z , respectively. (A body's orientation can always be specified by these three angles.) Each of these 6 coordinates can be paired with momenta that store the particle's energy: 3 linear momenta p_x, p_y, p_z for the spatial displacements, and 3 angular momenta L_1, L_2, L_3 for the angular displacements about the three principal axes (the subscripts “1, 2, 3” denote these axes). Each particle thus has 6 pairs of position–momentum coordinates, and so $D = 6$. The dx^{6N} in (2.25) represents these linear and angular displacements of the N particles, and the dp^{6N} represents the linear and angular momenta. As before, the particles do not interact with each other, and so the position and momentum in the integral (2.25) decouple. The space integral is

$$\begin{aligned} \int dx^{6N} &= \underbrace{dx_1 dy_1 dz_1 d\phi_1 d\theta_1 d\psi_1}_{\text{particle 1}} \dots \underbrace{dx_N dy_N dz_N d\phi_N d\theta_N d\psi_N}_{\text{particle } N} \\ &= V^N (2\pi)^{3N}. \end{aligned} \quad (2.53)$$

The momentum integral in (2.25), $\int dp^{6N}$, is the volume of an ellipsoid in $6N$ dimensions, similar to (2.42), but now including angular momenta. Particle i has mass m_i and moments of inertia pertaining to the principal axes “1, 2, 3” of I_{1i}, I_{2i}, I_{3i} , respectively.⁷ The angular momentum about a princi-

⁷ In all generality, the moment of inertia I is actually a tensor whose elements can be written as a 3×3 matrix. It is often mistakenly thought to be a number that is defined relative to a given axis. In fact, it is defined relative to a *point*, not an axis; its value does not depend on any choice of axis. But it turns out—and this is not meant to

pal axis can be written in the simple form “ $L^2/(2I)$ ”. The total energy of the N particles is then

$$\sum_{i=1}^N \frac{p_{xi}^2}{2m_i} + \frac{p_{yi}^2}{2m_i} + \frac{p_{zi}^2}{2m_i} + \frac{L_{1i}^2}{2I_{1i}} + \frac{L_{2i}^2}{2I_{2i}} + \frac{L_{3i}^2}{2I_{3i}} = E. \quad (2.54)$$

We require the volume of this $6N$ -dimensional ellipsoid.

The volume of a d -dimensional ellipsoid $x_1^2/a_1^2 + \cdots + x_d^2/a_d^2 = 1$ is the generalisation of (2.43) to the case of d arbitrary “radii” (that is, semi-axes lengths). This volume happens to be

$$\text{volume} = \frac{\pi^{d/2} a_1 a_2 \cdots a_d}{(d/2)!}. \quad (2.55)$$

To deal simply with this ellipsoid in $d = 6N$ dimensions, rather than include all the factors of mass and moment of inertia, we simply note that each of the semi-axes lengths $a_1, \dots, a_{6N}$ contributes a factor of $\sqrt{E}$ to the volume (because $a_1^2 = 2m_1 E$, and so on). Hence,

$$\int dp^{6N} = \text{volume of ellipsoid in momentum space} \propto (\sqrt{E})^{6N}. \quad (2.56)$$

Finally, (2.25) yields

$$\Omega_{\text{tot}}(E) = \frac{1}{h^{6N}} \int_{\text{energies 0 to } E} dx^{6N} dp^{6N} \propto V^N E^{6N/2}. \quad (2.57)$$

The “6” in the exponent of E results from each particle contributing 6 terms to the total energy in (2.54).

Ω_{tot} for an Ideal Gas of Rotating, Oscillating Diatomic Molecules

Suppose that the rotating particles of the last few paragraphs can also oscillate. We will examine the simplest such example here: a diatomic molecule, and we’ll allow each molecule to have its own mass. Similar to (2.53), molecule i ’s centre of mass has position coordinates x_i, y_i, z_i . We’ll see in Section 5.6.1 that diatomic molecules don’t rotate about the line joining the two atoms, because the atoms don’t have any “handles” that can be grabbed

be obvious—that when we are dealing with principal axes, we can treat the moment of inertia as a number for each axis. This is what I have done above. The widespread belief that I refers to a specific axis probably arose because the eigenvectors of I define the preferred axes of spin for commonly used bodies in engineering, such as wheels. But even for these, a reference point must be specified. For example, a wheel spinning at the end of an axle presents different dynamics to a wheel spinning in the middle of that same axle. See Section 5.6.1 for more discussion.

by collisions with other atoms to spin the molecule around that line. It follows that we can describe the orientation of the molecule using just two angles instead of three: say, ϕ_i and θ_i . We must also include a sixth spatial coordinate, r_i , which is the distance at each moment that the two atoms have stretched from their equilibrium separation, as the molecule oscillates. Each of these 6 spatial coordinates is paired with a momentum coordinate, as we'll see in a moment. So, $D = 6$ here.

Mimicking (2.53), the integration over the N position triplets and the $2N$ angular coordinates contributes a factor of $V^N (2\pi)^{2N}$ to Ω_{tot} . What about the “stretch” coordinates, $r_1, r_2, \dots, r_N$? The values of these are affected by collisions with other molecules, and these collisions change those molecules' momenta. It follows that we must include the integration over the stretch coordinates in the momentum integration. Equation (2.25) produces

$$\Omega_{\text{tot}}(E) = \frac{V^N (2\pi)^{2N}}{h^{6N}} \int_{\text{energies 0 to } E} dr^N dp^{6N}. \quad (2.58)$$

The $D = 6$ momentum coordinates of molecule i are as follows. For the entire molecule, the usual p_{xi}, p_{yi}, p_{zi} appear. The rotation around just two axes requires L_{1i}, L_{2i} . The oscillation is specified by $\mu_i \dot{r}_i$, where μ_i is molecule i 's *reduced mass*, examined in detail in Section 5.6.1. Similar to (2.54), these coordinates are tied to the total energy in the following way:

$$\sum_{i=1}^N \frac{p_{xi}^2}{2m_i} + \frac{p_{yi}^2}{2m_i} + \frac{p_{zi}^2}{2m_i} + \frac{L_{1i}^2}{2I_{1i}} + \frac{L_{2i}^2}{2I_{2i}} + \frac{k_i r_i^2}{2} + \frac{\mu_i \dot{r}_i^2}{2} = E, \quad (2.59)$$

where the two atoms in molecule i are modelled as being joined by a spring with spring constant k_i . Note that, as required, the stretch coordinates r_i are tied to momenta here. Equation (2.59) defines an ellipsoid in $7N$ dimensions, and the integral in (2.58) is the volume of this ellipsoid. Each of the $7N$ semi-axes lengths contributes a factor of $\sqrt{E}$ to this volume, and hence this volume is proportional to $(\sqrt{E})^{7N}$. Equation (2.58) yields

$$\Omega_{\text{tot}}(E) \propto V^N E^{7N/2}. \quad (2.60)$$

The “7” in the exponent of E results from each molecule contributing 7 terms to the total energy in (2.59).

Ω_{tot} for a Lattice of Point Oscillators in One Dimension

Figure 2.10 shows a lattice of N oscillators constrained to move in one dimension. The position of the i^{th} point oscillator is determined by its displacement x_i from its equilibrium position, and its momentum is p_i . This single

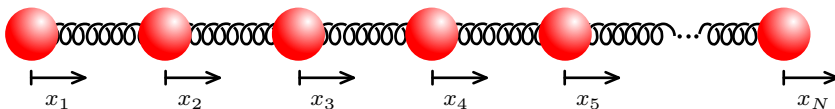


Fig. 2.10 A one-dimensional lattice of N point oscillators. Particle i has a displacement x_i from its equilibrium position in the lattice

position–momentum pair sets $D = 1$. Equation (2.25) then requires

$$dx^N \equiv dx_1 \dots dx_N, \quad dp^N \equiv dp_1 \dots dp_N, \quad (2.61)$$

and

$$\Omega_{\text{tot}}(E) = \frac{1}{h^N} \int dx_1 \dots dx_N dp_1 \dots dp_N. \quad (2.62)$$

But the displacement and momentum of an oscillator are *not* independent, and so (2.62) does not separate into position and momentum integrals. This means we don’t integrate over position separately, meaning that no volume term (or rather lattice-length term) is produced. What we can do is note that the lattice’s total energy E is given by

$$\sum_{i=1}^N \frac{k_i x_i^2}{2} + \frac{p_i^2}{2m_i} = E, \quad (2.63)$$

where particle i has spring constant k_i and mass m_i . This equation has $2N$ terms on its left-hand side, and defines an ellipsoid in a phase space of $2N$ dimensions. The volume of this ellipsoid is the sought-after integral (2.25). Each of the $2N$ semi-axes lengths contributes a factor of $\sqrt{E}$ to this ellipsoid volume, and so this volume is proportional to $(\sqrt{E})^{2N}$. Then, (2.25) becomes (without cancelling the 2’s—we’ll see why in the next sentence)

$$\Omega_{\text{tot}}(E) \propto E^{2N/2}. \quad (2.64)$$

The first “2” in the exponent of (2.64) denotes each particle having contributed 2 terms to the total energy in (2.63).

Ω_{tot} for a Lattice of Point Oscillators in Three Dimensions

This is the expected extension of the one-dimensional case. Each particle moves in three dimensions, and thus has three associated position–momentum pairs of variables, giving $D = 3$. Equation (2.25) uses

$$dx^{3N} \equiv \underbrace{dx_1 dy_1 dz_1}_{\text{particle 1}} \dots \underbrace{dx_N dy_N dz_N}_{\text{particle N}},$$

$$dp^{3N} \equiv \underbrace{dp_{x1} dp_{y1} dp_{z1}}_{\text{particle 1}} \cdots \underbrace{dp_{xN} dp_{yN} dp_{zN}}_{\text{particle } N}, \quad (2.65)$$

and

$$\Omega_{\text{tot}}(E) = \frac{1}{h^{3N}} \int_{\text{energies 0 to } E} dx^{3N} dp^{3N}. \quad (2.66)$$

Again, each particle's position and momentum are not independent; thus no volume term arises. The lattice's total energy describes an ellipsoid in a phase space of $6N$ dimensions:

$$\sum_{i=1}^N \frac{k_{xi}x_i^2}{2} + \frac{k_{yi}y_i^2}{2} + \frac{k_{zi}z_i^2}{2} + \frac{p_{xi}^2}{2m_i} + \frac{p_{yi}^2}{2m_i} + \frac{p_{zi}^2}{2m_i} = E, \quad (2.67)$$

where k_{xi} is the spring constant in the x direction for particle i (and similarly for k_{yi} and k_{zi}). As before, each of the $6N$ semi-axes lengths contributes a factor of $\sqrt{E}$ to the ellipsoid's volume. This volume is then proportional to $(\sqrt{E})^{6N}$, and so

$$\Omega_{\text{tot}}(E) \propto E^{6N/2}. \quad (2.68)$$

The “6” in the exponent of (2.68) arises from each particle having contributed 6 terms to the total energy in (2.67).

Ω_{tot} for Complex Molecules

Complex molecules have a great number of modes of motion, and so their value of D is large. These modes are not as readily classified and counted as they were for the simple particles we described above. Experiments on complex molecules can be difficult to perform: the molecules might be rendered simpler to examine if they are formed into a gas, but complex molecules can seldom be coaxed into gaseous form without breaking up. Various modes of oscillation can be excited by illuminating the molecules with, say, laser light; but this procedure is, by its nature, selective of the modes to be excited. This selectivity prevents any straightforward analysis of the molecules' motion, and so we won't attempt to calculate Ω_{tot} for them.

Summary of the Above Results

The calculations of Ω_{tot} in the last few pages have probably been a little bewildering, and so we will summarise the results of Section 2.4 here.

1. **Free Point Particle in One Dimension:** The particle has a single space coordinate, x , and a single momentum coordinate, p . These constitute the $D = 1$ position–momentum coordinate pair for the par-

ticle. Its energy is $E = p^2/(2m)$. This single term that is quadratic in momentum contributed a single factor of $\sqrt{E}$ to $\Omega_{\text{tot}}(E)$ in (2.27). We introduce an important new quantity here:

$$\nu \equiv \left[\begin{array}{l} \text{the number of quadratic terms, per particle,} \\ \text{that appear in a system's energy.} \end{array} \right] \quad (2.69)$$

We'll refer to ν as “the number of quadratic energy terms per particle”. So, a free point particle in one dimension has $\Omega_{\text{tot}}(E) \propto E^{\nu/2}$, with $\nu = 1$.

2. Free Point Particle in Three Dimensions: The particle has

$$\begin{aligned} &\text{space coordinates: } x, y, z, \\ &\text{momentum coordinates: } p_x, p_y, p_z. \end{aligned} \quad (2.70)$$

There are thus $D = 3$ position–momentum coordinate pairs for the particle. Its energy is

$$E = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m}. \quad (2.71)$$

These $\nu = 3$ quadratic energy terms each contributed a factor of $\sqrt{E}$ to $\Omega_{\text{tot}}(E)$ in (2.33). So, $\Omega_{\text{tot}}(E) \propto E^{\nu/2}$.

3. Ideal Gas of N Point Particles: Particle i has

$$\begin{aligned} &\text{space coordinates: } x_i, y_i, z_i, \\ &\text{momentum coordinates: } p_{xi}, p_{yi}, p_{zi}. \end{aligned} \quad (2.72)$$

There are thus $D = 3$ position–momentum coordinate pairs per particle. The gas's energy is

$$E = \sum_{i=1}^N \frac{p_{xi}^2}{2m} + \frac{p_{yi}^2}{2m} + \frac{p_{zi}^2}{2m}. \quad (2.73)$$

These $\nu = 3$ quadratic energy terms per particle each contributed a factor of $\sqrt{E}$ to $\Omega_{\text{tot}}(E)$ in (2.45). So, $\Omega_{\text{tot}}(E) \propto E^{\nu N/2}$.

4. Ideal Gas of N Rotating Non-Point Particles: Particle i has

$$\begin{aligned} &\text{space coordinates: } x_i, y_i, z_i, \phi_i, \theta_i, \psi_i, \\ &\text{momentum coordinates: } p_{xi}, p_{yi}, p_{zi}, L_{1i}, L_{2i}, L_{3i}. \end{aligned} \quad (2.74)$$

There are thus $D = 6$ position–momentum coordinate pairs per particle. The gas’s energy is (2.54):

$$E = \sum_{i=1}^N \frac{p_{xi}^2}{2m_i} + \frac{p_{yi}^2}{2m_i} + \frac{p_{zi}^2}{2m_i} + \frac{L_{1i}^2}{2I_{1i}} + \frac{L_{2i}^2}{2I_{2i}} + \frac{L_{3i}^2}{2I_{3i}}. \quad (2.75)$$

These $\nu = 6$ quadratic energy terms per particle each contributed a factor of $\sqrt{E}$ to $\Omega_{\text{tot}}(E)$ in (2.57). So, $\Omega_{\text{tot}}(E) \propto E^{\nu N/2}$.

5. Ideal Gas of N Rotating, Oscillating Diatomic Molecules: Molecule i has

$$\begin{aligned} \text{space coordinates: } & x_i, y_i, z_i, \phi_i, \theta_i, r_i, \\ \text{momentum coordinates: } & p_{xi}, p_{yi}, p_{zi}, L_{1i}, L_{2i}, \mu_i \dot{r}_i. \end{aligned} \quad (2.76)$$

There are thus $D = 6$ position–momentum coordinate pairs per particle—note that a “particle” here is the whole molecule, and does not refer to the individual atoms. The gas’s energy is (2.59):

$$E = \sum_{i=1}^N \frac{p_{xi}^2}{2m_i} + \frac{p_{yi}^2}{2m_i} + \frac{p_{zi}^2}{2m_i} + \frac{L_{1i}^2}{2I_{1i}} + \frac{L_{2i}^2}{2I_{2i}} + \frac{k_i r_i^2}{2} + \frac{\mu_i \dot{r}_i^2}{2}. \quad (2.77)$$

These $\nu = 7$ quadratic energy terms per particle each contributed a factor of $\sqrt{E}$ to $\Omega_{\text{tot}}(E)$ in (2.60). So, $\Omega_{\text{tot}}(E) \propto E^{\nu N/2}$.

6. Lattice of N Point Oscillators in One Dimension: Particle i has

$$\begin{aligned} \text{space coordinate: } & x_i, \\ \text{momentum coordinate: } & p_i. \end{aligned} \quad (2.78)$$

There is thus $D = 1$ position–momentum coordinate pair per particle. The lattice’s energy is (2.63):

$$E = \sum_{i=1}^N \frac{k_i x_i^2}{2} + \frac{p_i^2}{2m_i}. \quad (2.79)$$

These $\nu = 2$ quadratic energy terms per particle each contributed a factor of $\sqrt{E}$ to $\Omega_{\text{tot}}(E)$ in (2.64). So, $\Omega_{\text{tot}}(E) \propto E^{\nu N/2}$.

7. Lattice of N Point Oscillators in Three Dimensions: Particle i has

$$\begin{aligned} \text{space coordinates: } & x_i, y_i, z_i, \\ \text{momentum coordinates: } & p_{xi}, p_{yi}, p_{zi}. \end{aligned} \quad (2.80)$$

There are thus $D = 3$ position–momentum coordinate pairs per particle. The lattice’s energy is (2.67):

$$E = \sum_{i=1}^N \frac{k_{xi}x_i^2}{2} + \frac{k_{yi}y_i^2}{2} + \frac{k_{zi}z_i^2}{2} + \frac{p_{xi}^2}{2m_i} + \frac{p_{yi}^2}{2m_i} + \frac{p_{zi}^2}{2m_i}. \quad (2.81)$$

These $\nu = 6$ quadratic energy terms per particle each contributed a factor of $\sqrt{E}$ to $\Omega_{\text{tot}}(E)$ in (2.68). So, $\Omega_{\text{tot}}(E) \propto E^{\nu N/2}$.

The above values of $\Omega_{\text{tot}}(E)$ for systems of increasing complexity are summarised in Table 2.1.

Table 2.1 Values of the various parameters for the calculations of $\Omega_{\text{tot}}(E)$ in this chapter. D is the number of phase-space dimension pairs allocated to each particle. ν is the number of quadratic energy terms per particle

System	D	cell “volume”	$\Omega_{\text{tot}} \propto$	ν
free particle in one dimension, (2.27)	1	h	$E^{1/2}$	1
free particle in three dimensions, (2.33)	3	h^3	$E^{3/2}$	3
ideal gas of N point particles, (2.45)	3	h^{3N}	$E^{3N/2}$	3
ideal gas of N rotating non-point particles, (2.57)	6	h^{6N}	$E^{6N/2}$	6
ideal gas of N rotating, oscillating diatoms, (2.60)	6	h^{6N}	$E^{7N/2}$	7
lattice of N point oscillators in 1D, (2.64)	1	h^N	$E^{2N/2}$	2
lattice of N point oscillators in 3D, (2.68)	3	h^{3N}	$E^{6N/2}$	6
arbitrary system of N particles	D	h^{DN}	$E^{\nu N/2}$	ν

2.4.1 A System’s Quadratic Energy Terms

The parameter ν is the number of the system’s dynamical coordinates that appear *quadratically* in the energy of each of the system’s particles. It turns up widely in statistical mechanics, and is usually called the number of *degrees of freedom* per particle in the system. As explained below, I avoid this term, and use *quadratic energy terms* per particle instead. The total number of quadratic energy terms for the system, νN , appears in the energy dependence, $E^{\nu N/2}$, of $\Omega_{\text{tot}}(E)$. The reason for ν appearing in this way is that the

system's quadratic dependence on such a coordinate allows that coordinate to contribute a dimension to the higher-dimensional ellipsoid that describes the system's energy in phase space. Whereas the parameter D sets the volume of a cell in phase space to be h^{DN} , the really important indicator of a system's ability to store energy is its number of quadratic energy terms, νN .

Quadratic Energy Terms, Not Degrees of Freedom

In *classical* mechanics, a degree of freedom is defined to be any independent coordinate used in describing the positions of a system's constituents. Thus, a free particle in one dimension has one degree of freedom: its position. Similarly, the position of a particle that oscillates in one dimension is also its sole degree of freedom.

In contrast, statistical mechanics defines a degree of freedom to be any coordinate that contributes *quadratically* to a system's energy. This coordinate might have nothing to do with position. Hence, a free particle moving in one dimension has energy $1/2 mv^2$, and is then said to have $\nu = 1$ degree of freedom, which is its velocity (or momentum). Likewise, a particle that oscillates in one dimension with a spring constant k has energy $1/2 mv^2 + 1/2 kx^2$, and is thus said to have $\nu = 2$ degrees of freedom: its position and its velocity.

Not surprisingly, these different uses of “degree of freedom” in classical and statistical mechanics give rise to some perplexity. I have chosen to replace “degree of freedom” with “quadratic energy term”.

As is evident in Table 2.1, in simple systems ν may or may not equal D , the number of pairs of position–momentum coordinates in the system's phase space.

To demonstrate, consider a lattice of point oscillators in one dimension, such as in Figure 2.10. The i^{th} displacement term x_i of the position–momentum pairs in (2.62) gives the potential energy $k_i x_i^2/2$ of the oscillation of particle i in (2.63). The i^{th} momentum term p_i gives the kinetic energy $p_i^2/(2m_i)$ of the *same* oscillation (there is only one oscillation per particle). Here, $D = 1$ (one position–momentum pair for the particle), but $\nu = 2$ (two terms contributing quadratically to the energy of the particle: kinetic and potential).

Note that when our particle is, say, a diatomic molecule, the number of “quadratic energy terms per particle” corresponding to the oscillation of the two atoms about their centre of mass is still $\nu = 2$, because we are concerned only with the molecule as a whole: it is still treated as a “single particle”. These two quadratic energy terms are the $k_i r_i^2/2$ and $\mu_i \dot{r}_i^2/2$ terms in (2.59). If we insist on treating the molecule as *two* particles, then in the expression for the system's energy (2.59), we will write *two* terms for potential energy (one term for each particle) and *two* terms for kinetic energy, giving *four*

terms for these two particles. But there are still two terms per particle here, and so again we say that each particle (now an atom) has $\nu = 2$ quadratic energy terms.

The important point is that the total number of quadratic energy terms νN gives the energy dependence of Ω_{tot} as $E^{\nu N/2}$. We summarise:

Ω_{tot} for a Gas and a Lattice

For a gas of N non-interacting particles occupying a volume V and with total energy E , where each particle has ν quadratic energy terms,

$$\Omega_{\text{tot}} \propto V^N E^{\nu N/2}. \quad (2.82)$$

For a lattice of oscillators with total energy E , where each particle has ν quadratic energy terms,

$$\Omega_{\text{tot}} \propto E^{\nu N/2}. \quad (2.83)$$

2.4.2 When Particles are Identical Classical

Up until now, we have assumed the particles of a gas to be distinguishable: able to be numbered. Recall from Figure 2.4 that each *permutation* of such numbered particles in position–momentum space is associated with its own unique cell in phase space. $\Omega_{\text{tot}}(E)$ is the number of cells in phase space with energy between 0 and E , and that means this number has so far treated the particles as distinguishable.

It follows that we must revise our counting scheme when considering the identical-classical particles described in Section 1.1.1. Recall that these particles are really *identical*: they cannot be numbered. Now refer to the calculation that produced (2.23), where we saw that for every particle in a standard room, several million unoccupied cells of position–momentum space exist.⁸ In this case, the chance is overwhelming that, at most, only *one* particle will ever be found in any particular cell of that position–momentum space.

With this sparseness of occupied position–momentum cells in mind, refer to Figure 2.11. This shows $N = 2$ distinguishable particles occupying two random cells of position–momentum space. Two occupations of the space are evident here: the top- and bottom-left pictures in the figure. When the

⁸ Remember the distinction between position–momentum space and phase space: at any moment, N particles will occupy N points of position–momentum space, and hence occupy, at most, N cells of that space when it is partitioned into cells; whereas, by construction, they are always represented by a *single* point of phase space, and hence occupy a single cell of that space when it is partitioned into cells.

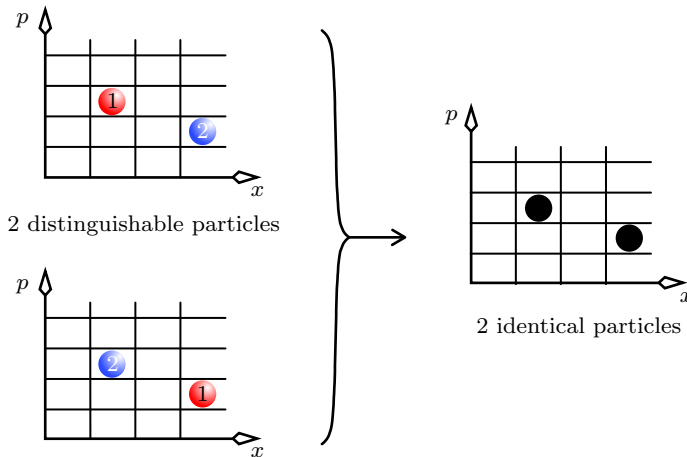


Fig. 2.11 **Left:** Suppose that our system contains just two particles, “1” and “2”, and they are distinguishable. When the number of position–momentum cells (about 20 here) is vastly greater than the number of particles, those particles will almost certainly occupy two different cells. When they *do* occupy two different cells, they can do so in $2!$ ways, as shown. **Right:** When the particles are identical, the two distinct occupations on the left must be counted as one: we must divide the number of left-hand configurations by $2!$

particles are identical classical, these two permutations must be counted as one. That is, we must count *combinations* instead of permutations. So, when the particles are identical classical, the above calculations of Ω_{tot} will have over-counted by a factor of $N!$, just as we saw in Section 1.1.1 for the ink in the bathtub. For such particles, we must divide the expressions for Ω_{tot} by $N!$.

If the number of position–momentum cells were *not* very much larger than the number of gas particles, there would be a high chance that two or more particles would occupy the same cell. Figure 2.12 shows the case of $N = 2$ particles with only two position–momentum cells available. The number of such microstates would not need modifying if the particles were considered to be identical, and so we would *not* divide that number of states by $2!$. In general, to convert Ω_{tot} for N distinguishable particles into Ω_{tot} for N identical-classical particles, the number of distinguishable microstates for which all occupied cells had one particle would have to be divided by $N!$, but the smaller number of microstates for which some occupied cells had *more than* one particle would have to be treated differently, making the overall calculation of Ω_{tot} more complicated. We will treat such “crowding” in Chapter 7 for the case of identical *quantum* particles, which really do exist. For now, we will assume that the number of position–momentum cells *is* much larger than the number of particles—just as we saw in (2.23)—so that we *do* only divide the number of microstates by $N!$ when dealing with a gas. On the

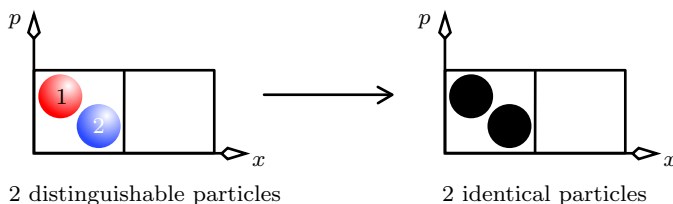


Fig. 2.12 Left: Once again we have $N = 2$ distinguishable particles, but now very few position–momentum cells are available. Particles 1 and 2 might well occupy the same cell, as shown. **Right:** If the particles are considered as identical, then nothing changes from the left-hand picture: we do not divide the number of configurations by $2!$

other hand, the particles of a solid are distinguished by their locations at the various lattice sites; so for them, there is no dividing by $N!$.

For an ideal gas of a large number of point particles, dividing $\Omega_{\text{tot}}(E)$ in (2.50) by $N! \approx N^N e^{-N}$ gives the following expression, with superscript “ic” for “identical classical”:

$$\Omega_{\text{tot}}^{\text{ic}} \approx \frac{V^N}{N^N e^{-N}} \left(\frac{2\pi emkT}{h^2} \right)^{3N/2} = \left(\frac{V}{N} \right)^N e^{5N/2} \left(\frac{2\pi mkT}{h^2} \right)^{3N/2}. \quad (2.84)$$

We’ll call on this expression in Section 3.8.1 when introducing the Sackur–Tetrode equation, as well as in later chapters.

Does treating a room full of gas particles as identical classical change the number of microstates significantly? Return to the room full of air that produced (2.52). Dividing the number of microstates for that distinguishable case by $10^{27}!$ gives the number of microstates for identical-classical particles as

$$\Omega_{\text{tot}}^{\text{ic}} \approx \frac{10^{3.5 \times 10^{28}}}{10^{27}!} \approx 10^{8.4 \times 10^{27}} \text{ microstates}. \quad (2.85)$$

Despite the division by a staggeringly large number here, the result remains staggeringly large.

2.5 The Density of States

Treating the spread of energies of a system’s microstates as forming a continuum is analogous to treating the mass of a ruler as though it were distributed continuously along the ruler’s length. The ruler’s mass is not really a continuum; this mass is located in the nuclei of the atoms that comprise the ruler. We treat the mass as a continuum merely to make the physics tractable, because real numbers tend to be easier to manipulate than whole numbers. You will recall from Section 1.6.2’s discussion of probability density that, once we

Table 2.2 Finding the mass in a ruler as an analogy to counting energy states

Finding the amount of mass	Counting the number of states
$M(x) \equiv$ total mass in 0 to x	$\Omega_{\text{tot}}(E) \equiv$ number of states in 0 to E
$\lambda(x) \equiv$ mass density at x	$g(E) \equiv$ density of states at E
$\Delta M \simeq \lambda(x) \Delta x$ = mass in Δx around x	$\Omega(E) = \Delta \Omega_{\text{tot}} \simeq g(E) \Delta E$ = number of states at E , where ΔE is energy-level spacing.
$dM = \lambda(x) dx$ = infinitesimal mass at x	$d\Omega_{\text{tot}} = g(E) dE$ = “infinitesimal number of states” at E

accept this approximation of a continuous spread of mass, we can work with the ruler’s linear mass density $\lambda(x)$ from (1.147), its mass per unit length at x . Given $\lambda(x)$, the amount of mass located in some small ruler segment Δx near x is $\Delta M \simeq \lambda(x) \Delta x$. We even refer to an infinitesimal mass $dM = \lambda(x) dx$ “at x ”, even though, strictly speaking, this has no physical meaning for a ruler made of atoms. But it’s a useful concept for our continuum approximation of the ruler.

The individual states of a system are analogous to the individual nuclei of a ruler. Just as discussing the individual nuclei at the “ x notch” is problematic, so too the number of states $\Omega(E)$ at some energy E of a system can be problematic to deal with when we are treating energy as a continuum. Instead, we “coarse-grain” $\Omega(E)$ like the piece of mass $\Delta M(x)$ of the ruler, as shown in Table 2.2.

In the same way that we model the ruler as a continuum, and thus work with the mass *density* at a point, we model the energy spread E of a system’s states as continuous. Analogous to the total mass of $M(x)$ of a ruler in 0 to x is the total number of states $\Omega_{\text{tot}}(E)$ in 0 to E . Analogous to the ruler’s linear mass density $\lambda(x) = M'(x)$ is the *density of states* or *spectrum of accessible states* $g(E)$:

$$g(E) = \Omega'_{\text{tot}}(E). \quad (2.86)$$

We calculate the density of states $g(E)$ by first finding $\Omega_{\text{tot}}(E)$ for various systems in the manner of Section 2.4, and then differentiating that with respect to E . This density of states will play a key role in the *quantum statistics* of Chapter 7.

When a system’s energy is truly continuous, its number of states “of” energy E is $d\Omega_{\text{tot}} = g(E) dE$. Otherwise, if some natural non-infinitesimal choice of energy spacing ΔE is available, we write the number of states at E as

$$\Omega(E) = \Delta \Omega_{\text{tot}} \simeq g(E) \Delta E. \quad (2.87)$$

A suitable choice of ΔE might come from the sort of analysis that produced (2.12). Let's calculate $g(E)$ for a monatomic gas.

Density of States of a Monatomic Gas

For N large, calculate $g(E)$ by differentiating $\Omega_{\text{tot}}(E)$ from (2.49), using the fact that $3N/2 = E/(kT)$ (proved in the next chapter), to arrive at the second line below:

$$\begin{aligned} g(E) &= \Omega'_{\text{tot}}(E) \approx \frac{V^N}{\sqrt{2\pi}} \left(\frac{4\pi em}{3Nh^2} \right)^{3N/2} \frac{3N}{2} E^{3N/2-1} \\ &= \frac{V^N}{\sqrt{2\pi}} \left(\frac{4\pi em}{3Nh^2} \right)^{3N/2} \frac{E}{kT} E^{3N/2-1} = \frac{\Omega_{\text{tot}}(E)}{kT} \quad (N \text{ large}). \end{aligned} \quad (2.88)$$

Combine this with (2.87), to write

$$\Omega(E) \simeq g(E) \Delta E \stackrel{(2.88)}{=} \frac{\Omega_{\text{tot}}(E) \Delta E}{kT} \quad (N \text{ large}). \quad (2.89)$$

Equation (2.52) gave a typical value of $\Omega_{\text{tot}}(E)$. What about $\Delta E/(kT)$? Recall (2.14): for the example in Section 2.2 where we increased an energy quantum number n by 1 for a room of gas at an everyday temperature, the resulting ΔE satisfied

$$\frac{\Delta E}{kT} = \frac{1}{n} \simeq 5 \times 10^{10}. \quad (2.90)$$

For this example of a gas at room temperature, (2.89) becomes

$$\Omega(E) \simeq 10^{3.5 \times 10^{28}} \times 5 \times 10^{10} \simeq 10^{3.5 \times 10^{28}} = \Omega_{\text{tot}}(E). \quad (2.91)$$

We see why $\Omega_{\text{tot}}(E)$ and $\Omega(E)$ are usually treated interchangeably in statistical mechanics. But we might add that in most derivations of (2.88), the $3N/2 - 1$ in the first line is approximated as $3N/2$ (since $N \gg 1$), which has the effect of removing kT from the final expression for $g(E)$. The result is that $\Omega_{\text{tot}}(E)$ and $g(E)$ are also treated interchangeably in many textbooks, so that

$$\Omega(E) \approx \Omega_{\text{tot}}(E) \approx g(E). \quad (2.92)$$

It is physically meaningless to replace either $\Omega_{\text{tot}}(E)$ or $\Omega(E)$ with $g(E)$, because Ω and g have different dimensions: Ω is a number of states, whereas g is a number of states per unit energy. Equation (2.87) ties $\Omega(E)$ and $g(E)$ together via some representative energy width ΔE , which tends to act as a factor in the maths without changing the physics. Thus, this ΔE is almost never written explicitly. Still, we should be ever mindful that such a strange bit of mathematics is often implicitly present in many analyses.

2.6 Ω_{tot} for Massless Particles

The calculations of $\Omega_{\text{tot}}(E)$ in Section 2.4 used a non-zero rest mass for the particles. The analogous calculations for particles of *zero* rest mass are, in fact, very similar. We first take a moment to recall the definition of zero rest mass. All particles of matter have energy $E = \gamma m_0 c^2$, where

- $\gamma \equiv 1/\sqrt{1 - v^2/c^2}$ is the “gamma factor” determined by their speed v and the speed of light c .
- $m_0 > 0$ is their *rest mass*: their mass (resistance to being accelerated) when at rest, where the zero subscript indicates zero speed. Matter particles become increasingly more difficult to accelerate as their speed increases; so, by the very definition of mass, this means they become more massive at high speed. Their mass m at speed, or “relativistic mass”, is given by $m = \gamma m_0$.

Special relativity defines the momentum (vector) $\mathbf{p}$ of a velocity- $\mathbf{v}$ particle as $\mathbf{p} \equiv \gamma m_0 \mathbf{v}$, which has magnitude p . The particle’s energy $E = \gamma m_0 c^2$ can be written in squared form as

$$E^2 = p^2 c^2 + m_0^2 c^4. \quad (2.93)$$

“Particles” of light, *photons*, have energy $E = pc$, and so don’t appear to fit into the above scheme; indeed, they are not particles of matter. But they can be incorporated by realising that $E = pc$ is just an instance of (2.93) for $m_0 = 0$. This is why photons are said to have zero rest mass. Their mass at their one allowed speed, c , can be defined as

$$m = E/c^2 = p/c. \quad (2.94)$$

We will calculate a density of states for the photon in the following pages, using the arguments of Section 2.4. When studying blackbody radiation in Chapter 9, we’ll require the photon’s density of states as a function of its frequency f , in both one and three dimensions, but calculated for a *single* photon. As well as calculating that here, we’ll recalculate it in Chapter 9 with a slightly different approach that introduces the very useful concept of a “wave number”.

Beside photons, we deal here with *phonons*. The phonon quantises the sound waves that transport the energy of lattice vibrations in three dimensions in a solid, and we will introduce these properly when studying the Debye model of heat capacity in Chapter 7. Debye’s model uses the density of states of a *single* phonon, and we derive that here. Phonons don’t travel at the speed of light, but neither can they be brought to rest, and they are often described as particles of zero rest mass. Like photons, the energy of a phonon is also often written as $E = pc$; but here, c is the speed of the phonon, not of light.

We will need to incorporate photons' and phonons' modes of vibration, or *polarisations* (which become *spins* in quantum-mechanical language):

- **One** polarisation is exhibited by *phonons* travelling through a liquid: this is longitudinal, describing vibration along the sound wave's direction of travel. A liquid cannot support transverse waves.
- **Two** polarisations are exhibited by *photons*, corresponding to the electric (or magnetic) field vector of the associated electromagnetic wave being decomposable along two orthogonal directions that are both orthogonal to the wave's direction of travel.
- **Three** polarisations are exhibited by *phonons* travelling through a crystal: along with the longitudinal polarisation above, two transverse polarisations arise, because a crystal *can* support transverse waves.

Motion in One Dimension

In Chapter 9, we will examine photons moving in one dimension to calculate the electrical noise they cause in a one-dimensional resistor. It might at first seem counter-intuitive that we can speak of a photon moving in one dimension, when the associated electromagnetic field is necessarily three dimensional. But we mean only that all the photons, or light waves, move in the same direction.

For one particle ($N = 1$) moving in one dimension ($D = 1$) and with one polarisation, (2.25) becomes

$$\Omega_{\text{tot}}^{1 \text{ pol}}(E) = \frac{1}{h} \int dx \int dp, \quad (2.95)$$

where the momentum integral is taken over all energies 0 to E . The particle is confined to a box of length L , and so $\int dx = L$. Its range of energies is represented by a one-dimensional momentum space extending from $p = -E/c$ to E/c , similar to the case shown in Figure 2.5. Hence, $\int dp = 2E/c$, and (2.95) becomes

$$\Omega_{\text{tot}}^{1 \text{ pol}}(E) = \frac{2LE}{hc}. \quad (2.96)$$

When multiple polarisations are allowed, each contributes a term of the form (2.96) to the number of states, although the speed c might differ for each polarisation. Photons of each polarisation move at the same speed c ; so for these, we simply double (2.96) to obtain the total number of states:

$$\Omega_{\text{tot}}(E) = \frac{4LE}{hc} \quad (\text{photons, 1 dimension}). \quad (2.97)$$

Their density of states is thus

$$g(E) = \Omega'_{\text{tot}}(E) = \frac{4L}{hc} \quad (\text{photons, 1 dimension}). \quad (2.98)$$

It is more usual to work with photons' density of *frequency* states $g(f)$.

Re-using Function Names

Should this density of frequency states $g(f)$ be given a different name, such as $\tilde{g}(f)$, to prevent confusion with the *different* function $g(E)$? In *generic* use, no confusion should arise from using the same function name “ g ” for energy and frequency. Of course, using the same symbol means that the expression “ $g(4)$ ” is ambiguous: is the “4” energy or frequency? If we really do want to work with *specific* values of energy or frequency, we might write “ $g(E = 4)$ ” and “ $g(f = 4)$ ”. To be really clear, we could certainly write the functions as $g(E)$ and $\tilde{g}(f)$; but better yet would be $g_E(E)$ and $g_f(f)$, since this notation is extendable to any number of other variables, and it results in the easy-to-read expressions “ $g_E(4)$ ” and “ $g_f(4)$ ”. Some physicists refer to their own use of a single function symbol in a grandiose self-deprecating way as “being sloppy”. But there is nothing sloppy about the use of a single function symbol; it is economical in speech, compact in notation, and streamlines the writing of long calculations. That is what good notation is all about.

By definition of a density,

$$g(f) df \equiv g(E) dE. \quad (2.99)$$

Setting $E = hf$ for a photon, (2.99) becomes

$$g(f) = g(E) dE/df = 4L/c \quad (\text{photons, 1 dimension}). \quad (2.100)$$

In Chapter 9 we will use a somewhat different approach to counting states for photons, but will obtain this same value of $g(f)$ in (9.20).

Motion in Three Dimensions

We will treat both phonons and photons in three dimensions. For one particle moving in three dimensions ($N = 1, D = 3$), (2.25) becomes, for each polarisation,

$$\Omega_{\text{tot}}^1{}^{\text{pol}}(E) = \frac{1}{h^3} \int dx^3 \int dp^3. \quad (2.101)$$

Put the particle in a box of volume V , obtaining $\int dx^3 = V$. Consider that its momentum $\mathbf{p}$ is a vector with squared length $p^2 = p_x^2 + p_y^2 + p_z^2$. Similar to the right-hand picture in Figure 2.6, all energies 0 to E are represented by

vectors in momentum space whose lengths run from 0 to E/c ; hence, these vectors occupy a sphere of radius E/c in that space. It follows that $\int dp^3$ is the volume of this sphere:

$$\int dp^3 = \frac{4\pi}{3} \frac{E^3}{c^3}. \quad (2.102)$$

Equation (2.101) becomes

$$\Omega_{\text{tot}}^{\text{pol}}(E) = \frac{V}{h^3} \frac{4\pi}{3} \frac{E^3}{c^3}. \quad (2.103)$$

- For **phonons in a liquid**, their single polarisation allows for the direct use of (2.103):

$$\Omega_{\text{tot}}(E) = \frac{4\pi V E^3}{3h^3 c^3} \quad (\text{phonons in liquid, 3 dimensions}). \quad (2.104)$$

This gives a density of states

$$g(E) = \Omega'_{\text{tot}}(E) = \frac{4\pi V E^2}{h^3 c^3} \quad (\text{phonons in liquid, 3 dimensions}). \quad (2.105)$$

- For the **photons in an oven** described in Chapter 9, each polarisation has the number of accessible states in (2.103)—and both of these polarisations move with the same speed c ; thus, we simply double the number of accessible states in (2.103), to obtain

$$\Omega_{\text{tot}}(E) = \frac{8\pi V E^3}{3h^3 c^3} \quad (\text{photons in oven, 3 dimensions}). \quad (2.106)$$

The calculation of $g(f)$ runs as it did for the one-dimensional case (2.100):

$$g(f) = g(E) dE/df = 8\pi V f^2/c^3 \quad (\text{photons in oven, 3 dimensions}). \quad (2.107)$$

We will rederive this result in Chapter 9 using a somewhat different approach, ultimately arriving at (9.37).

- In contrast to photons, the **phonons in a solid** of Chapter 7 have generally different speeds for each polarisation. Call these speeds c_1, c_2, c_3 , and write the total number of states from (2.103) as

$$\Omega_{\text{tot}}(E) = \sum_{i=1}^3 \Omega_{\text{tot}}^{\text{pol } i}(E) = \frac{V}{h^3} \frac{4\pi}{3} E^3 \left(\frac{1}{c_1^3} + \frac{1}{c_2^3} + \frac{1}{c_3^3} \right). \quad (2.108)$$

We can abbreviate such an unwieldy function of three wave speeds by defining a mean speed via the discussion in the following grey box.

Generalised Averages

The everyday idea of an “average”—the *arithmetic mean*—is only one of an infinite number of ways in which an average can be defined. A more general type of average of some data set can be defined for any operation that is homogeneous in the data (meaning that this operation treats the data democratically, as we’ll see shortly), such that the same result is obtained by replacing each item of data with this generalised average.

An example will make this idea obvious. Suppose, without loss of generality, that we have three numbers x_1, x_2, x_3 , and wish to define a mean using the simplest operation: addition. Addition is homogeneous, meaning the order of the numbers being added doesn’t affect the final result: $x_1 + x_2 + x_3 = x_1 + x_3 + x_2$, and so on. The mean m defined from this operation is thus required to satisfy

$$m + m + m \equiv x_1 + x_2 + x_3. \quad (2.109)$$

It follows that $m = (x_1 + x_2 + x_3)/3$, which is the usual expression for the *arithmetic mean*. If we replace addition with multiplication, we have

$$m m m \equiv x_1 x_2 x_3, \quad (2.110)$$

which yields the *geometric mean*, $m = (x_1 x_2 x_3)^{1/3}$. Similarly, for adding reciprocals,

$$\frac{1}{m} + \frac{1}{m} + \frac{1}{m} \equiv \frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} \quad (2.111)$$

defines the *harmonic mean* m . We can invent any new mean m by the same procedure: even something as convoluted and seemingly useless as

$$\sin^m m + \sin^m m + \sin^m m \equiv \sin^{x_1} x_1 + \sin^{x_2} x_2 + \sin^{x_3} x_3. \quad (2.112)$$

The important question is always “Will it be useful?”

A glance at (2.108) suggests that a “cubic harmonic” mean c can be usefully defined here, such that

$$\frac{3}{c^3} \equiv \frac{1}{c_1^3} + \frac{1}{c_2^3} + \frac{1}{c_3^3}. \quad (2.113)$$

This converts (2.108) into

$$\Omega_{\text{tot}}(E) = \frac{4\pi V E^3}{h^3 c^3} \quad (\text{phonons in solid, 3 dimensions}). \quad (2.114)$$

The corresponding density of states is

$$g(E) = \Omega'_{\text{tot}}(E) = \frac{12\pi V E^2}{h^3 c^3} \quad (\text{phonons in solid, 3 dimensions}). \quad (2.115)$$

We will call on this last density of states in (7.38).

It has been said that the classical subject of statistical mechanics became easier to learn when quantum mechanics arrived on the physics scene. The reason is that quantum mechanics gave us a *discrete* picture of the world with a granularity given by Planck’s constant h ; this allowed phase space to be discretised, partitioned into cells—microstates—that could be counted, giving the subject the familiar feel of simply enumerating a system’s possible configurations. Statistical mechanics is built on a knowledge of the relative numbers of microstates that represent each of a system’s “interesting” (macro)states. We will often be required to analyse this number of microstates Ω_{tot} , and hence an understanding of it as covered in this chapter is crucial for much of what is to come.



Chapter 3

The Laws of Thermodynamics

In which we investigate the concepts of energy and work, define temperature and entropy, write down the laws of thermodynamics, show how entropy growth is reflected in the everyday behaviour of systems, and delve into the interior of a star.

3.1 The Concept of Energy for a Central Force

The discussion in the previous chapter focussed heavily on the amount of energy in a system. But how did that energy get there in the first place—and what *is* energy, anyway? Energy is a central idea in physics, but like other basic precepts of the world in which we live, we really have no idea what it is. Perhaps the question has no meaning or can never be answered, since an answer presumably must be expressed in terms of other quantities, forming a chain that eventually arrives at a quantity that cannot be defined in terms of anything else.¹ Perhaps energy *is* just such an end-point. We know only that when we define this quantity called “energy” in some restricted way such that it remains constant over time in simple systems, we also find that it remains constant over time in very complex systems.

At this point, we will back up to explore the idea of energy from the viewpoint of knowing absolutely nothing about it. We place ourselves in the shoes of one who knows Newton’s laws and something about gravity, who thus knows that planets orbit the sun under its gravitational pull. We might even be aware of the electrostatic (Coulomb) force between charges; but we have no concept of energy. The following discussion is tangential to the main subject of statistical mechanics, but it serves to place our ideas of energy into something of a historical perspective.

We will study a mass m that is moving under the influence of some other mass M . These two masses are either points or spherically symmetric, and each might hold some charge; they might even possess some as-yet-undiscovered quantity that induces an additional interaction. We will simplify the scenario by requiring that $M \gg m$, because then, M becomes—to any required level of approximation—the system’s centre of mass. Newton’s laws

¹ We are familiar with this idea for all languages—including mathematics. Our dictionaries must have finite size, which implies that some words have no definition or are defined in a circular way; in practice, we must derive their meaning from experience.

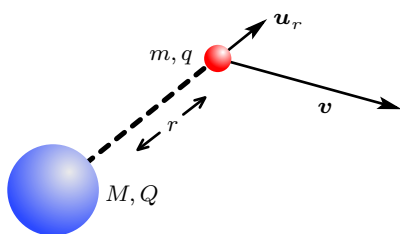


Fig. 3.1 The interaction of two point masses, or spherically symmetric masses, a distance r apart. These might hold charges Q and q , or indeed, some other as-yet-unknown force-causing quantity. A unit vector $\mathbf{u}_r$ is shown, pinned to m and pointing away from M . Mass m moves with some velocity $\mathbf{v}$ in the inertial frame in which M can be treated as being at rest when $M \gg m$

apply primarily to inertial frames, and the centre of mass is guaranteed to be stationary in an inertial frame. With M stationary in that inertial frame, we can ignore it in the following calculation, and that simplifies the mathematics to reveal the core concepts more easily.

These two interacting charged masses are shown in Figure 3.1. To determine how mass m moves, we require the total force acting on it. This might be, say, the sum of the gravitational and Coulomb forces exerted by M :

$$\text{total force on } m = \underbrace{\frac{-GMm}{r^2} \mathbf{u}_r}_{\text{gravity}} + \underbrace{\frac{kQq}{r^2} \mathbf{u}_r}_{\text{Coulomb}}, \quad (3.1)$$

where G is the gravitational constant and k is the Coulomb constant. Or indeed, the total force might be something else that depends on some other property of the masses that we know nothing about. But we will demand that, whatever it is, the force on m is *central*, meaning it acts along the line joining the two masses. We'll also suppose that its magnitude contains no angular dependence. So write this force as

$$\text{total force on } m = b(r) \mathbf{u}_r, \quad (3.2)$$

where $b(r)$ is some known function. Mass m has velocity $\mathbf{v}$ and acceleration $\mathbf{a}$. Its acceleration is given by Newton's "force = mass $\times$ acceleration":

$$b(r) \mathbf{u}_r = m\mathbf{a}. \quad (3.3)$$

Now consider "dotting" each side of (3.3) with $\mathbf{v}$:

$$b(r) \mathbf{u}_r \cdot \mathbf{v} = m\mathbf{a} \cdot \mathbf{v}. \quad (3.4)$$

The dot product $\mathbf{u}_r \cdot \mathbf{v}$ of a unit radial vector with velocity equals $\dot{r}$, the time rate of increase of the distance between the masses. (We will always use an

overdot to denote time differentiation: $\dot{r} \equiv dr/dt$.) We'll prove that identity in the following aside.

Some Short Manipulations with Vectors

The expression $\mathbf{u}_r \cdot \mathbf{v} = \dot{r}$ holds in one, two, and three dimensions. You can check it in the following way. Mass M is at $r = 0$, and thus defines the origin. Write the position of m as $\mathbf{s} = r\mathbf{u}_r$. The velocity $\mathbf{v}$ of the mass m is

$$\mathbf{v} \equiv \dot{\mathbf{s}} = \dot{r}\mathbf{u}_r + r\dot{\mathbf{u}}_r. \quad (3.5)$$

The required dot product is then

$$\begin{aligned} \mathbf{u}_r \cdot \mathbf{v} &= \mathbf{u}_r \cdot (\dot{r}\mathbf{u}_r + r\dot{\mathbf{u}}_r) \\ &= \dot{r} + r\mathbf{u}_r \cdot \dot{\mathbf{u}}_r. \end{aligned} \quad (3.6)$$

In one dimension, $\mathbf{u}_r$ points to the right everywhere to the right of $r = 0$, and it points to the left everywhere to the left of $r = 0$. Since we don't allow the mass m to pass through M (that is, we don't allow m to pass through the origin), we can say that $\mathbf{u}_r$ never changes as m moves. In that case, $\dot{\mathbf{u}}_r = \mathbf{0}$. Equation (3.6) becomes $\mathbf{u}_r \cdot \mathbf{v} = \dot{r}$, as we set out to prove.

In two and three dimensions, $\mathbf{u}_r$ always points radially away from M (which is the origin, $r = 0$), and so $\mathbf{u}_r$ certainly *does* change as m moves. Now, consider that $\mathbf{u}_r$ always has unit length. Then imagine taking snapshots of $\mathbf{u}_r$ at successive locations of m , and arranging all of these $\mathbf{u}_r$ vectors so that their tails sit on a common point. Then it's clear from these evolving snapshots that $\mathbf{u}_r$ is rotating: its head is moving on a circle whose centre is its tail. It follows that $\dot{\mathbf{u}}_r$ is always tangential to this circle. That means $\dot{\mathbf{u}}_r$ is always perpendicular to $\mathbf{u}_r$, and so $\mathbf{u}_r \cdot \dot{\mathbf{u}}_r = 0$. Again, (3.6) becomes $\mathbf{u}_r \cdot \mathbf{v} = \dot{r}$.

Equation (3.4) becomes

$$b(r) \dot{r} = m\mathbf{a} \cdot \mathbf{v}. \quad (3.7)$$

For notational brevity, introduce the function $B(r)$ whose r -derivative is $B'(r) \equiv b(r)$. It follows that $\dot{B} = b(r) \dot{r}$. Also notice that, with speed $v \equiv |\mathbf{v}|$,

$$\frac{d}{dt}(v^2) = \frac{d}{dt}(\mathbf{v} \cdot \mathbf{v}) = \mathbf{a} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{a} = 2\mathbf{a} \cdot \mathbf{v}. \quad (3.8)$$

(This last equation makes no reference to any force, and it shows that the speed of an object is constant if and only if its acceleration is orthogonal to its velocity.) Equation (3.7) now becomes

$$\dot{B} = \frac{m}{2} \frac{d}{dt} (v^2). \quad (3.9)$$

Rearranging the terms yields

$$\frac{d}{dt} \left[\frac{mv^2}{2} - B(r) \right] = 0. \quad (3.10)$$

Something remarkable has emerged here: a “*constant of the motion*”. When the mass m is given some initial velocity and is subject to a perhaps-complicated central force $b(r) \mathbf{u}_r$ in (3.2), then no matter what sort of complication motion it follows as a result, the quantity $1/2 mv^2 - B(r)$ stays constant with time. This constant of the motion is called the *energy* E of the system of the two masses.

The $1/2 mv^2$ in (3.10) concerns only the motion of m , and so is called the *kinetic energy* of m . The $-B(r)$ relates to the interaction of the masses, and is called the *potential energy* of the entire system (but often is just called the potential energy of mass m):

$$\text{kinetic energy of } m = 1/2 mv^2, \quad \text{potential energy} = -B(r). \quad (3.11)$$

Just why the word “potential” is appropriate becomes clear when we write (3.10) as

$$\frac{d}{dt} (\text{kinetic energy} + \text{potential energy}) = 0, \quad (3.12)$$

or,

$$\frac{d}{dt} (\text{kinetic energy}) = -\frac{d}{dt} (\text{potential energy}). \quad (3.13)$$

That is, the rate of *gain* of kinetic energy equals the rate of *loss* of potential energy. Or indeed, by switching the signs of both sides of (3.13), the rate of *loss* of kinetic energy equals the rate of *gain* of potential energy. Potential energy thus acts as a storage that can hold various amounts of the total energy. It releases that energy to become kinetic energy when “required” by the system.²

Note that $B(r)$ is defined by its derivative $b(r) = B'(r)$. It follows that $B(r)$ is determined only up to an added constant of integration. For the case

² *Potential* is defined as potential energy per unit of whatever it is that creates the potential energy: mass for gravity, and charge for electrostatic. Thus, we have

$$\begin{aligned} \text{gravitational potential} &\equiv \text{gravitational potential energy}/m = -GM/r, \\ \text{electrostatic potential} &\equiv \text{electrostatic potential energy}/q = kQ/r. \end{aligned} \quad (3.14)$$

Unfortunately, the minus sign in the gravitational potential is omitted by some practitioners in the astronomy, geodesy, and precise-timing communities. It then gets inserted into other equations, in an ad hoc attempt to get things to work. This only results in confusion, in both the maths and the physics of these fields.

of gravity and the Coulomb force applying in Figure 3.1,³

$$b(r) = \frac{-GMm}{r^2} + \frac{kQq}{r^2}. \quad (3.15)$$

Integrating this produces

$$B(r) = \frac{GMm}{r} - \frac{kQq}{r} + \text{constant}. \quad (3.16)$$

The constant is conventionally set equal to zero; this sets the potential energy $-B(r)$ to the intuitively reasonable value of zero when the masses' separation is infinite and they are thus not interacting at all. The two-mass system's energy becomes

$$E \equiv \frac{mv^2}{2} - B(r) = \frac{mv^2}{2} - \frac{GMm}{r} + \frac{kQq}{r}. \quad (3.17)$$

Next, we ask the question: how does the energy of the above system change if we interact with the mass m ? We could apply a force $\mathbf{F}$ to m for some long time, but as m continues to move, the kinetic and potential energy both change continuously. It's simpler, then, to apply $\mathbf{F}$ only for an *infinitesimal* time dt , and ask how the energy E changes in this time.

Consider first the simpler situation when M is absent, so that m moves with constant velocity $\mathbf{v}$. Its energy is purely kinetic: $E = 1/2 mv^2$. Now apply a force $\mathbf{F}$ to it for a time dt , during which m moves through $d\mathbf{x} = \mathbf{v} dt$. This force accelerates m by an amount $\mathbf{a} = \mathbf{F}/m$. After the time dt , the mass's total energy (which equals its kinetic energy) is⁴

$$\begin{aligned} E(t + dt) &= 1/2 mv^2(t + dt) = 1/2 m\mathbf{v}(t + dt) \cdot \mathbf{v}(t + dt) \\ &= 1/2 m(\mathbf{v} + \mathbf{a} dt) \cdot (\mathbf{v} + \mathbf{a} dt) = 1/2 m(\mathbf{v} \cdot \mathbf{v} + 2\mathbf{a} \cdot \mathbf{v} dt) \\ &= 1/2 mv^2 + m\mathbf{a} \cdot \mathbf{v} dt = E(t) + \mathbf{F} \cdot d\mathbf{x}. \end{aligned} \quad (3.18)$$

It follows that the mass's total energy (all kinetic) has increased by

$$dE = \mathbf{F} \cdot d\mathbf{x}. \quad (3.19)$$

If the force is applied for a non-infinitesimal time Δt (during which $\mathbf{F}$ might vary), the increase in E is

$$\Delta E = \int \mathbf{F} \cdot d\mathbf{x} = \int_t^{t+\Delta t} \mathbf{F} \cdot \mathbf{v} dt. \quad (3.20)$$

³ Take care to get the minus signs right in equations (3.15)–(3.17)!

⁴ We are applying Taylor's theorem here; that is, $E(t + dt)$ means E evaluated at $t + dt$, not $E \times (t + dt)$. The same remark applies to $\mathbf{v}(t + dt)$ and $v^2(t + dt)$.

We see here why we could not have derived the non-infinitesimal (and incorrect) expression “ $\Delta E = \mathbf{F} \cdot \Delta \mathbf{x}$ ” with an argument like (3.18): because the angle between $\mathbf{F}$ and $d\mathbf{x}$ might change continuously as the mass is accelerated, making the dot product $\mathbf{F} \cdot d\mathbf{x}$ change from moment to moment. We *could* work with non-infinitesimals and a single dot product via an appropriately defined *mean* value of $\mathbf{F}$; but that definition of the mean would involve the integration in (3.20) anyway, making the reasoning somewhat circular. The bottom line is that we must integrate (3.19) when we wish to calculate a non-infinitesimal increase in energy.

Now consider the only slightly more complex situation when M is present and supplying the central force $b(r) \mathbf{u}_r$. Recall from (3.11) that the system’s total energy is

$$E(t) = -B(r) + 1/2 mv^2. \quad (3.21)$$

Again, we apply the force $\mathbf{F}$ to m for a time dt , during which m moves through $d\mathbf{x} = \mathbf{v} dt$. The total force on m is now $b(r) \mathbf{u}_r + \mathbf{F}$, and this accelerates m by an amount $\mathbf{a}$ for a time dt . Mimicking (3.18), the system’s new total energy (kinetic *plus* potential) is then

$$E(t + dt) = -B(r + dr) + 1/2 mv^2 + m\mathbf{a} \cdot \mathbf{v} dt. \quad (3.22)$$

But m ’s acceleration is now given by $m\mathbf{a} = b(r) \mathbf{u}_r + \mathbf{F}$, and (3.22) becomes

$$\begin{aligned} E(t + dt) &= -B(r) - dB + 1/2 mv^2 + [b(r) \mathbf{u}_r + \mathbf{F}] \cdot \mathbf{v} dt \\ &= E(t) - dB + [b(r) \mathbf{u}_r + \mathbf{F}] \cdot \mathbf{v} dt. \end{aligned} \quad (3.23)$$

The system’s total energy has thus increased by

$$\begin{aligned} dE &= -dB + [b(r) \mathbf{u}_r + \mathbf{F}] \cdot \mathbf{v} dt = -dB + b(r) \dot{r} dt + \mathbf{F} \cdot \mathbf{v} dt \\ &= -dB + b(r) dr + \mathbf{F} \cdot d\mathbf{x} = -dB + B'(r) dr + \mathbf{F} \cdot d\mathbf{x} \\ &= -dB + dB + \mathbf{F} \cdot d\mathbf{x} = \mathbf{F} \cdot d\mathbf{x}. \end{aligned} \quad (3.24)$$

We see again that the total energy of the system has increased by $\mathbf{F} \cdot d\mathbf{x}$. This quantity is called the *work* done on the system by the force $\mathbf{F}$.

The above calculations applied to a central force that was a function of the particles’ separation alone. A deeper study of classical mechanics ties together conserved quantities with the concepts of a lagrangian and a hamiltonian, and allows us to posit that, no matter how complex the system, and no matter how complex the nature of the forces acting, its energy can always be defined. This energy will remain fixed if no external force acts on the system, and it will increase by $\mathbf{F} \cdot d\mathbf{x}$ when this amount of work is performed on the system. These two ideas of a fixed energy and a work done form the central tenets of classical dynamics.

The above gravitational potential energy of $-GMm/r$ gives a simple example of energy conservation through the performance of work. Suppose we lift a stationary mass m through a small distance such that the force on m due to Earth's gravity can be treated as constant from start to finish. We lift m through a height h , and we require the final speed of the mass to be zero, so that only potential energy is being studied. Thus, we must lift the mass slowly, applying a force only infinitesimally greater than gravity's pull on the mass. If Earth has mass M and radius R , the force we exert near Earth's surface and throughout the lift is approximately $GMm/R^2 \mathbf{u}_r \equiv mg \mathbf{u}_r$, where $g \equiv GM/R^2$ is the "gravitational field strength" (and we ignore the slight correction to this that Earth's rotation imposes⁵). We thus do work

$$\int \mathbf{F} \cdot d\mathbf{x} \simeq \int_0^h \underbrace{mg \mathbf{u}_r}_{\simeq \mathbf{F}} \cdot \underbrace{d\mathbf{h} \mathbf{u}_r}_{d\mathbf{x}} = mg \int_0^h dh = mgh. \quad (3.25)$$

This must be the increase in the potential energy of the Earth-mass system. Is it? The increase in this energy is

$$\begin{aligned} \frac{-GMm}{r_{\text{final}}} - \frac{-GMm}{r_{\text{initial}}} &= \frac{-GMm}{R+h} - \frac{-GMm}{R} \\ &= -GMm \left[\frac{1}{R+h} - \frac{1}{R} \right] = -GMm \left[\frac{R - (R+h)}{(R+h)R} \right] \\ &\simeq \frac{GMmh}{R^2} = mgh, \end{aligned} \quad (3.26)$$

as expected. We could, of course, have used a more exact expression for $\mathbf{F}$, replacing (3.25) with

$$\begin{aligned} \int \mathbf{F} \cdot d\mathbf{x} &= \int_R^{R+h} \underbrace{\frac{GMm}{r^2} \mathbf{u}_r}_{\simeq \mathbf{F}} \cdot \underbrace{dr \mathbf{u}_r}_{d\mathbf{x}} = \left[\frac{-GMm}{r} \right]_R^{R+h} \\ &= \frac{-GMm}{R+h} - \frac{-GMm}{R}. \end{aligned} \quad (3.27)$$

This last expression indeed appears on the first line of (3.26) as the increase in the potential energy of the Earth-mass system. We are back to seeing, from (3.26), that this increase is approximately mgh .

⁵ Earth's rotation doesn't affect the gravitational force it exerts on a mass, but it *does* affect the "weight" of the mass, meaning the force required to support the mass. If Earth spun so quickly that objects on its Equator were *only just* in orbit, they would hover over the ground, and so have no weight: they would register nothing on a set of weighing scales (which is how "weight" is defined). Standing on the Equator, we could then nudge them upward using only a negligible force, according to "force equals mass times acceleration".

An important property of kinetic energy is worth proving here. First, a particle's kinetic energy $\frac{1}{2}mv^2$ is a function of its *speed* v . But if we define three separate kinetic energies, $\frac{1}{2}mv_x^2$, $\frac{1}{2}mv_y^2$, $\frac{1}{2}mv_z^2$, for the particle's motion along each of the cartesian axes, we can show that these separate energies add to produce $\frac{1}{2}mv^2$:

$$\frac{1}{2}mv_x^2 + \frac{1}{2}mv_y^2 + \frac{1}{2}mv_z^2 = \frac{1}{2}mv^2. \quad (3.28)$$

Prove this by invoking Pythagoras's theorem. In a time dt , the particle moves through an infinitesimal displacement vector (dx, dy, dz) . Pythagoras then says that it travels a distance $d\ell$, where

$$d\ell^2 = dx^2 + dy^2 + dz^2. \quad (3.29)$$

Dividing this expression by dt^2 converts these displacements to velocities:

$$\begin{aligned} v^2 &= (d\ell/dt)^2 = (dx/dt)^2 + (dy/dt)^2 + (dz/dt)^2 \\ &= v_x^2 + v_y^2 + v_z^2. \end{aligned} \quad (3.30)$$

Multiplying all terms here by $\frac{1}{2}m$ gives us (3.28).

3.2 Force and Potential Energy

Discussing potential energy is a good place to segue into a proof of the standard expression relating potential energy to force. Suppose that a particle has potential energy U in some possibly complicated field.⁶ What force $\mathbf{F}$ (not necessarily central) does the field produce on the particle? While the field moves the particle through some arbitrary displacement $d\mathbf{x}$, we will apply a counteracting force infinitesimally weaker than $-\mathbf{F}$: call it $-\mathbf{F} + \boldsymbol{\varepsilon}$. This allows the field to push the particle through $d\mathbf{x}$ without the particle's kinetic energy increasing; this lets us isolate the potential energy for analysis. We know from the foregoing that when *we* apply the force $-\mathbf{F} + \boldsymbol{\varepsilon}$ to the particle, its total energy increases by the work we do:

$$d(\text{total energy}) = (-\mathbf{F} + \boldsymbol{\varepsilon}) \cdot d\mathbf{x} = -\mathbf{F} \cdot d\mathbf{x}, \quad (3.31)$$

where the $\boldsymbol{\varepsilon} \cdot d\mathbf{x}$ is of second order, and so can be ignored. Because the particle's kinetic energy is held constant, then, after this procedure, the potential energy—and hence the total energy—will have increased by

⁶ As mentioned above shortly after (3.10), the potential energy U really belongs to the entire system. But imagining that energy to be “owned” by the particle appeals to our intuition when discussing the particle's dynamics in a frame in which the rest of the system is fixed.

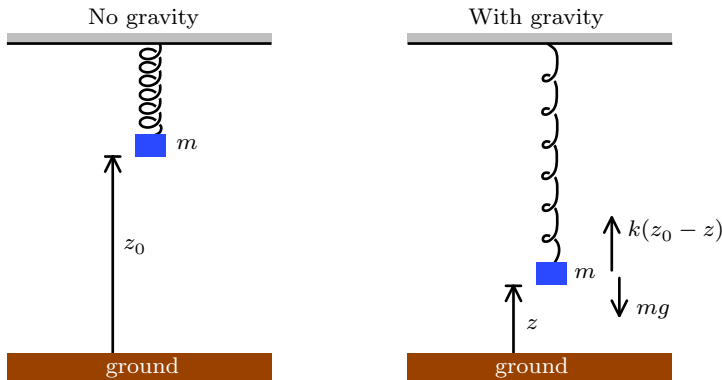


Fig. 3.2 **Left:** A relaxed spring is placed vertically in the absence of gravity. The mass at its end lies at height z_0 above the ground. **Right:** Now switch gravity g on. Gravity stretches the spring (with spring constant k), giving the mass a new equilibrium position at height z . What is z in terms of z_0, m, k, g ?

$$dU = d(\text{total energy}) = -\mathbf{F} \cdot d\mathbf{x}. \quad (3.32)$$

But (1.170) tells us that $dU = \nabla U \cdot d\mathbf{x}$. Since $d\mathbf{x}$ is arbitrary, we conclude that

$$\mathbf{F} = -\nabla U. \quad (3.33)$$

The above might be rephrased by saying that when the field moves the particle through some arbitrary $d\mathbf{x}$, the work done by the field equals the loss in the potential energy U :

$$\begin{aligned} \mathbf{F} \cdot d\mathbf{x} &= \text{work done by field} = \text{loss in } U \\ &= -dU \stackrel{(1.170)}{=} -\nabla U \cdot d\mathbf{x}. \end{aligned} \quad (3.34)$$

This must hold for *any* $d\mathbf{x}$, and so again it follows that $\mathbf{F} = -\nabla U$.

Now, this is all very well, but what can be said when more than one field is present? For example, refer to Figure 3.2, which shows a spring held vertically without and then with gravity present. Without gravity, its equilibrium length means that the mass m at its end lies at a height z_0 above the ground. When we switch gravity g on, the mass is pulled down to a new equilibrium position at a height z above the ground. The question is, what is z ? This is a standard question in the theory of statics. As gravity applies a downward force mg on the mass, the spring stretches, and responds according to Hooke's Law with an upward tension force of $k \times \text{stretch} = k(z_0 - z)$, where k is its spring constant. In its new position, the mass doesn't accelerate; hence, the two forces must balance, giving

$$k(z_0 - z) = mg. \quad (3.35)$$

This is easily solved for z —the precise result doesn't concern us. The question is, what can be said about potential energy here? Forces are vectors, so let's be more precise by introducing the unit basis vector in the z direction, $\mathbf{u}_z$. The forces on the mass are then

$$\mathbf{F}_{\text{spring}} = k(z_0 - z) \mathbf{u}_z, \quad \text{and} \quad \mathbf{F}_{\text{gravity}} = -mg \mathbf{u}_z. \quad (3.36)$$

Equation (3.35) is really a statement that the total force on the mass is zero:

$$\mathbf{F}_{\text{total}} = \mathbf{F}_{\text{spring}} + \mathbf{F}_{\text{gravity}} = [k(z_0 - z) - mg] \mathbf{u}_z = \mathbf{0}. \quad (3.37)$$

But the force applied by the spring, $\mathbf{F}_{\text{spring}}$, is related to the potential energy due to the spring, U_{spring} , and similarly for gravity:

$$\mathbf{F}_{\text{spring}} = -\nabla U_{\text{spring}}, \quad \mathbf{F}_{\text{gravity}} = -\nabla U_{\text{gravity}}. \quad (3.38)$$

The total force on the mass is then

$$\begin{aligned} \mathbf{F}_{\text{total}} &= -\nabla U_{\text{spring}} - \nabla U_{\text{gravity}} \\ &= -\nabla(U_{\text{spring}} + U_{\text{gravity}}) \quad (\text{because } \nabla \text{ is linear}) \\ &= -\nabla U_{\text{total}}, \end{aligned} \quad (3.39)$$

where

$$U_{\text{total}} \equiv U_{\text{spring}} + U_{\text{gravity}}. \quad (3.40)$$

This says that we can define a total potential energy of the system to be the *sum* of the potential energies arising from the various fields involved. This might seem like a natural thing to do, but it is certainly not something obvious that can be stated a priori and taken for granted. In equilibrium, when the total force must equal zero, we have

$$-\nabla U_{\text{total}} = \mathbf{0}, \quad \text{or simply} \quad \nabla U_{\text{total}} = \mathbf{0}. \quad (3.41)$$

It follows that the position of the mass in equilibrium is that which *extremises the total potential energy*.⁷

To see how this approach of extremising the total potential energy reproduces (3.35), we need expressions for U_{spring} and U_{gravity} . We saw above, in (3.25), that $U_{\text{gravity}} = mgz$; check this by calculating $\mathbf{F}_{\text{gravity}}$:

$$\mathbf{F}_{\text{gravity}} = -\nabla U_{\text{gravity}} = -\nabla mgz = -mg \mathbf{u}_z. \quad (3.42)$$

This agrees with (3.36). Next, what is U_{spring} ? Most springs follow the Hooke's Law assumption that “tension is proportional to extension”. Figure 3.3 shows the spring before and after being stretched through a displacement x , which

⁷ A stable equilibrium will equate to *minimising* the total potential energy. But that is a topic for a course on classical mechanics.

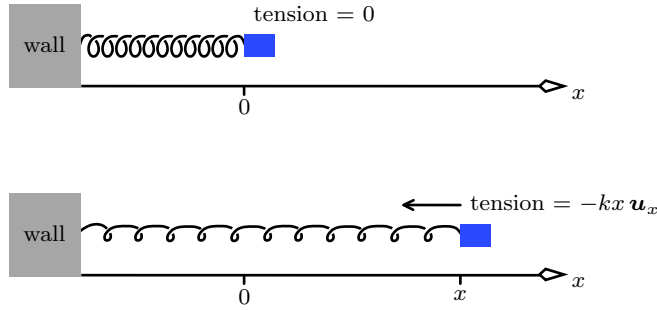


Fig. 3.3 Top: The spring and mass in their natural position with no stretch, and thus no tension in the spring. **Bottom:** The mass is pulled through a displacement x slowly, to give it only potential energy. If the spring conforms to Hooke's law (which all well-behaved springs do), its tension is always proportional to x

can be positive or negative. (We draw it horizontally here to prevent any confusion with gravity.) Regardless of the sign of x , the tension in the spring acts oppositely to the displacement from its equilibrium position:

$$\text{tension} = -kx \mathbf{u}_x, \quad (3.43)$$

where $k > 0$ is the spring constant and $\mathbf{u}_x$ is a unit basis vector pointing in the direction of increasing x . To do work slowly on the spring, and thus alter its potential energy (and give it no kinetic energy), *we* must apply a force that is equal but opposite to the tension.⁸ We move the mass through a displacement x :

$$\text{force we apply} = kx \mathbf{u}_x, \text{ in increments of } d\mathbf{x} = dx \mathbf{u}_x. \quad (3.44)$$

If U_{spring} is defined to be zero at the spring's equilibrium position $x = 0$, then U_{spring} equals the work we do when we stretch or compress the spring from that equilibrium position:

$$\begin{aligned} U_{\text{spring}} &= \int_{\text{initial}}^{\text{final}} (\text{force we apply}) \cdot d\mathbf{x} \\ &= \int_0^x kx \mathbf{u}_x \cdot dx \mathbf{u}_x = \int_0^x kx dx = 1/2 kx^2. \end{aligned} \quad (3.45)$$

Check this by calculating the tension:

$$\text{tension} = -\nabla U_{\text{spring}} = \frac{-dU_{\text{spring}}}{dx} \mathbf{u}_x = -kx \mathbf{u}_x. \quad (3.46)$$

⁸ Remember that “equal but opposite” is a common vector expression that is meant to be descriptive, and translates simply to “minus”.

This matches (3.43), as expected. In general, we can write

$$U_{\text{spring}} = 1/2 k \times (\text{length of stretch or compression})^2. \quad (3.47)$$

Now that we have an expression for the spring's potential energy, return to the scenario of Figure 3.2. Write

$$U_{\text{spring}} = 1/2 k(z_0 - z)^2, \quad U_{\text{gravity}} = mgz. \quad (3.48)$$

The total potential energy U_{total} is the sum of these. Extremising it via $dU_{\text{total}}/dz = 0$ yields

$$\frac{dU_{\text{total}}}{dz} = -k(z_0 - z) + mg = 0. \quad (3.49)$$

This expression matches (3.35). (Also, $d^2U_{\text{total}}/dz^2 = k > 0$, which shows that we have indeed found a *minimum* in the total potential energy.) We see how the approach of using forces to solve this task of finding the equilibrium height z tallies with the approach of minimising potential energy.⁹

3.3 Interaction Types and the Zeroth Law of Thermodynamics

Energy conservation works perfectly well for gravity and the Coulomb force. For example, a falling mass can be made to do work: the kinetic energy that is extracted from the gravitational potential can be converted into a force that does useful work for us. Historically, this was the idea behind the water wheel and, in modern times, hydroelectric power.

But the invention of the steam engine in the seventeenth century brought a complication to the playing field. Useful work could also be extracted from, say, heating water and allowing the steam to push a piston. And yet, it was not clear what the necessary heat really was or did.

In the eighteenth century, heat was thought to be a substance called *caloric* that was conserved in interactions: when a hot body was placed in contact with a cold body, caloric was thought to pass from the hot into the cold body until they reached a common temperature. But by the end of that century, experiments with friction showed that caloric cannot be conserved: friction could create heat indefinitely without anything apparently being used up. At this time, Count Rumford conducted experiments in boring cannon, and found that the heat generated seemed to match the mechanical work done

⁹ The common expression “minimising a system’s energy” in textbooks should be understood to mean minimising its *potential* energy, as we did above to find the equilibrium position z of the mass on the spring.

by the boring tool. By the 1840s, experiments performed mainly by James Joule produced the modern view that caloric does not exist; rather, heat is a manifestation of internal kinetic energy, which must be quantified and worked into the book-keeping if energy is to remain a conserved quantity.

The result of this research up until the mid nineteenth century was the *First Law of Thermodynamics*, which focusses on three ways in which energy can be transferred:

1. Heating: *thermal interactions* (conduction, convection, radiation).
2. Performing work: *mechanical interactions* (pressure, electromagnetic forces).
3. Transferring particles: *diffusive interactions* (chemical changes, atmospheres, permeable membranes).

Classical statistical mechanics always assumes that interacting systems are either in equilibrium or are very close to it; so, naturally, it's of primary importance to relate the different interaction types to the overriding concept of equilibrium. Hence, before we can discuss the First Law in detail, we require the *Zeroth Law of Thermodynamics*:

The Zeroth Law of Thermodynamics

If two systems are in thermal/mechanical/diffusive equilibrium with a third system, then they're in thermal/mechanical/diffusive equilibrium with each other.

The Zeroth Law is given priority over the First Law in the numbering scheme because it lays the foundation for our even beginning to speak of systems interacting. Once we analyse such interactions, we find that the relevant energy is always conserved across an interaction. The ways in which a system can acquire energy are described by the First Law. This law involves the inexact differentials of Section 1.8, and so first, we place these into context.

Suppose we are given a container of hot gas whose history is unknown. Thermodynamics deals with the processes this gas could have undergone to bring it to its present state. In particular, we ask how the gas acquired its current temperature. Perhaps it was heated over a stove ("thermal" in the Zeroth Law), or had work done on it ("mechanical"), or its chemical environment was changed ("diffusive"), or some combination of these three. Knowing nothing of the gas's history, we cannot ascertain just how it reached its current state. To make progress, we need, at least, to label the ways in which energy can be added to it. This is not a labelling of any types of energy, but rather a labelling of the processes via which energy can be transferred to the gas. Nevertheless, the following labels *are* routinely used to quantify the various amounts of energy transferred via different processes:

1. Heating the gas over a stove while doing no mechanical work or transferring particles involves a transfer of internal energy in a process we quantify by labelling that energy as Q . An infinitesimal amount of this “thermally transferred energy” is written as the inexact differential $\mathrm{d}Q$.
2. Performing mechanical work on the gas involves a transfer of energy that we label as W . An infinitesimal amount of work done on the system is written as the inexact differential $\mathrm{d}W$. (Take note: some texts write our $\mathrm{d}W$ as $-\mathrm{d}W$.)
3. Transferring particles or changing their chemical environment can be called a chemical transfer of energy, and an infinitesimal amount of this energy transferred to the system is written as the inexact differential $\mathrm{d}C$.

Because these processes have no effect other than to increase the internal energy of the gas, the answer as to how this energy was transferred—whether thermally, mechanically, or chemically—is ultimately of no consequence. *All* energy added to the gas becomes indistinguishable internally. This is entirely analogous to transferring money to a bank account: the end result is the same whether we deposit a cheque, cash, or move the funds electronically. These funds lose their identity when banked, and all of these transfer processes simply increase an internal parameter called “money” in the bank account. We would even be hard pressed to locate this “money” were we to search for it. If the amount of money in our bank account is (for some reason) fixed, then the funds transferred into and out of it through all of these processes must be related, and cannot be treated independently.

Likewise, we cannot maintain that the gas has unique values of “heat” Q , “work” W , and “chemical energy” C associated with it: none of these quantities describe the state of the gas, just as a bank account has no notion of dividing its money into cash and “cheque money”. The gas has no state variables called Q , W , C , and this is why $\mathrm{d}Q$, $\mathrm{d}W$, $\mathrm{d}C$ are written as inexact differentials. For that reason, just as in the discussion immediately following (1.195) of the distance s covered by a hiker, we will always reserve the symbol Q —and never ΔQ —for a non-infinitesimal amount of energy transferred thermally to a system. Analogous comments apply to W and C .

3.4 The First Law of Thermodynamics

The union of the apparently unrelated concepts of heat transfer with mechanical and chemical energy marked a watershed in the history of physics, and produced the *First Law of Thermodynamics*:

The First Law of Thermodynamics

The infinitesimal increase $\mathrm{d}E$ in a system’s internal energy E is the sum

of thermal, mechanical, and diffusive contributions:

$$dE = dQ + dW + dC, \quad (3.50)$$

where

- dQ = energy put **into** the system thermally by an environment, such as a stove;
- dW = mechanical work performed **on** the system by forces arising from pressure, electromagnetism, etc.;
- dC = energy brought **into** the system, either by particles that arrive as a result of environmental changes, or by the environment itself. This includes large-scale potential energy, such as that due to gravity when we are treating a large system such as an atmosphere.

But whereas the First Law allowed for a proper book-keeping of heat as “just” another form of energy transfer, analytically, the law was practically useless, on account of (3.50) being expressed using *inexact* differentials. It was far more desirable to express the law using only exact differentials of some choice of state variables, since that would bring calculus to bear on First Law analyses. Of course, referring to state variables requires a state to *exist*, by which we mean a set of external variables can be defined: pressure, volume, and so on.

Such a set of variables cannot always be chosen. For example, Figure 3.4 shows the *free expansion* of an ideal gas initially confined by a removable partition to the left part of an empty box. An ideal gas has no inter-particle forces; hence, its energy is solely kinetic. The walls of the box are *adiabatic*,

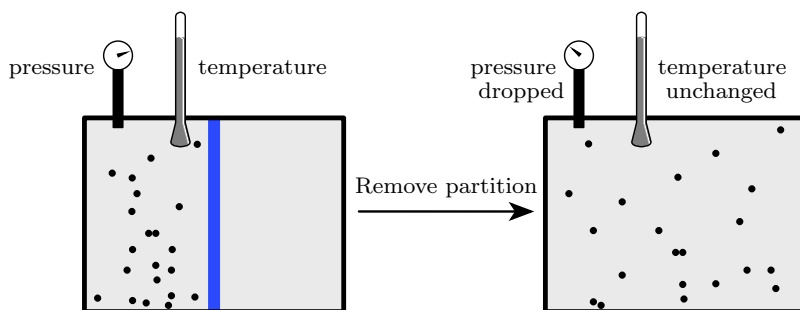


Fig. 3.4 Free expansion of an ideal gas, which is initially confined to the left part of an empty box by a removable partition (shown in blue). After the partition is removed, the gas spreads out to occupy the whole box. Experimentally, its pressure is found to drop, while its temperature remains fixed

meaning thermally insulating: the gas cannot exchange energy thermally with its environment.¹⁰ The partition is suddenly removed, and the gas particles expand freely to occupy the larger volume. Experimentally, we find that the gas quickly comes to equilibrium at a reduced pressure but with the *same* temperature. We'll see later that the temperature is a measure of the gas particles' speeds; and since these speeds cannot be affected by the partition's removal, the gas's temperature does not change in the process. But although the gas's temperature is well defined throughout the expansion, its pressure and volume are another case entirely. These variables have well-defined initial and final values, but are not defined *during* the expansion. We'll see in Section 3.6 that pressure is defined by the rate at which the particles of a system interact with a set of confining walls; so, when these walls do not all exist—such as during the above free expansion—pressure cannot be defined. This lack of a full set of confining walls during the expansion also means that the freely expanding gas lacks a well-defined volume.

The thermodynamics of non-equilibrium processes such as free expansion is often analysed by examining small regions of the system for which state variables *can* be defined to some precision. As stated in Section 2.1, we will always assume our processes to be quasi-static: ones that are always very close to equilibrium while they evolve. Being very close to equilibrium means that a well-defined set of state variables exists for the duration of the process being studied. For example, letting a gas expand quasi-statically can be accomplished by allowing its pressure to move a piston whose other side contacts an external gas that always supplies a slightly lower back pressure. The confined gas does work on the external gas as it pushes on the piston, and thus transfers some of its energy to the external gas. To mimic the above case of free expansion in which the confined gas's energy remains constant, we could heat the confined gas just enough to supply that energy, that must be handed over to the external gas. At all times, such a system has a well-defined pressure, volume, and temperature.

Returning to the First Law (3.50), our goal is to produce a version of it that applies to quasi-static processes, one that uses only exact differentials of state variables. The most difficult of its terms to tackle is the first, $\mathrm{d}Q$, and so we'll leave that until we have addressed its other terms, $\mathrm{d}W$ and $\mathrm{d}C$.

3.4.1 Expressions for Quasi-Static Mechanical Work

Section 3.1 showed that the key expression for the mechanical work performed by a force $\mathbf{F}$ acting on a mass that moves through displacement $\mathrm{d}\mathbf{x}$ is $\mathbf{F} \cdot \mathrm{d}\mathbf{x}$. In Section 3.2, we explored the potential energy of a stretched spring that

¹⁰ “Adiabatic” comes from the Greek “*adiabatos*”, meaning “impassable”. Be aware that some authors describe a process as adiabatic when they really mean quasi-static.

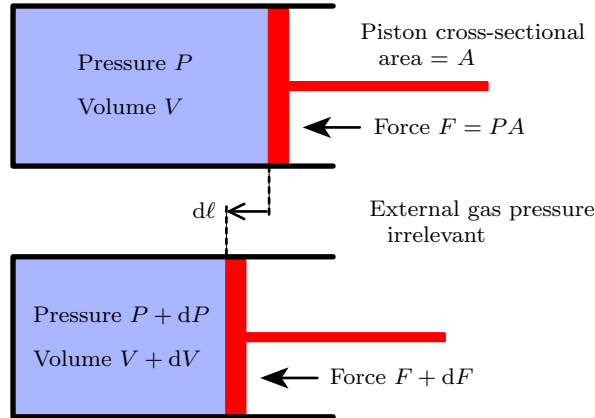


Fig. 3.5 Quasi-static work is done by a piston on the gas in a cylinder by applying infinitesimally more pressure than that of the gas, P . The work done on the gas is $dW = F d\ell$. The loss in the gas's volume is $-dV = A d\ell$, so it follows that $dW = F \times -dV/A = -PdV$

follows Hooke's Law of "tension proportional to extension", as most springs do. We calculated the work done in stretching the spring in (3.45). That equation had the following form:

increase in potential energy = work we do

$$= \int dW = \int_{\text{initial}}^{\text{final}} (\text{force we apply}) \cdot d\mathbf{x} = \int \mathbf{F} \cdot d\mathbf{x}. \quad (3.51)$$

Here, we have managed to replace the inexact differential dW with the exact differential $\mathbf{F} \cdot d\mathbf{x}$.

Another classic example of a mechanical interaction is the work performed on a gas by compressing it in a cylinder, as shown in Figure 3.5. The gas in the cylinder has pressure P and volume V . Irrespective of the pressure of any atmosphere outside the cylinder, a pressure P must be applied to the piston just to hold it still. When we increase this pressure infinitesimally, the piston moves to compress the gas, which retains a well-defined value of pressure throughout. The piston of cross-sectional area A exerts a force $F = PA$ on the gas, and that piston moves through a distance $d\ell$. We thus do work $dW = F d\ell$. As the gas compresses, it loses volume $-dV = A d\ell$, from which it's clear that we do work on the gas equal to

$$dW = F d\ell = F \times -dV/A = -PdV. \quad (3.52)$$

(This is equivalent to saying that the gas does work PdV .) The volume of the gas is a state variable; so again, we have managed to replace the inexact differential dW with the exact differential $-PdV$.

A point about the infinitesimals is relevant here. Because the compression is carried out quasi-statically, the pressure we apply to the piston can be taken to be exactly P . We might say instead that a slightly higher pressure $P + dP$ must really be applied to overcome the pressure P of the gas, and this will change (3.52) to $dW = -(P + dP)dV$. But this expression differs from the infinitesimal in (3.52) only by a second-order term $-dP dV$, and so—recalling the discussion just after (1.133)—this additional term can be ignored.

In contrast to (3.52), you will sometimes see the expression “ $dW \geq -PdV$ ” in discussions of thermodynamics. The origin of the “ $\geq$ ” sign can be somewhat obscure: it is meant to signal that other systems are being included in the energy book-keeping, making us do more work than just the $-PdV$ that it takes to compress the gas. But it can create confusion about which system is being discussed. For (3.52), the dW specifically refers to the gas confined by the piston and cylinder, and not to anything else. Of course, the cylinder might be a little bit rusty and lacking a perfect lubricating oil, so that we must do some extra work in scraping the piston down the cylinder bore: those who write “ $dW \geq -PdV$ ” are including that extra work in the dW . This work of scraping could be described mathematically by other terms, but it simply is not being discussed in (3.52), which remains a strict equality that refers to the *gas* only. This is an important point to realise, because thermodynamics discussions often involve inequalities relating to the work done, and it’s of central importance always to know which system or systems are having the work done on them. Our discussion of the First Law always refers to a single well-defined system. In the case of the gas being compressed, the quasi-stasis¹¹ ensures that the gas always has a well-defined pressure and volume, and it’s these that the quasi-static expression $dW = -PdV$ refers to.

Other types of work exist, such as that performed by an electric field $\mathbf{E}$ when it rotates an electric dipole $\mathbf{p}$, or the work performed by a magnetic field $\mathbf{B}$ when it rotates a magnetic dipole $\boldsymbol{\mu}$. It’s insightful to calculate these two amounts of work.

Examine first the electric case, shown in Figure 3.6. An *electric dipole* is composed of two equal-but-opposite charges, $q > 0$ and $-q$, that are rigidly connected, in the sense that their separation a is fixed while they are free to rotate. This simple system is characterised by its *electric dipole moment* $\mathbf{p}$, a vector pointing from $-q$ to q with magnitude qa . The electric field $\mathbf{E}$ tries to rotate the dipole to align $\mathbf{p}$ with $\mathbf{E}$. We do quasi-static work on the dipole by turning it against the field with just enough force $\mathbf{F}$ to override the field’s force on the charges. Without loss of generality, set $\mathbf{E}$ parallel to the x axis, with $\mathbf{p}$ at angle θ to $\mathbf{E}$. We do work dW on the dipole by turning it through angle $d\theta$ against the field, as shown in the figure. This work is the sum of the work done on each charge. We apply a force $\mathbf{F} = -\mathbf{E}q$ to q (moving it through $d\ell$) and $-\mathbf{F}$ to $-q$ (moving it through $-d\ell$). The work we do is

¹¹ Quasi-stasis means “quasi-staticity”.

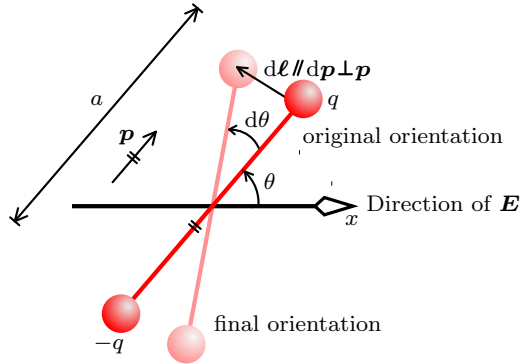


Fig. 3.6 An electric dipole is a pair of charges $q > 0$ and $-q$ separated by a distance a . It is free to rotate in an electric field $\mathbf{E}$. Its “dipole moment” $\mathbf{p}$ points from $-q$ to q , with magnitude qa . The infinitesimal vector $d\boldsymbol{\ell}$ that the top charge rotates through is always parallel to $d\mathbf{p}$, and these are perpendicular to $\mathbf{p}$

$$\begin{aligned} dW &= \mathbf{F} \cdot d\boldsymbol{\ell} + (-\mathbf{F}) \cdot (-d\boldsymbol{\ell}) = 2\mathbf{F} \cdot d\boldsymbol{\ell} = -2Eq \cdot d\boldsymbol{\ell} \\ &= -2q \begin{bmatrix} E \\ 0 \end{bmatrix} \cdot \frac{a}{2} d\theta \begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix} = Ep \sin \theta d\theta. \end{aligned} \quad (3.53)$$

But $p d\theta = |d\mathbf{p}|$, and so¹²

$$dW = E |d\mathbf{p}| \sin \theta = -E |d\mathbf{p}| \cos(\theta + \pi/2) = -\mathbf{E} \cdot d\mathbf{p}. \quad (3.54)$$

For completeness, we can associate a potential energy $U(\theta)$ with this orientation of the dipole:

$$\begin{aligned} U(\theta) &= U(0) + \int_0^\theta dW = U(0) + \int_0^\theta Ep \sin \theta d\theta \\ &= U(0) - Ep \cos \theta + Ep. \end{aligned} \quad (3.55)$$

Potential energy is, of course, only defined up to an additive constant, which allows us to set $U(0)$ to any value we choose. It is conventionally set to $-Ep$ to simplify (3.55), resulting in

$$U = -Ep \cos \theta = -\mathbf{E} \cdot \mathbf{p}, \quad (3.56)$$

where we have replaced “ $U(\theta)$ ” simply with “ U ”, because the dot product is independent of the coordinate system chosen.

¹² Be aware of the difference between $|d\mathbf{p}| \neq 0$ and $d\mathbf{p} \equiv d|\mathbf{p}| = 0$. We are not interested in $d\mathbf{p}$.

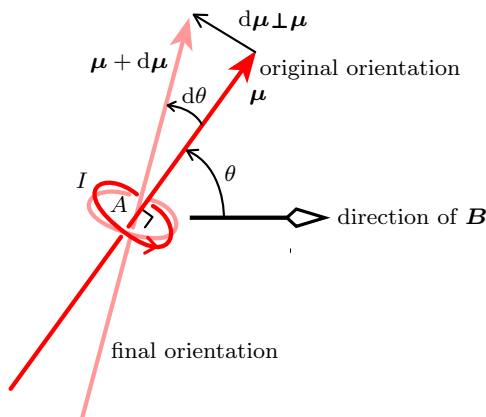


Fig. 3.7 A magnetic dipole is composed of a planar loop of area A carrying current $I > 0$, and is represented by a vector $\boldsymbol{\mu}$ of length IA that points normal to the current plane in a right-handed sense. The dipole can rotate in a magnetic field $\mathbf{B}$

The dipole moment of a macroscopic collection of dipoles is defined as the sum of its individual dipole moments. We can imagine performing work on such a collection one dipole at a time, and so the expression $dW = -\mathbf{E} \cdot d\mathbf{p}$ in (3.54) holds even in this macroscopic case.

Performing work on a *magnetic dipole* follows a similar treatment to the electric case above. A magnetic dipole is a closed planar loop of area A and carrying current $I > 0$, as shown in Figure 3.7. Its *magnetic dipole moment* $\boldsymbol{\mu}$ is a vector whose length is defined to be IA and which points normal to the planar loop, right-handed to the direction of the current. Unlike the electric case with its two charges and two forces, the magnetic case involves an infinite number of forces $d\mathbf{F} = I d\boldsymbol{\ell} \times \mathbf{B}$ to be calculated for the infinitesimal line segments $d\boldsymbol{\ell}$ of current that comprise the loop. It becomes easier here to use the language of torque: we will call on a standard result of magnetostatics, which says that the torque exerted by the field on the current loop is $\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B}$. Performing quasi-static work dW against this torque increases the angle θ in Figure 3.7 to $\theta + d\theta$:

$$dW = \tau d\theta = |\boldsymbol{\mu} \times \mathbf{B}| d\theta = \mu B \sin \theta d\theta. \quad (3.57)$$

But $\mu d\theta = |d\boldsymbol{\mu}|$, and so

$$dW = B |d\boldsymbol{\mu}| \sin \theta = -B |d\boldsymbol{\mu}| \cos(\theta + \pi/2) = -\mathbf{B} \cdot d\boldsymbol{\mu}. \quad (3.58)$$

Just as in the electric case, for completeness, we can associate a potential energy $U(\theta)$ with this orientation of the magnetic dipole:

$$\begin{aligned}
 U(\theta) &= U(0) + \int_0^\theta \mathrm{d}W = U(0) + \int_0^\theta \mu B \sin \theta \, \mathrm{d}\theta \\
 &= U(0) - \mu B \cos \theta + \mu B .
 \end{aligned}
 \tag{3.59}$$

$U(0)$ is conventionally set to $-\mu B$ to simplify (3.59), yielding

$$U = -\boldsymbol{\mu} \cdot \mathbf{B} , \tag{3.60}$$

where we again omit the dependence on θ , following the comment after (3.56).

The total moment of a set of magnetic dipoles is the sum of its individual dipole moments. Just as for the electric case above, we can consider the work performed on such a collection to be done one dipole at a time, and so conclude that $\mathrm{d}W = -\mathbf{B} \cdot \mathrm{d}\boldsymbol{\mu}$ also holds for a macroscopic set of dipoles.

The sum $\mathrm{d}W$ of the various ways in which a system can absorb energy mechanically and quasi-statically is sometimes written as a single generic term involving the relevant exact differentials. For example, a system that combines electric and magnetic effects might have

$$\mathrm{d}W = -\mathbf{E} \cdot \mathrm{d}\mathbf{p} - \mathbf{B} \cdot \mathrm{d}\boldsymbol{\mu} . \tag{3.61}$$

We could expand this expression as

$$\begin{aligned}
 \mathrm{d}W &= \underbrace{-(E_x, E_y, E_z, B_x, B_y, B_z)}_{\equiv \mathbf{f}} \cdot \underbrace{(\mathrm{d}p_x, \mathrm{d}p_y, \mathrm{d}p_z, \mathrm{d}\mu_x, \mathrm{d}\mu_y, \mathrm{d}\mu_z)}_{\equiv \mathrm{d}X} \\
 &= \mathbf{f} \cdot \mathrm{d}X ,
 \end{aligned}
 \tag{3.62}$$

where the dot product is understood as though it were acting on six-dimensional vectors expressed in cartesian form. The term $\mathbf{f}$ is a *generalised force* (such as pressure: it need not have the dimension of force) and $\mathrm{d}X$ is a *generalised displacement* (such as volume: it need not have the dimension of distance). Writing¹³ $\mathrm{d}W = \mathbf{f} \cdot \mathrm{d}X$ is a little too generic for the calculations that we will do in the chapters to come, and so we will represent the various types of work described above by the compression of gas in a cylinder, since this type of work dates from the earliest days of thermodynamics, when the subject was completely rooted in the industrial applications of the time. That is, we'll replace $\mathrm{d}W$ in a *quasi-static* version of the First Law with $-P\mathrm{d}V$, as a generic way of representing the work performed quasi-statically on a system. It's clear in the examples analysed above that we have always man-

¹³ Be aware that some authors might insert a minus sign into this expression, if they define $\mathrm{d}W$ as our $-\mathrm{d}W$.

aged to replace the inexact differential dW with an exact differential, now represented generically by $-P dV$.

3.4.2 The dC Term and Chemical Potential

The dC in (3.50) is the energy brought into the system by incoming particles via chemical changes or the influence of a field such as gravity. This energy can be expressed using the “particle potential energy” μ , conventionally called the *chemical potential*, because it often appears from bond rearrangements between atoms in chemical reactions. The number N of interacting particles is always a whole number, of course, but it can be treated as continuous for the large numbers that are standard in statistical mechanics, and so it’s normal to speak of adding dN particles to an N -particle system. These particles bring in an average energy of μ per particle. The total energy added is then

$$dC = \mu dN. \quad (3.63)$$

In the chapters to come, we will explore how μ relates to an atmosphere’s pressure variation with height, its role in chemical reactions, and its very important position in quantum statistics. We’ll also investigate how μ determines the details of electric current and a metal’s heat capacity.

Recall the difference between “potential” and “potential energy”: potential is defined as potential energy per unit of the relevant quantity. For example, gravitational potential is gravitational potential energy per unit mass; electrostatic potential is electrostatic potential energy per unit charge. Similarly, chemical potential is chemical potential energy per particle. Thus, unlike the gravity and electrostatic cases, chemical potential has units of energy.

One example of a changing chemical potential occurs when we add water to a concentrated acid: the resulting mixture can quickly grow dangerously hot. (Always add a concentrated acid to water to dilute it, never the reverse.)

An easy way to derive insight into the meaning of μ is to relate it to the fact that a gas has weight: the weight of air in our atmosphere means that air higher up pushes on air lower down and compresses it, causing the density and pressure to be higher closer to the ground. Begin by picturing two equal volumes of gas with identical particle densities ν (that is, numbers of particles per unit volume), where each gas particle has mass m . Figure 3.8’s left-hand picture shows these volumes held in boxes sitting at ground level $z = 0$. We write μ as a function of the box height z and the particle-number density ν ; thus, each box has a chemical potential $\mu(z = 0, \nu)$.

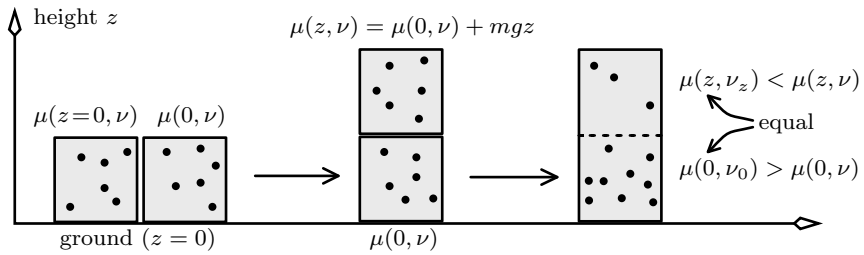


Fig. 3.8 **Left:** Two identical boxes of gas. **Middle:** One box is placed on top of the other, increasing the top box's chemical potential by the gravitational potential energy mgz given to each particle. **Right:** The walls separating the boxes are now removed to allow the particles to flow freely. In equilibrium, more particles are now present in the bottom box than in the top box, and the two boxes again have equal chemical potentials

Now lift one box up to height z and place it on the other, as shown in the middle in Figure 3.8. The particle density in each box cannot change, but each particle in the top box has been given gravitational potential energy mgz , and this just adds to the top box's value of μ :

$$\mu(z, \nu) = \mu(0, \nu) + mgz. \quad (3.64)$$

Finally, remove the walls separating the two boxes so that particles can flow freely between them, shown at the right in the figure. The weight of the particles in the top box causes them to settle downward until the pressure in the lower box has increased sufficiently to halt any further settling. We can view this initial settling of the upper particles as indicating that the upper box initially has a higher value of μ than the lower box. When equilibrium is reached, the upper box's particle density has dropped to a new value ν_z and its chemical potential has also dropped. The lower box's particle density has increased to a new value ν_0 and its chemical potential has also increased. These two new chemical potentials are equal in equilibrium, because particles can only flow under a *gradient* of chemical potential (which we show in Section 3.13). Hence,

$$\mu(z, \nu_z) = \mu(0, \nu_0). \quad (3.65)$$

But (3.64) holds quite generally, and so can be written with ν_z in place of ν :

$$\mu(z, \nu_z) = \mu(0, \nu_z) + mgz. \quad (3.66)$$

Comparing (3.65) and (3.66) shows that their right-hand sides must be equal:

$$\mu(0, \nu_0) = \mu(0, \nu_z) + mgz. \quad (3.67)$$

We will return to this idea in Section 4.3.3 to derive the standard exponential drop of pressure with height in an atmosphere.

In the same way as pressure, the chemical potential is not defined for a non-quasi-static process. When the two boxes in Figure 3.8 were allowed to exchange particles, their chemical potentials were only well defined before and after the mixing process. So, provided we only consider quasi-static processes, we can replace dC with μdN .

Summarising the last few pages, we have succeeded in replacing the inexact differentials dW and dC of the First Law (3.50) with the exact differentials $-PdV$ and μdN , respectively, for quasi-static processes. For these processes, the First Law so far has the form

$$dE = dQ - PdV + \mu dN. \quad (3.68)$$

Finding an exact differential to replace the thermal transfer dQ will be a more difficult task. Because heat is quantified by temperature, we begin by defining temperature.

3.5 The Definition of Temperature

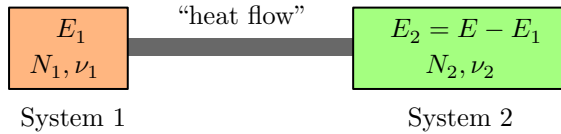
The counting of microstates described in Chapter 2 sheds light on the historically elusive notion of “heat transfer”. We’ll study the thermal transfer of energy in Section 4.1.2; but that requires a definition of temperature. It turns out that temperature can be defined by examining the number of states available to two thermally connected systems that are evolving toward a shared thermal equilibrium.

The emergence of thermodynamics in the middle of the nineteenth century changed our view of heat from being a substance in its own right, “caloric”, to being a manifestation of the transfer of energy at a microscopic level; caloric does not exist, and instead, what we perceive as heat reflects the fact that the distribution of internal energy tends to become uniform (slowly or quickly) throughout the combined system when a hot body and a cold body come into contact. This modern view treats heat flow as that which results when a huge number of “billiard-ball collisions” between atoms transfer internal energy from the hot to the cold body. This approach essentially renders the noun “heat” outdated, while retaining the verb “to heat”. Even so, the use of “heat” as a noun along with the term “heat flow” are so well understood in a modern context that no harm results from using them. In fact, using “heat” as a noun is an efficient way for a physicist to speak and think. We all understand that when “heat flows” between two systems, those systems are “interacting thermally”, meaning energy is able to flow between them with no mechanical or diffusive interaction necessary. We examine this type of energy flow now, and will use it to define temperature.

3.5.1 Accessible Microstates for Thermally Interacting Systems

Begin with two systems, “1” and “2”, that interact thermally, but not mechanically or diffusively: apparently, “heat is flowing” from the hot body to the cold body, and our job here is to quantify what we mean by “heat is flowing”, “hot”, and “cold”. The combined system is isolated from the rest of the world, and so its total energy is fixed at some E . It’s important to note that there is no background potential energy that could drive particles or energy in some direction.

At a given moment, system 1 has total energy E_1 and N_1 particles, each with ν_1 quadratic energy terms; and system 2 has total energy E_2 and N_2 particles, each with ν_2 quadratic energy terms. Given that the total energy $E = E_1 + E_2$ is fixed, we can eliminate E_2 by writing $E_2 = E - E_1$ from the outset:



We ask: how does the total number of accessible microstates Ω of the combined system vary as a function of E_1 ? (We will refer to these microstates simply as “states” for brevity.)

Recall from Table 2.1 and equations (2.82) and (2.83) that the relevant quantity for calculating the individual numbers of accessible states of each system is $\nu N/2$. We borrow the language of γ just before (2.48) to define

$$\gamma_1 \equiv \nu_1 N_1/2, \quad \gamma_2 \equiv \nu_2 N_2/2. \quad (3.69)$$

Now refer to (2.82) or (2.83) [and recall (2.92)], to write the numbers of states accessible to each system as

$$\Omega_1 \propto E_1^{\gamma_1}, \quad \Omega_2 \propto E_2^{\gamma_2}. \quad (3.70)$$

The total number of accessible states (that is, to the entire system) at energy E is the product of the numbers for each system, and can be expressed as a function of E_1 :

$$\Omega(E_1) = \Omega_1 \Omega_2 \propto E_1^{\gamma_1} (E - E_1)^{\gamma_2}, \quad 0 \leq E_1 \leq E. \quad (3.71)$$

Figure 3.9 shows a plot of $\Omega(E_1)$ versus E_1 . This curve turns out to be *extremely* sharply peaked, as we will now show. Its stationary points occur when $\Omega'(E_1) = 0$:

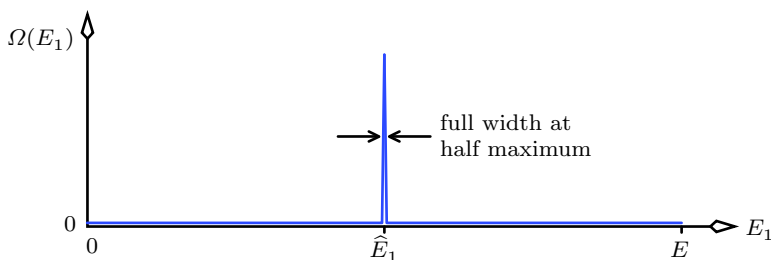


Fig. 3.9 $\Omega(E_1)$ is an *incredibly* sharply peaked function of E_1 . We can quantify this sharpness by calculating its “full width at half maximum”

$$\Omega'(E_1) = E_1^{\gamma_1-1} (E - E_1)^{\gamma_2-1} [\gamma_1 E - (\gamma_1 + \gamma_2) E_1] \stackrel{\text{req}}{=} 0. \quad (3.72)$$

Solving this equation for E_1 (at fixed E) yields three stationary points. Two of these are minima, at $E_1 = 0$ and $E_1 = E$, which are, of course, the extreme values that E_1 can have. The sole maximum occurs at

$$E_1 = \hat{E}_1 \equiv \gamma_1 E / (\gamma_1 + \gamma_2), \quad (3.73)$$

and this peak in the number of accessible states is the really interesting part of the function $\Omega(E_1)$.

How wide is this peak at $E = \hat{E}_1$? A useful measure of the width is a dimensionless number α that gives the relative step away from the peak to where $\Omega(E_1)$ has dropped to half of its peak value:

$$\Omega(\hat{E}_1 + \alpha \hat{E}_1) \equiv \frac{1}{2} \Omega(\hat{E}_1). \quad (3.74)$$

So, $2\alpha\hat{E}_1$ is a measure of the peak’s *full width at half maximum*. (This measure is still only approximate because the peak isn’t necessarily symmetric, but it’s sufficient for our purpose.)

To calculate α , consider first from (3.71), that Ω is a product of powers. In that case, the *logarithm* of Ω will not only be far easier to work with, but, because the logarithm transforms a product into a sum, it gives us the possibility of isolating something new that combines *additively*—meaning it might possibly be envisaged as some kind of substance (perhaps caloric!)—even if only as a mental aid in analyses. We won’t quite reach that picture in what is to come with the definition of *entropy*, because (a) entropy will turn out not to be a substance, and (b) entropy won’t be a conserved quantity. But picturing entropy as a substance that can be transferred, or can even grow, might well allow the gaining of some intuition about it. Indeed, this idea is the modern evolution of the old caloric idea of heat. But aside from this conceptual idea of transforming a product into a sum, mathematically, a logarithm will produce a *far* better approximation of the shape of the peak in Figure 3.9 than if we simply Taylor-expand $\Omega(E_1)$ about the peak’s location

$\hat{E}_1$. This is because the logarithm of a strongly peaked function is not strongly peaked itself, and thus needs fewer Taylor terms to describe it.

And so we work with the logarithm of the number of states accessible to the entire system:

$$\sigma(E_1) \equiv \ln \Omega(E_1). \quad (3.75)$$

Taking the logarithm of (3.71) and differentiating twice yields¹⁴

$$\begin{aligned} \sigma(E_1) &= \text{constant} + \gamma_1 \ln E_1 + \gamma_2 \ln(E - E_1), \\ \sigma'(E_1) &= \frac{\gamma_1}{E_1} - \frac{\gamma_2}{E - E_1}, \\ \sigma''(E_1) &= \frac{-\gamma_1}{E_1^2} - \frac{\gamma_2}{(E - E_1)^2}. \end{aligned} \quad (3.76)$$

Equation (3.74) becomes

$$\sigma(\hat{E}_1 + \alpha \hat{E}_1) = -\ln 2 + \sigma(\hat{E}_1). \quad (3.77)$$

This Taylor-expands to second order in $\alpha \hat{E}_1$, where α is assumed to be small:

$$\cancel{\sigma(\hat{E}_1)} + \underbrace{\sigma'(\hat{E}_1)}_{=0} \alpha \hat{E}_1 + \sigma''(\hat{E}_1) \frac{\alpha^2 \hat{E}_1^2}{2} \simeq -\ln 2 + \cancel{\sigma(\hat{E}_1)}, \quad (3.78)$$

where we note that $\sigma'(\hat{E}_1) = 0$ because $\sigma(E_1)$, like $\Omega(E_1)$, attains its maximum at $\hat{E}_1$. Recalling (3.73), equation (3.76) produces

$$\sigma''(\hat{E}_1) = \frac{-(\gamma_1 + \gamma_2)^3}{E^2 \gamma_1 \gamma_2}. \quad (3.79)$$

Now, some rearranging of (3.78) yields

$$\alpha \simeq \sqrt{\frac{2 \gamma_2 \ln 2}{\gamma_1 (\gamma_1 + \gamma_2)}}. \quad (3.80)$$

For real systems with $\gamma_1 = \gamma_2 \simeq 10^{24}$, we find that $\alpha \simeq 10^{-12}$, and so the full width at half maximum of the peak in Figure 3.9 is around $2 \times 10^{-12} \hat{E}_1$. This is *tiny* compared with $\hat{E}_1$. We conclude that system 1 is extremely likely to have energy $\hat{E}_1$ —and then, of course, system 2 is extremely likely to have the rest of the available energy:

$$\hat{E}_1 = \frac{\gamma_1 E}{\gamma_1 + \gamma_2} = \frac{\nu_1 N_1 E}{\nu_1 N_1 + \nu_2 N_2},$$

¹⁴ Recall Section 1.9.2, which gives the rationale behind what appears to be the logarithm of a dimensional quantity such as energy.

$$\widehat{E}_2 \equiv E - \widehat{E}_1 = \frac{\gamma_2 E}{\gamma_1 + \gamma_2} = \frac{\nu_2 N_2 E}{\nu_1 N_1 + \nu_2 N_2}. \quad (3.81)$$

By what factor does the total number of accessible states Ω drop if E_1 should exceed $\widehat{E}_1$ by, say, one part per million? Call this factor f :

$$f \equiv \frac{\Omega(\widehat{E}_1)}{\Omega([1 + 10^{-6}]\widehat{E}_1)}. \quad (3.82)$$

Calculate f by Taylor-expanding $\ln f$ to second order in $10^{-6}\widehat{E}_1$:

$$\ln f \stackrel{(3.82)}{\simeq} \sigma(\widehat{E}_1) - \sigma(\widehat{E}_1 + 10^{-6}\widehat{E}_1) \simeq -\sigma''(\widehat{E}_1) \frac{10^{-12}\widehat{E}_1^2}{2}. \quad (3.83)$$

Using our known expressions for $\widehat{E}_1$ and $\sigma''(\widehat{E}_1)$ results in $\ln f \simeq 10^{12}$, or, using $e \simeq 10^{0.4343}$,

$$f \simeq e^{10^{12}} \simeq 10^{0.4343 \times 10^{12}} = 10^{434,300,000,000}. \quad (3.84)$$

This is a huge drop. With the combined system equally likely to be in any of its accessible states (by the fundamental postulate of statistics mechanics), the chance of a 1 part per million fluctuation away from energies $\widehat{E}_1$ and $\widehat{E}_2$ is so minuscule that we can discount it from *ever* happening.¹⁵

3.5.2 Temperature and the Equipartition Theorem

Thermal equilibrium is defined as the state in which the internal energy E has become distributed over the two systems as the $\widehat{E}_1$ and $\widehat{E}_2$ of (3.81). When this has occurred, we notice that the average internal energy per particle per quadratic energy term is the same for both systems:

$$\frac{\widehat{E}_1}{\nu_1 N_1} = \frac{\widehat{E}_2}{\nu_2 N_2} = \frac{E}{\nu_1 N_1 + \nu_2 N_2}. \quad (3.85)$$

The two systems in thermal equilibrium are now defined as having a common *temperature* that is proportional to their common average internal energy per particle per quadratic energy term in (3.85). When we derive the ideal-gas law in the next section, we'll show, by attaching the proportionality constant $k/2$ —where $k \simeq 1.381 \times 10^{-23}$ is *Boltzmann's constant*—that this statistical definition of temperature becomes identical to the “everyday”

¹⁵ We might really want to consider a fluctuation of *at least* 1 ppm here. That requires an integral; but the above calculation serves to give a good idea of the numbers involved.

temperature T that was already known in thermodynamics before the advent of statistical mechanics:

$$\frac{kT}{2} \equiv \frac{\hat{E}_1}{\nu_1 N_1} = \frac{\hat{E}_2}{\nu_2 N_2} = \frac{E}{\nu_1 N_1 + \nu_2 N_2}. \quad (3.86)$$

To put it another way, we could define a new quantity called, say, “statistical temperature” to be the expressions in (3.85), and then observe later, from the ideal-gas law, that this statistical temperature is proportional to the everyday thermodynamics temperature T . The constant of proportionality would be called “ $k/2$ ”, where, historically, k has acquired the name “Boltzmann’s constant”.

This definition of temperature, (3.86), implies that the temperature of the systems that we have considered here must always be positive. Its SI unit is the *kelvin*. As with all SI units, this is treated as a noun, and so has a lower-case initial letter in English. In keeping with the SI system, because the kelvin is named after a person (William Thomson of the nineteenth century, also known as Baron Kelvin), its abbreviation K is capitalised. In addition, note that “K” is not a “Kelvin degree”; a temperature of 100 K is “one hundred kelvins”, and not “one hundred degrees Kelvin”, nor “one hundred Kelvin”.

In practice, we must provide something extra to disentangle temperature from Boltzmann’s constant k . This is conventionally accomplished by defining a temperature of $T \equiv 273.16$ K at the *triple point* of water, a unique temperature at which water can exist in solid, liquid, and gaseous forms. Additionally, the modern definition of the Celsius scale then defines one Celsius degree to be one kelvin, along with the triple point at exactly 0.01° on the Celsius scale.¹⁶ The triple point is easier to work with experimentally—and to base a definition on—than the more well-known states of water, ice, and steam, because the triple-point state needs no reference to any standard pressure, and so is more easily reproduced in the laboratory. Basing a practical temperature scale on water’s triple point sets the absolute zero of temperature at $T = -273.15^\circ\text{C}$, the lowest, coldest temperature that a system can ever have.

Note that any attempt to set the melting, triple, and boiling points of water at some pressure to be the temperatures 0°C , 0.01°C , and 100°C respectively, will over-specify the Celsius scale. Hence, we can only say that the melting and boiling points of water at one atmosphere are *approximately* 0°C and 100°C .

Systems 1 and 2 can be intermixed. For example, system 1 might refer to translation ($\nu_1 = 3$) of a set of diatomic molecules, while system 2 refers to the rotation ($\nu_2 = 2$) of the same molecules.¹⁷ Equation (3.86) tells us how

¹⁶ On the use of language, “one Celsius degree” is a kelvin: an increment of one degree on the Celsius scale, at any temperature. Contrast this with “one degree Celsius”, which is a temperature of 274.15 K.

¹⁷ Diatomic molecules don’t rotate about the axis joining the atoms, and so have only two quadratic energy terms, not three.

the internal energy is spread amongst these two systems at equilibrium:

$$\hat{E}_1 = \nu_1 N_1 \frac{kT}{2}, \quad \hat{E}_2 = \nu_2 N_2 \frac{kT}{2}. \quad (3.87)$$

It follows that each of the $\nu_1 N_1$ translational quadratic energy terms can be said to contribute internal energy $kT/2$, and likewise each of the $\nu_2 N_2$ rotational quadratic energy terms can be said to contribute internal energy $kT/2$. This idea that each quadratic energy term, irrespective of its type, contributes $kT/2$ to the system's energy is called the *equipartition theorem*. We'll use it often.

The Equipartition Theorem

If the equilibrium distribution is

- the most probable distribution consistent with fixed total energy and fixed particle number, and
- there is no restriction on the number of particles in any one state, and
- the internal energy varies continuously with a coordinate u and depends on u^2 ,

then the internal energy associated with this coordinate is $kT/2$.

What if the internal energy depends on some other power of u ? We'll tackle that analysis in Section 5.8, where we examine the more general and realistic scenario of a non-isolated system, one that is in thermal contact with a large environment. Because that system's energy is able to vary, we must focus on its *average* value. It will turn out that when the system's energy depends on u^α for some positive α , this energy has an average value of kT/α .

3.6 The Ideal Gas and Temperature Measurement

We are obliged, as early as possible, to make the connection between temperature as defined above in statistical mechanics, and the thermodynamical temperature T that appears in the ideal-gas law, $PV = NkT$. We do that in this section by deriving the statistical-mechanics version of the ideal-gas law from first principles, and then invoking the equipartition theorem to conclude that the statistical-mechanical temperature is identical to the thermodynamical temperature T in “ $PV = NkT$ ”.

The ideal-gas law originated as an experimental observation in thermodynamics. A system's measurable properties, such as its pressure, volume, or length, were observed to vary as that system was heated. Historically, tem-

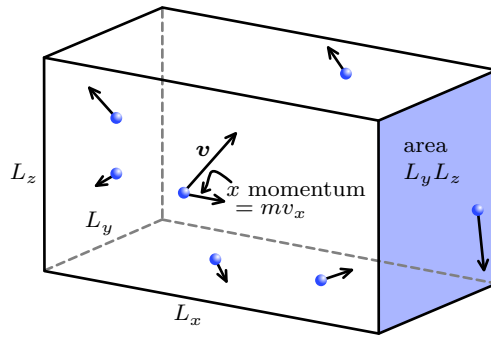


Fig. 3.10 To derive the ideal-gas law, focus on a single representative gas particle in the box and calculate the momentum it transfers to the blue wall in a collision

perature was *quantified* in thermodynamics in such a way as to require these properties to vary as simply as possible with that temperature quantity. Thus, a system's temperature could first be fixed by ensuring that these properties remained constant. Experiments then showed that at fixed temperature, the product PV of a dilute gas was a constant, and this constant increased with temperature. Temperature could then most simply be quantified as some " T " such that $PV \propto T$. It was also expected that a gas's pressure and volume should both be proportional to its number of particles, N ; hence, $PV \propto NT$. This proportionality could then be turned into an equality by defining a constant, which we now call Boltzmann's constant k : thus, $PV = NkT$.

The above is the historical definition of temperature as a numerical quantity T . How do we make a connection to statistical mechanics? Begin with the idea that a gas's pressure P is a measure of how quickly (and with what momentum) each gas particle bounces off the walls of its volume- V container. We will relate PV to the average translational energy of the gas particles.

Figure 3.10 shows an oblong box of sides L_x, L_y, L_z that holds the gas whose N particles, each of mass m , exert a pressure P on the walls through their random motions. Relate this pressure to the momentum changes of all the particles when they strike the blue wall $x = \text{constant}$:

$$P = \frac{\text{force on blue wall due to all particles}}{\text{area of blue wall } (= L_y L_z)}. \quad (3.88)$$

The force on the blue wall due to a single particle is

$$\text{force on blue wall} = \frac{\text{total momentum transferred to blue wall}}{\text{time of transfer}}. \quad (3.89)$$

A particular particle collides with this blue wall with an x velocity of v_x , rebounds elastically, eventually hits the opposite wall a distance L_x away, and then returns and hits the blue wall again. The momentum transferred to the blue wall equals that lost by the particle, which is

$$\begin{aligned}
 -\Delta \text{ momentum} &= \text{initial momentum} - \text{final momentum} \\
 &= mv_x - (-mv_x) = 2mv_x.
 \end{aligned}
 \tag{3.90}$$

The force applied by the particle to the blue wall equals this momentum transfer divided by the *interaction* time, which can be tiny. Consider that the stone that is flicked up from a gravel road to impact your car’s windscreen is an example of a very hard object striking a very hard surface. Since neither stone nor windscreen acts like a tiny trampoline, bending gracefully with the blow, the interaction time is extremely short. The result is a very high “impulsive” force that cracks the window. But we are not really interested in calculating this impact force of each particle on the blue wall—for which we would, at any rate, need to know something of the particle’s internal structure and elasticity. The pressure in the box arises from a large number of particle–wall impacts per second, and we can calculate this pressure equally well by imagining that the effect of each impact is spread over the entire time between successive impacts on the blue wall by the same particle. The time interval between these successive bounces is $2L_x/v_x$, and thus the *average* force on the blue wall due to a single particle is

$$\begin{aligned}
 \left[\begin{array}{l} \text{average force due} \\ \text{to a single particle} \end{array} \right] &= \frac{\text{momentum transferred to blue wall}}{\text{time between successive impacts}} \\
 &= \frac{2mv_x}{2L_x/v_x} = \frac{mv_x^2}{L_x}.
 \end{aligned}
 \tag{3.91}$$

A real particle does not tend to traverse the entire length of the box without colliding with other particles and having its velocity altered drastically. But if particle 1 strikes the blue wall and then strikes it again very soon after as a result of colliding with particle 2 (which ends up being deflected away from the blue wall), then we will consider particle 1 to have taken the place of particle 2 in the wall collision. The situation is then *as if* there had been no collision between the particles: instead, they narrowly missed colliding, but swapped their identities as they passed each other. But the box in Figure 3.10 doesn’t care about their identities; it only registers the pressure. In that sense, the particles can be treated as non-interacting *points*—although they may have rotational and vibrational motion due to collisions with the walls. On average, in the time interval $2L_x/v_x$, we can say that *every* particle has bounced off the blue wall. So, referring to (3.91), we will suppose each particle to contribute an average force on the blue wall of $m\langle v_x^2 \rangle / L_x$. The total force on the blue wall is then this number multiplied by the N particles:

$$\text{force on blue wall due to all particles} = Nm\langle v_x^2 \rangle / L_x.
 \tag{3.92}$$

The pressure is then

$$P \stackrel{(3.88)}{=} \frac{Nm \langle v_x^2 \rangle / L_x}{L_y L_z} = \frac{Nm \langle v_x^2 \rangle}{L_x L_y L_z}. \quad (3.93)$$

But the box's volume is $V = L_x L_y L_z$, so

$$PV = Nm \langle v_x^2 \rangle. \quad (3.94)$$

Now note that $v^2 = v_x^2 + v_y^2 + v_z^2$; and, because no direction is preferred, it follows that $\langle v^2 \rangle = 3 \langle v_x^2 \rangle$. Equation (3.94) then becomes

$$PV = 1/3 Nm \langle v^2 \rangle = 2/3 N \langle mv^2/2 \rangle, \quad (3.95)$$

where $\langle mv^2/2 \rangle$ is the average translational energy of a particle. Next, invoke the equipartition theorem: because translational motion has just three quadratic energy terms, a particle's average translational energy is $3kT/2$. Equation (3.95) then becomes $PV = 2/3 N \times 3kT/2$, or

$PV = NkT.$

(3.96)

This is the celebrated ideal-gas law, upon which our everyday notion of temperature is based. The T in (3.96) came from the equipartition theorem, (3.86) and (3.87), and is thus the temperature defined in (3.86) from the statistical-mechanics idea of microstates. We see now that this definition of temperature is identical to the everyday temperature used in the ideal-gas law from the early days of thermodynamics.

Physicists usually write the ideal-gas law in the form (3.96). Chemists more often deal with large numbers of particles; and for these, the idea of a mole is useful, as discussed around (1.222): a mole of any quantity is Avogadro's number of that quantity's fundamental units, or about 6.022×10^{23} of those units.¹⁸ We define the *gas constant* R :

$$R \equiv N_A k \simeq 8.314 \text{ JK}^{-1} \text{ mol}^{-1}. \quad (3.97)$$

This leads to the *molar form* of the ideal-gas law:

$$PV = NkT = \frac{N}{N_A} N_A kT = nRT, \quad (3.98)$$

where n is the number of moles of the gas present. It's usually convenient in chemical calculations to replace Boltzmann's constant k with R/N_A : the gas constant R is a conveniently simple number, and N_A allows well-known molar quantities to be introduced. Examples of this occur in various places in the chapters ahead.

¹⁸ The first determination of Avogadro's number a century ago involved the work of Johann Josef Loschmidt, leading to his name also sometimes being attached to N_A .

What is the volume of 1 kg of O_2 gas at 1 atmosphere of pressure and 20°C (101,325 Pa and 293.15 K)?

Since one mole of O_2 has a mass of about 32 g, we are dealing with $n = 1000/32$ moles of what is essentially an ideal gas. Then the required volume is

$$V = \frac{nRT}{P} \simeq \frac{1000 \times 8.314 \times 293.15}{32 \times 101,325} \text{ m}^3 = 0.752 \text{ m}^3 = 752 \ell, \quad (3.99)$$

where “ ℓ ” denotes a litre.

When the particles of an ideal gas each have ν quadratic energy terms, the equipartition theorem states that the gas’s internal energy is $kT/2$ per quadratic energy term per particle, giving the gas a total energy of

$$E = \nu NkT/2. \quad (3.100)$$

If the gas has a background potential energy per particle of U_0 —taking as an example Earth’s atmosphere—then we must include this in the total energy E :

$$E = \nu NkT/2 + NU_0. \quad (3.101)$$

In Section 4.3.3, we’ll set $U_0 = mgz$ at height z in our atmosphere, where m is the mass of a typical air molecule.

One quick check is worth making: if we combine two volumes (not necessarily equal) of ideal gases composed of the same substance and with the same temperature T , the temperature of the mixture should also be T , since equal temperatures means no heat flow. What does the equipartition theorem say here? The two volumes have N_1 and N_2 particles, and the particles each have ν quadratic energy terms and can be immersed in a background potential energy of U_0 per particle. Equation (3.101) says that the two volumes have total energies of

$$E_1 = \nu N_1 kT/2 + N_1 U_0, \quad E_2 = \nu N_2 kT/2 + N_2 U_0. \quad (3.102)$$

It follows from (3.102) that

$$E_1 + E_2 = \nu(N_1 + N_2)kT/2 + (N_1 + N_2)U_0. \quad (3.103)$$

If we write the mixture’s total energy and particle number as $E' \equiv E_1 + E_2$ and $N' \equiv N_1 + N_2$, then it follows from (3.103) that

$$E' = \nu N' kT/2 + N' U_0. \quad (3.104)$$

But this is just (3.101) again. Mixing the gases is completely consistent with the equipartition theorem.¹⁹

3.6.1 Measuring Temperature: the Constant-Volume Gas Thermometer

The *constant-volume gas thermometer* is a classic device that uses the ideal-gas law to measure temperature. Consider that the ideal-gas law (3.96) says that the pressure of an ideal gas is proportional to its temperature at fixed volume and particle number. This implies that at fixed volume, the pressure and temperature P, T of an ideal gas are related to a possibly different pressure and temperature P_0, T_0 of an ideal gas via

$$\frac{P}{P_0} = \frac{T}{T_0}. \quad (3.105)$$

Now examine the apparatus shown in Figure 3.11. Suppose we enclose some gas in a chamber that is connected to a reservoir of liquid: mercury is conventionally used, because its high density keeps the size of the apparatus small. We calibrate the thermometer by setting the “hot substance” in the picture to be, say, water–ice at its triple point, whose temperature is defined as $T_0 = 273.16$ K. As the gas comes to thermal equilibrium with the water–ice, we continually adjust the height of the mercury reservoir so as to keep the level of mercury in the thin vertical tube at a fixed pre-selected level, which can be chosen to be anywhere along the tube. When no further adjustment needs to be made, we measure the height $h = h_0$ of mercury that the gas supports due to its having temperature $T_0 = 273.16$ K. The pressure P_0 inside the gas chamber exerts a force $P_0 A$ on the area- A interface with the supported mercury column. Equating this force with the column’s weight yields

$$P_0 A = \text{mass of column} \times g = \varrho h_0 A g, \quad (3.106)$$

where ϱ is the density of the mercury and g is the acceleration due to gravity.

Recording the height h_0 has calibrated the thermometer. We now replace the water–ice with the substance whose temperature T is to be measured, and re-adjust the height of the mercury reservoir to keep the mercury in the thin vertical tube at its pre-selected level. This ensures the gas volume is fixed—which means we don’t have to measure that volume, which removes

¹⁹ This is a point worth emphasising: temperature is a measure of *average* energy (energy per quadratic energy term per particle), not *total* energy. Mixing two cups of water that each have temperature 50°C will, of course, give a mixture at 50°C . This is no different from saying that when two classes of students mix, with each class having the same mean exam score of some m per student, then the mean exam score of the entire group will also be m .

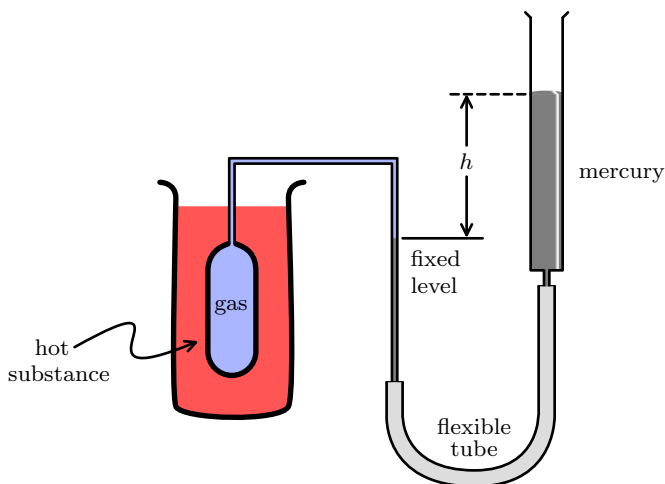


Fig. 3.11 The constant-volume gas thermometer measures the temperature of a hot substance by comparing the pressure of the gas inside a chamber in thermal equilibrium with the substance with the pressure that results when the hot substance is replaced by calibrating material. The pressures are found by measuring the height h of a heavy liquid such as mercury

one possible source of error from the apparatus. The new pressure P due to the hot substance is related to the new height h by

$$PA = \rho hAg. \quad (3.107)$$

Equations (3.106) and (3.107) produce

$$\frac{P}{P_0} = \frac{h}{h_0}. \quad (3.108)$$

If the gas in the bulb is ideal, (3.105) can now be written as

$$\frac{h}{h_0} = \frac{T}{T_0}. \quad (3.109)$$

The only unknown in (3.109) is T , which can now be calculated.

Real gases are not quite ideal, because their particles *do* experience long-range attractions to each other. (This is discussed further in Sections 3.7 and 4.2.2.) When we use the constant-volume gas thermometer filled with different gases at different pressures, we find a small variation of typically less than a percent in the measured values of T of a test substance. This variation across different gases decreases when the amount of gas (and thus both P_0 and P) is reduced at the given fixed volume—provided the temperatures T_0 and T do not also drop to very low values. We infer that a real gas becomes increasingly ideal as its pressure drops at a fixed volume.

3.6.2 Temperature of Our Upper Atmosphere

The above definition of temperature was built on a real gas with no background potential energy, and having a large number of particles. When a background potential energy exists (such as the gravity acting throughout our atmosphere), we cannot say that the entire body of gas has a single temperature, even though it's in thermal equilibrium. Indeed, the temperature of our atmosphere drops with increasing altitude, and yet no heat flow exists in the ideal case of a weatherless atmosphere.

The number of particles in a system of interest might not be large. A case in point is our upper atmosphere: hundreds of kilometres above Earth's surface, atmospheric temperatures can reach 1000°C or more; and yet you will freeze in such an environment. The air temperature is high because the gas molecules gain much kinetic energy by absorbing solar radiation: we use the equipartition theorem to equate this energy per particle to $\frac{1}{2}\nu kT$ (where ν depends on the type of particle), with the result being a very high value of T . If we are in the upper atmosphere, any one of these fast-moving particles can bump into us and transfer perhaps most of its energy to us; but there are simply too few of these collisions happening per second to give us enough energy to balance what we lose by radiating energy in the “blackbody” manner discussed in Chapter 9. In contrast, although the air around us at sea level is much cooler, and so transfers less energy per collision, many more collisions occur per second; and we are thus able to absorb enough kinetic energy to balance roughly what we radiate.

For an extreme example, picture a gas of just half a dozen particles moving in random directions at such high speeds that their temperature is in the millions of kelvins.²⁰ Some vestige of thermodynamics still applies, but little should be read into the high temperature here: the gas has only a tiny amount of energy available to heat anything else up.

This idea of an ultra-low-density atmosphere recalls the discussion of free expansion in Section 3.4. Recall Figure 3.4, and suppose that the empty volume that the gas expands into is arbitrarily large—even the size of the universe. We then have a huge volume that is, to all intents and purposes, a vacuum; and yet the gas it holds has the same temperature that it had originally, before it expanded freely. This scenario is completely valid, because a gas's temperature is a measure of its total kinetic energy, and this energy does not change during a free expansion, because the gas does no work as it expands. But if you were to place a sealed smaller box of cooler gas into this

²⁰ Temperature is proportional to the particles' kinetic energy. We might say that temperature is a measure of particle speeds—but *that* dependence is not linear. Even though the speeds are bounded by the speed of light c , there is no upper limit to the particles' kinetic energies. [Remember that $\frac{1}{2}mv^2$ is a non-relativistic approximation of a particle's kinetic energy, while the true (relativistic) expression is $(\gamma - 1)mc^2$, where $\gamma \equiv 1/\sqrt{1 - v^2/c^2}$, and m is the particle's rest mass.] It follows that temperature has no upper limit.

larger volume, the two temperatures would take a very long time to equalise, because the rate of collisions of the outside particles with the box would be extremely low. That is, the overall system would have a very long relaxation time, which reduces the effectiveness of thermodynamics in discussing it.

Returning to the upper atmosphere: even when the density of particles is high, the idea of temperature can still be mis-applied if the particles are not moving randomly. A box of cold gas that moves past us at ultra-high speed is not defined as having a high temperature, because the kinetic energy of its centre-of-mass motion must not be included when applying the equipartition theorem. Rather, its temperature is a measure of the particle speeds in its centre-of-mass frame. That means all observers (even those moving past the box at high speed) agree on the value of that temperature—because each of them performs essentially the same calculation, transforming the measured particle velocities into the box’s unique centre-of-mass frame.²¹ It is meaningless, for example, to assign a temperature T to a projectile or a set of particles in a particle accelerator, by setting the individual kinetic energies equal to $kT/2$ and solving for T .

3.7 The Non-Ideal Gas and van der Waals’ Equation

A gas’s departure from being ideal can be described with a framework that enables a better equation of state than $PV = NkT$ to be formulated, at least in principle. This framework is the *virial theorem*, which finds widespread application in physics.

Begin with the idea that the long-time mean of the time-derivative of a bounded function is zero. This stands to reason: it says that when a train moves in some arbitrary way on a track of finite length, its velocity will average out to zero over long times; after all, if that wasn’t the case, then the train would drift forever in one direction—which it cannot do because the track has finite length. We can also prove this theorem using calculus. Call the function $f(t)$, and find the mean of its time-derivative over a time T , where $T \rightarrow \infty$. Denote this time-mean by “ $\langle \cdot \rangle$ ”:

$$\lim_{T \rightarrow \infty} \langle f'(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f'(t) dt = \lim_{T \rightarrow \infty} \frac{1}{T} [f(T) - f(0)] = 0. \quad (\text{QED})$$

(3.110)

²¹ When observers in all frames agree on the value of a numerical quantity, it is called a *scalar*; temperature is an example of a scalar. Similarly, when observers in all frames agree on the value of a quantity that includes a sense of direction, it is called a *vector*; velocity is an example of a vector. A “scalar” is often treated, incorrectly, as a synonym for a “number”, but the fact is that not all numbers are scalars. Only numbers that all frames agree on are scalars. In particular, components of vectors are numbers but not scalars, because different frames gives those components different values.

Now consider a particle at some point $\mathbf{r}$ in a container, with velocity $\mathbf{v}$ and acceleration $\mathbf{a}$. Apply the above idea to the case of $f(t) = \mathbf{r} \cdot \mathbf{v}$, and understand all angle brackets below to apply in the long-time limit:

$$0 = \left\langle \frac{d}{dt}(\mathbf{r} \cdot \mathbf{v}) \right\rangle = \langle v^2 + \mathbf{r} \cdot \mathbf{a} \rangle. \quad (3.111)$$

Multiply this by the particle's mass m , and recognise that the force on it is $\mathbf{F} = m\mathbf{a}$, and that it has kinetic energy $E_k = 1/2 mv^2$:

$$0 = \langle mv^2 + \mathbf{r} \cdot \mathbf{F} \rangle = 2\langle E_k \rangle + \langle \mathbf{r} \cdot \mathbf{F} \rangle. \quad (3.112)$$

This is the virial theorem: $2\langle E_k \rangle + \langle \mathbf{r} \cdot \mathbf{F} \rangle = 0$. (In fact, we can formulate any number of analogous expressions by starting with other choices for $f(t)$, such as $f(t) = \mathbf{r} \cdot \mathbf{a}$.) For a system of N particles with total kinetic energy E_k , apply the virial theorem to each particle individually and sum the result:

$$0 = \sum_i 2\langle E_{ki} \rangle + \langle \mathbf{r}_i \cdot \mathbf{F}_i \rangle = 2\langle E_k \rangle + \left\langle \sum_i \mathbf{r}_i \cdot \mathbf{F}_i \right\rangle. \quad (3.113)$$

This last expression is the stepping-off point for the application of the virial theorem to diverse areas of physics. For our case of a real gas, we suppose that the force $\mathbf{F}_i$ on particle i is the sum of the forces due to every other particle, along with the force due to the wall:

$$\mathbf{F}_i = \sum_j \mathbf{F}_{ij} + \mathbf{F}_{i,\text{wall}}, \quad (3.114)$$

where

$$\begin{aligned} \mathbf{F}_{ij} &\equiv \text{force on particle } i \text{ due to particle } j \text{ (taking } \mathbf{F}_{ii} \equiv 0), \\ \mathbf{F}_{i,\text{wall}} &\equiv \text{force on particle } i \text{ due to wall.} \end{aligned} \quad (3.115)$$

The virial theorem (3.113) now says

$$\begin{aligned} -2\langle E_k \rangle &= \left\langle \sum_i \mathbf{r}_i \cdot \mathbf{F}_i \right\rangle = \left\langle \sum_i \mathbf{r}_i \cdot \left[\sum_j \mathbf{F}_{ij} + \mathbf{F}_{i,\text{wall}} \right] \right\rangle \\ &= \left\langle \sum_{ij} \mathbf{r}_i \cdot \mathbf{F}_{ij} \right\rangle + \left\langle \sum_i \mathbf{r}_i \cdot \mathbf{F}_{i,\text{wall}} \right\rangle. \end{aligned} \quad (3.116)$$

Ideal gases have no inter-particle interactions, and so for them, (3.116) becomes (with subscript “IG” denoting an ideal gas)

$$-2\langle E_k \rangle_{\text{IG}} = \left\langle \sum_i \mathbf{r}_i \cdot \mathbf{F}_{i,\text{wall}} \right\rangle. \quad (3.117)$$

This enables (3.116) to be written as

$$-2\langle E_k \rangle = \left\langle \sum_{ij} \mathbf{r}_i \cdot \mathbf{F}_{ij} \right\rangle - 2\langle E_k \rangle_{\text{IG}}. \quad (3.118)$$

But the equipartition theorem says that $\langle E_k \rangle = 3/2 NkT$, and we also know from (3.95) that $\langle E_k \rangle_{\text{IG}} = 3/2 PV$. Substitute these into (3.118):

$$-2 \times 3/2 NkT = \left\langle \sum_{ij} \mathbf{r}_i \cdot \mathbf{F}_{ij} \right\rangle - 2 \times 3/2 PV. \quad (3.119)$$

This rearranges to

$$PV = NkT + \frac{1}{3} \left\langle \sum_{ij} \mathbf{r}_i \cdot \mathbf{F}_{ij} \right\rangle. \quad (3.120)$$

This, then, is a fairly raw extension of the ideal-gas law to real gases. Calculating the long-term average on the right-hand side of (3.120) requires knowledge of each $\mathbf{r}_i$, which, in turn, needs a model of the interaction potential between the particles. These calculations are specialised, but some generic statements can be made here. First, writing the particle density as $\nu \equiv N/V$, the relevant calculations lead to a modification of the ideal-gas law “ $P = \nu kT$ ” as

$$P = kT[\nu + B_2(T)\nu^2 + B_3(T)\nu^3 + \dots] \quad (3.121)$$

for some functions $B_2, B_3, \dots$ of temperature. This equation is not meant to be obvious, and is usually called the *virial expansion*.

Here is a sketch of one way to understand something qualitative of the most important modification to the ideal-gas equation, which was produced by Johannes van der Waals in the latter part of the nineteenth century. Consider modifying the ideal-gas law $PV = NkT$ to incorporate the behaviour of real particles in the following way. If the real particles’ actual pressure and actual volume are P and V , respectively, start from a template that tries to model the real gas as an ideal gas:

$$\text{modified } P \times \text{modified } V = NkT. \quad (3.122)$$

- **Modifying pressure P :** The particles in a real gas experience a long-range attraction to each other, and so a particle that is about to collide with a container wall will be slightly pulled back by the attraction of the other particles—assuming it isn’t also attracted to the wall, which is an assumption that van der Waals made. The resulting pressure P is thus lower than the ideal-gas law predicts, so we expect to have to add a positive term to P to mimic the ideal-gas pressure. The higher the gas density N/V , the higher the attractive force pulling a particle back into the gas. It turns out that we adjust the lower pressure P back to the ideal-gas

prediction by adding to P a term that is proportional to $(N/V)^2$: so, the modified pressure in (3.122) becomes $P + aN^2/V^2$ for some positive a .

- **Modifying volume V :** Real gases do not have point particles; each particle occupies some volume b of its own. But the volume appearing in the ideal-gas law is the space between the particles; and so to “simulate” that volume, we subtract from the container volume V the volume of the particles themselves: for N particles, we subtract a volume bN . The modified volume in (3.122) thus becomes $V - bN$.

Placing these modifications into (3.122) gives van der Waals' modification of the ideal-gas law, now known as *van der Waals' equation*:

$$\left(P + \frac{aN^2}{V^2}\right)(V - bN) = NkT, \quad (3.123)$$

for positive constants a and b that depend on the gas in question. Approximate measured values of a and b for common gases are:

$$a \approx 1.5 \times 10^{-48} \text{ Pa m}^6, \quad b \approx 8 \times 10^{-29} \text{ m}^3. \quad (3.124)$$

In particular, the value of b corresponds to a sphere of radius 6 nm.

Van der Waals' equation can easily be rearranged into one instance of the virial expansion (3.121). Start by dividing (3.123) by V :

$$(P + a\nu^2)(1 - b\nu) = \nu kT. \quad (3.125)$$

Solve this for P by invoking a geometric series in the first line below, which we are able to do because, typically, $b\nu \approx 0.002 \ll 1$:

$$\begin{aligned} P &= \frac{\nu kT}{1 - b\nu} - a\nu^2 = \nu kT(1 + b\nu + b^2\nu^2 + \dots) - a\nu^2 \\ &= kT[\nu + b\nu^2 + b^2\nu^3 + \dots - a\nu^2/(kT)] \\ &= kT\left[\nu + \left(b - \frac{a}{kT}\right)\nu^2 + b^2\nu^3 + \dots\right]. \end{aligned} \quad (3.126)$$

This last series indeed matches the virial expansion (3.121).

Like the ideal-gas law, van der Waals' equation (3.123) can be treated as a function $P(V)$ for given values of T . Plots of this function of very low values of V for a selection of temperatures are shown in Figure 3.12. At high temperature, the van der Waals gas is approximately ideal, and the plot of pressure versus volume approximates the ideal-gas hyperbola $P = NkT/V$. The van der Waals plot has a vertical asymptote at $V = bN$, which can be seen by rearranging (3.123) to write P as a function of V . The asymptote represents the situation when all N molecules have been squeezed together with no extra room available to move. For a mole of gas, $bN \approx 0.05$ litres,

$$\begin{aligned} V_c &\simeq 2.1 \times 10^{-10} \text{ m}^3 = 0.2 \text{ mm}^3, \\ T_c &\simeq 403 \text{ K} \simeq 129^\circ\text{C}. \end{aligned} \quad (3.129)$$

Below the critical temperature T_c , the plot of $P(V)$ develops a dip that, for a low-enough temperature T_0 , just touches $P = 0$. We can find T_0 by setting $P = 0$ in (3.123) and rewriting the result as a quadratic in V :

$$kTV^2 - aNV + abN^2 = 0. \quad (3.130)$$

Now, demanding that this quadratic has a single root (or, what might be called a repeated root) leads to

$$T_0 = \frac{a}{4kb}. \quad (3.131)$$

At this temperature, the pressure reaches zero at a volume of $V = 2bN$. For the a and b in (3.124), $T_0 \simeq 340 \text{ K}$, or about 67°C . At colder temperatures, (3.130) has two roots, the larger of which is denoted V_1 in Figure 3.12:

$$V_1 = \frac{N}{2kT} \left[a + \sqrt{a^2 - 4abkT} \right]. \quad (3.132)$$

In the region of these lower temperatures and volumes, the gas is beginning to condense into a liquid.

As a last comment, note that (3.123) can be written in a molar form as

$$\left(P + \frac{a_{\text{mol}} n^2}{V^2} \right) (V - b_{\text{mol}} n) = nRT \quad (3.133)$$

for n moles of gas, where R is the gas constant as usual, and where a_{mol} and b_{mol} are given by

$$a_{\text{mol}} = N_A^2 a, \quad b_{\text{mol}} = N_A b. \quad (3.134)$$

These molar constants a_{mol} and b_{mol} are what you will usually find tabulated for various gases—but they tend to be written as “ a ” and “ b ” in those tables. The various van der Waals expressions in this section can be converted into molar form by making the replacements

$$a \rightarrow a_{\text{mol}}, \quad b \rightarrow b_{\text{mol}}, \quad N \rightarrow n, \quad k \rightarrow R. \quad (3.135)$$

3.8 Entropy and the Second Law of Thermodynamics

Let’s return to the ideal gas at temperature T with a large number N of particles. It follows from (2.82) and (2.92) that its number of accessible states is $\Omega \simeq f(V, N) E^{\nu N/2}$ for some function f . If we understand a large- N ap-

proximation to apply in the following equalities, then on recalling (3.75)’s “ $\sigma \equiv \ln \Omega$ ”, we conclude that the logarithm of the number of accessible states of the gas is

$$\sigma(E, V, N) = \ln \Omega = \ln f(V, N) + \frac{\nu N}{2} \ln E, \quad (3.136)$$

where the discussion in Section 1.9.2 allows us to write “ $\ln E$ ” even though E is not dimensionless. Now calculate

$$\left(\frac{\partial \sigma}{\partial E} \right)_{V, N} = \frac{\nu N}{2E} \frac{\text{equipartition}}{\text{theorem}} \frac{\nu N}{2 \nu N k T / 2} = \frac{1}{kT}. \quad (3.137)$$

This logarithm $\sigma \equiv \ln \Omega$ is so useful that we include a factor of Boltzmann’s constant for later utility, and define the *entropy* of a system to be $S \equiv k\sigma$:

entropy $S \equiv k \ln \Omega$.

(3.138)

Observe that a trivial rearrangement of (3.138) produces

$$\Omega = e^{S/k}. \quad (3.139)$$

We’ll use this in Section 5.1 when developing one of the core equations of statistical mechanics, the Boltzmann distribution. Equation (3.138) converts (3.137) to

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

(3.140)

Either of these expressions now *defines* the temperature T of an arbitrary system (not just an ideal gas), in terms of how that system’s energy and entropy increase in step with each other at fixed volume and particle number. Equation (3.140) says that for any system with a well-defined temperature T ,

$dE = T dS \quad \text{for fixed } V, N.$

(3.141)

This equation has exactly the same content as (3.140), but it lends itself easily to analyses using infinitesimals, as we’ll see soon.

Entropy is *additive for two systems that are placed in contact*. This is because the number of microstates available to the combined system (before an interaction occurs) is the product of the individual numbers of microstates. Systems 1 and 2 have entropies $S_1 = k \ln \Omega_1$ and $S_2 = k \ln \Omega_2$, and so the entropy of the combined system is

$$S = k \ln(\Omega_1 \Omega_2) = k \ln \Omega_1 + k \ln \Omega_2 = S_1 + S_2. \quad (3.142)$$

Now allow the two subsystems to interact. Once the whole has attained equilibrium, it is overwhelmingly likely to be found in one of the microstates in the set that maximises Ω . Hence, its total entropy is maximised, and we arrive at the *Second Law of Thermodynamics*:

The Second Law of Thermodynamics

When two systems interact, the odds are overwhelmingly likely that the entropy of the combined system will increase along the path to equilibrium. It attains its maximum value when equilibrium is reached.

Thermodynamical statements of this law exist that describe what an engine and/or a refrigerator working in a cycle can and cannot accomplish. We'll encounter these briefly in Section 4.2.1. They predate the above statement of the Second Law, and can be shown to be equivalent to it in an extensive study of thermodynamics.

Remember that the Second Law is not really a “law”: there is no invisible push that forces a system's entropy to increase over time. Entropy increase is simply overwhelmingly favoured by the probabilities of the various microstates accessible to the system. It's safe to say that no one has ever seen an isolated system's entropy decrease, even though it *could* decrease—and, in fact, “must” decrease if we wait long enough, just as the ink drop in the bath tub of Section 1.1 “must” eventually take on all appearances, ranging from Tutankhamen's iconic image to the face of every creature that has ever lived or will ever live. We would have to wait for much longer than the lifetime of the universe to be likely to have seen even a minuscule decrease in an isolated system's entropy. The Second Law is called a law because of this practical certainty in its operation, even though, strictly speaking, it does not *have* to apply. And yet we are so confident in its holding true that it is often viewed as being more of a law than the other laws of physics.

The Use of Ω_{tot} Instead of Ω in the Expression for Entropy

In Section 2.5, we saw that when analysing large systems whose energy E is effectively continuous, the density of states $g(E)$ is more easily defined than $\Omega(E)$. Given that $\Omega(E) \simeq g(E) \Delta E$, it's clear that discussing such a system requires an energy spacing ΔE ; and without this, calculating the entropy $S = k \ln \Omega$ becomes problematic.

The traditional fix to this difficulty is to apply the approximation of (2.92): $\Omega(E) \approx \Omega_{\text{tot}}(E) \approx g(E)$. In practice then, the entropy of a large, continuous system tends to be written as

$$S \simeq k \ln \Omega_{\text{tot}} . \quad (3.143)$$

This is very convenient, because it was Ω_{tot} that we found more natural to compute than Ω for the various systems in Chapter 2. We'll apply this approximation to calculating the entropy of an ideal gas next.

3.8.1 Entropy of an Ideal Gas of Point Particles

Two important expressions for entropy are those for ideal gases of distinguishable point particles and identical-classical point particles. To derive them, begin with (2.50) for distinguishable point particles and (2.84) for identical-classical point particles. We write strict equalities here, as is customary:

$$\begin{aligned}\Omega_{\text{tot}}^{\text{dist}} &= V^N e^{3N/2} \left(\frac{2\pi mkT}{h^2} \right)^{3N/2}, \\ \Omega_{\text{tot}}^{\text{ic}} &= \left(\frac{V}{N} \right)^N e^{5N/2} \left(\frac{2\pi mkT}{h^2} \right)^{3N/2}.\end{aligned}\tag{3.144}$$

The entropy of each gas is given by the usual $S = k \ln \Omega \simeq k \ln \Omega_{\text{tot}}$, so write

$$S^{\text{dist}} = Nk \left[\ln V + \frac{3}{2} + \frac{3}{2} \ln \frac{2\pi mkT}{h^2} \right],\tag{3.145}$$

$$S^{\text{ic}} = Nk \left[\ln \frac{V}{N} + \frac{5}{2} + \frac{3}{2} \ln \frac{2\pi mkT}{h^2} \right].\tag{3.146}$$

The expression for the entropy of an ideal gas of identical-classical point particles in (3.146) is known as the *Sackur–Tetrode equation*. Otto Sackur and Hugo Tetrode independently discovered this expression in 1911—which was a remarkable feat in a time when Planck's constant was still rather new and quantum physics lay a decade in the future. We will use the Sackur–Tetrode equation in Sections 4.3.3 and 7.4.

Finally, recall from the de Broglie discussion in Section 2.2, that the expressions for entropy in (3.145) and (3.146) cannot be expected to hold in the low-temperature limit. We'll return to this point in Section 7.5.

Here is a small check on the logical consistency of (3.145) and (3.146). Write both of those equations generically as

$$S = f(V, N) + \frac{3Nk}{2} \ln T\tag{3.147}$$

for some function f that we needn't focus on [and it is not the same f that appeared in (3.136)]. It follows that at fixed V and N ,

$$\frac{dS}{dT} = \frac{3Nk}{2T},\tag{3.148}$$

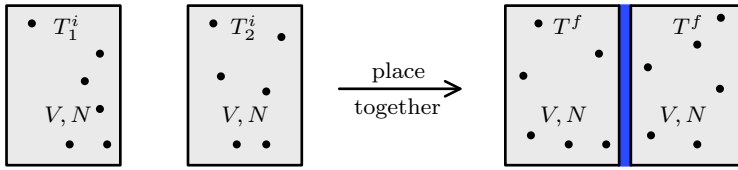


Fig. 3.13 Two boxes of ideal gas are placed in thermal contact (shown as the blue region), without being able to exchange volume or particles. What final temperature results, and what happens to the various entropies?

in which case

$$TdS = \frac{3Nk}{2} dT. \quad (3.149)$$

But equipartition says that the energy of the ideal gas of point particles is $E = 3NkT/2$. So, when N is fixed, $dE = 3Nk dT/2$. Now invoke (3.149) to conclude that

$$dE = TdS \quad \text{for fixed } V, N, \quad (3.150)$$

which we saw earlier in (3.141).

3.8.2 The Canonical Example of Entropy Growth

The generic equation (3.147) for the entropy of an ideal gas can be applied to the canonical set-up in Figure 3.13 to investigate the way in which entropy grows when two systems are placed in thermal contact.

The figure shows “box 1” and “box 2”, each of which holds an ideal gas of point particles. For simplicity, let them contain equal numbers of particles N and equal volumes V , but they have different initial temperatures T_1^i and T_2^i with $T_1^i < T_2^i$, with superscripts i and f in this discussion denoting initial and final states. Place the boxes together so that “heat can flow between them” (box 1 heats up and box 2 cools), but not volume or particles. This process needn’t be quasi-static, because the following equations use only initial and final states, and the final state isn’t affected by how quickly equilibrium was reached. What is the final common temperature T^f , how do the entropies of the boxes (or rather, the gases they hold) change, and how does the total entropy change?

Equipartition tells us that boxes 1 and 2 initially have respective energies

$$E_1^i = 3NkT_1^i/2, \quad E_2^i = 3NkT_2^i/2. \quad (3.151)$$

Their total energy must be the sum of these, since no energy enters or leaves the system:

$$\text{total energy} = E_1^i + E_2^i = 3Nk(T_1^i + T_2^i)/2. \quad (3.152)$$

Equipartition relates this energy to the final common temperature in the usual way as follows, where the final configuration contains $2N$ particles that each have 3 quadratic energy terms:

$$\text{total energy} = 3 \times 2N \frac{kT^f}{2} \stackrel{(3.152)}{=} \frac{3Nk(T_1^i + T_2^i)}{2}. \quad (3.153)$$

We infer that the final common temperature is $T^f = (T_1^i + T_2^i)/2$, the average of the two initial temperatures. Now, how does the entropy of each box evolve as they come to thermal equilibrium? The growth in entropy of box 1 is

$$\begin{aligned} \Delta S_1 &= S_1^f - S_1^i \stackrel{(3.147)}{=} f(V, N) + \frac{3Nk}{2} \ln T^f - f(V, N) - \frac{3Nk}{2} \ln T_1^i \\ &= \frac{3Nk}{2} \ln \frac{T^f}{T_1^i} > 0. \end{aligned} \quad (3.154)$$

Since $\Delta S_1 > 0$, the entropy of box 1 increased as it heated up to the final common temperature. Similarly, the growth in entropy of box 2 is

$$\Delta S_2 = \frac{3Nk}{2} \ln \frac{T^f}{T_2^i} < 0. \quad (3.155)$$

That is, the entropy of box 2 decreased as it cooled to the final common temperature. What about the entropy S of the *combined* system—how did that change? Recall from (3.142) that the entropies of subsystems add to give that of the combined system. The entropy increase of the combined system is then

$$\begin{aligned} \Delta S &= S^f - S^i = S_1^f + S_2^f - (S_1^i + S_2^i) = \Delta S_1 + \Delta S_2 \\ &= \frac{3Nk}{2} \left[\ln \frac{T^f}{T_1^i} + \ln \frac{T^f}{T_2^i} \right] = \frac{3Nk}{2} \ln \frac{(T_1^i + T_2^i)^2}{4T_1^i T_2^i} \\ &= \frac{3Nk}{2} \ln \left[\frac{1}{4} \left(\frac{T_1^i}{T_2^i} + 2 + \frac{T_2^i}{T_1^i} \right) \right]. \end{aligned} \quad (3.156)$$

Although we originally stipulated that $T_1^i < T_2^i$, this was for illustration only; that inequality was used above only in the “ > 0 ” and “ < 0 ” parts of (3.154) and (3.155). If the temperatures T_1^i, T_2^i were equal (so that nothing happened when the boxes were brought into thermal contact), (3.156) shows—not surprisingly—that the total entropy increase ΔS would be zero. If the temperatures were different, the total entropy increase would be positive; and if the temperatures were *very* different, this positive increase would be huge.

The result is, that although the entropy of the warming gas in box 1 increases while that of the cooling gas in box 2 decreases, the entropy of the entire system always increases—in line with what we found when defining

temperature in Section 3.5. This is the content of the Second Law: while entropy certainly can decrease in a *subsystem* (box 2 here), the total entropy of all interacting systems will be found either to remain constant or to increase throughout the interaction. It will become constant when equilibrium is reached.

The above example makes it clear that “ $dE = TdS$ ” cannot apply to the entire system, because $dE = 0$ always, but $dS > 0$ for unequal temperatures. This is not a matter of simply saying “No single temperature T is defined for the systems as they come to thermal equilibrium”, because no matter what representative temperature we might insert into “ $dE = TdS$ ”, we will still obtain the contradiction “ $0 = \text{a positive number}$ ” (unless we set that representative temperature to zero; but doing that has no meaning here).

We see now that “heat flow” differs from the old idea that heat, or caloric, was a conserved substance that flowed between bodies. *Energy* flows between the bodies according to $dE_1 = -dE_2$. If we slow the heat flow down to make the process quasi-static, then, at any moment, both gases have well-defined intermediate temperatures T_1, T_2 . Then, because $dE_1 = -dE_2$, it follows that $T_1 dS_1 = -T_2 dS_2$, and we do have a quantity TdS that “flows”—but that quantity is energy. In contrast, when $T_1^i < T_2^i$, and so $T_1 < T_2$ as the gases come to a common temperature, consider that $dE_1 = -dE_2$ implies

$$\frac{dE_1}{T_1} > \frac{-dE_2}{T_2}. \quad (3.157)$$

In other words,

$$dS_1 > -dS_2, \quad (3.158)$$

meaning that box 1 gains more entropy than box 2 loses. As the gases come to thermal equilibrium, entropy thus appears from nowhere, and that means it cannot be the old caloric. The “real” caloric is energy $dE = TdS$, which *does* flow and *is* conserved throughout the thermalisation. And because we already have this concept of energy, we can discard the old idea of a caloric.

No longer do we view heat as a conserved substance that can flow between bodies—or even a substance that can be created by doing work, such as when we rub our hands together on a cold morning, or when we bore out a cannon, as studied by Count Rumford. Our modern idea is that the heat we feel when holding a hot cup of tea results from the simple mechanical billiard-ball collisions that transfer energy and momentum from the random motions of the atoms comprising the “hot” tea and the ceramic cup to our hands; nerves in our skin are then triggered to send electrical signals to our brain, which interprets those signals as “this cup is hot”. There is no substance called “heat” that actually flows; there is only an initially non-uniform distribution of internal energy of atomic motion. When this distribution gradually evens out as the energy transfers back and forth between atoms, the result is a system whose internal energy of each particle in each of its quadratic energy terms has settled on a single value of $kT/2$. The changing distribution of

temperature is governed by the *diffusion equation* of Section 4.1.4, and even though that equation describes diffusion more generally, we now know that no new substance called “heat” is actually diffusing when “heat flows”.

Should the noun “heat” be discarded then, and the word used only as a verb or adjective? Should we speak only of “heating a cup” or holding a “hot cup”, but never say “I can feel the heat of the cup flowing into my cold hands”? Certainly we should be aware, when discussing thermodynamics, that we are not discussing a substance called “heat”. But the aim of physics is to give us insight into the workings of Nature, and not to hobble our language with obsessive linguistic detail. It would be tiresome to say “I can feel a transference of internal energy from the cup into my hands”—or perhaps we might allow ourselves to say “I can feel a transference of internal energy from the *hot* cup into my *cold* hands”. Once we know what “heat” is, using the word as a noun becomes purely a linguistic convenience to avoid comically tedious phrases such as “the transference of internal energy, where this energy might be transferred ‘thermally’, meaning not by way of doing work or transferring particles”.

The use of “heat” as a noun parallels our language of waves. We speak of waves on the ocean, and yet we know that the water doesn’t contain a separate substance called a wave that can be gathered up and put into a box. Water waves are an expression of the kinetic energy of water molecules that are doing nothing more than moving in small ellipses, and these individual elliptical motions manifest as waves travelling across/through/in the water. But we plainly have no difficulty picturing a wave as a separate entity with an existence of its own, and the use of “wave” as a noun is, of course, perfectly routine in physics. Likewise, there is no reason to avoid or ban the use of the noun “heat”.

3.8.3 Reversible and Cyclic Processes

A special type of quasi-static process that has played a major role in historical discussions of thermodynamics is the *reversible* process, being one whose operation can be reversed—at least in principle—when one or more of its state variables is altered infinitesimally. The classic example is a cylinder with a frictionless piston that confines a gas that is hotter than the outside air by some infinitesimal amount $\varepsilon > 0$, shown at the left in Figure 3.14. As the confined cylinder gas expands, it does work on the outside air by pushing the piston in the direction out of the cylinder. We can run the process in reverse with a single infinitesimal change: by cooling the cylinder gas infinitesimally relative to the surrounding atmosphere, as shown at the right in the figure. The atmosphere will now do work by pushing the piston into the cylinder and compressing the cylinder gas. The motion of the piston has been reversed by our making an infinitesimal change to the system parameters.



Fig. 3.14 A reversible process. **Left:** When the cylinder gas is infinitesimally warmer than the outside atmosphere (by $\epsilon > 0$), it expands against the piston (assumed frictionless) to do work quasi-statically on the atmosphere. **Right:** When the cylinder gas is made infinitesimally cooler than the atmosphere, the atmosphere pushes against the piston to do quasi-static work on the cylinder gas

This reversible motion introduces the *cyclic* thermodynamical process, in which a system of interest evolves in a cycle that restores it to its initial state, while other systems with which it interacts (such as an environment) are *not* restored to their initial states at the end of the cycle. An example is a car’s piston engine, where the pistons, valves, and crankshaft periodically return to the start of the fuel-intake cycle, while the atmosphere continually has hot gas exhausted into it.

Naturally, we want an engine’s components to be as frictionless as possible as they slide past each other during the combustion process. Any process involving friction is irreversible, because friction only ever works in one direction, dissipating energy thermally: we cannot perturb any state variable infinitesimally to “undo” the friction in a reverse process. The total entropy increase always equals the sum of the individual entropy increases of all subsystems, as in (3.156); but we are generally interested only in a few of those subsystems, whether or not friction is involved. So, we might write “ $\Delta S = \Delta S_1 + \Delta S_2 + \Delta S_3$ ” when three subsystems are involved; but perhaps subsystem 3 is an environment that is not being *explicitly* included in the analysis, because it is just unnecessary to deal with. We know only that subsystem 3’s entropy never decreases throughout the cycle, and so we write “ $\Delta S \geq \Delta S_1 + \Delta S_2$ ” instead. You will find such inequalities in discussions of cyclic processes in books on thermodynamics.

Friction is more difficult to treat in thermodynamics than in classical mechanics. When we apply Newton’s laws to a mechanical system, the presence of friction can be treated as just another force—albeit one whose precise form is not always well known. But the thermodynamical nature of friction and irreversible processes in general have proven historically to be difficult to pin down. The question of how quasi-static processes relate to reversible processes has a long history, one which has been muddled by concepts that are not always precisely defined when they are discussed, such as whether a process is assumed to happen cyclically, and what the precise connection is between a system and an environment that drives that system’s evolution. Different researchers have preferred different approaches. Gibbs and Carathéodory, both major players in the field a century ago, worked with *equilibrium states*, where

quasi-stasis is central. Others, such as Planck, made *thermodynamical processes* central to their arguments, where reversibility comes to the fore.

One difficulty in discussions of the subject's history has been the confusion arising from what has turned out to be an indiscriminate use of the word “reversible” in the original English translation of Planck's important work on the subject, *Vorlesungen über Thermodynamik*. An added problem is that some authors use the word “reversible” as if it means quasi-static. We will leave in-depth discussion of the word “reversible” to others who focus on thermodynamics and its history, particularly in the introduction of the concept of entropy via cyclic thermodynamical processes. These are thermodynamical ideas that lie, to some extent, outside the realm of statistical mechanics. The real demand that we make in the pages ahead is that any process being analysed must either be quasi-static, or else the creation of its end state must be *achievable* with a quasi-static process. Such processes can always be described by a set of state variables, meaning that our processes will always have, for example, a well-defined temperature. A case in point is the free expansion in Figure 3.4: although this is not a quasi-static process, we will analyse it in Section 3.11 by drawing a parallel between it and the quasi-static process of allowing a gas of well-defined volume to push with a well-defined pressure against a piston, while that gas is heated to maintain it at the constant temperature that exists in free expansion. Further discussion of reversibility appears in Section 4.2.1.

3.8.4 The Use of Planck's Constant for Quantifying Entropy

In Sections 2.3 and 2.4, we discussed the use of Planck's constant h to define the cell size when partitioning phase space. We had set a cell's volume using a factor of h for each pair of position–momentum variables; but, in fact, when calculating an entropy increase, we could just as well have replaced h with $h/10$ or $100h$.

To see why, recall that (2.24) expresses the general idea that the number Ω of accessible microstates of a system is proportional to the volume $\mathcal{V}$ of phase space consistent with that system's energy: $\Omega = \alpha \mathcal{V}$ for some α . Now consider the system evolving from a microstate of the same energy as the others in a volume $\mathcal{V}_i$, to a microstate belonging to a volume $\mathcal{V}_f$. As the system evolves from $\mathcal{V}_i$ to $\mathcal{V}_f$, its entropy increases by

$$\Delta S = S_f - S_i = k \ln \Omega_f - k \ln \Omega_i = k \ln \frac{\Omega_f}{\Omega_i} = k \ln \frac{\alpha \mathcal{V}_f}{\alpha \mathcal{V}_i} = k \ln \frac{\mathcal{V}_f}{\mathcal{V}_i}. \quad (3.159)$$

Clearly, our choice of α is immaterial to the resulting entropy increase: ΔS is determined only by a ratio of phase-space volumes. Settling on a specific

value of α can be viewed as a device that defines discrete microstates of a continuous system. This discretising of phase space is a modern way of approaching entropy. An alternative approach might define entropy via phase-space volume alone, without introducing any tiling of that space. But that would divorce entropy from the idea of counting accessible microstates; in contrast, defining entropy via the number of accessible microstates allows us to build an intuition about it, because we can then begin a study of the subject by counting the microstates of very simple discrete systems, as we did in Chapter 2.

3.9 Can Temperature Be Negative?

We began our discussion of temperature with (3.70)'s expressions of $\Omega \propto E^{\nu N/2}$ for each of two interacting gases. Temperature was introduced first in (3.86) for these gases, and then more generally in (3.140) or, equivalently, (3.141). This idea of temperature is practical, something that gives a meaningful and mathematically useful description of the physical world of thermodynamics. Its definition springs from the core idea that adding energy to a gas increases its number of available microstates. Thus, its energy E and entropy S either increase or decrease together. It follows from (3.140) or (3.141) that T is positive and well behaved in this everyday example of heat flow: two gases of unequal temperature that are placed in thermal contact both evolve toward a common equilibrium temperature that is somewhere between the two initial temperatures. The cooler gas gains energy and entropy, while the warmer gas loses energy and entropy.

What if E and S can be arranged so as no longer to increase or decrease in tandem? If a system's energy *decreases* while its entropy increases, its temperature $T = (\partial E / \partial S)_{V,N}$ will be negative. Although it's possible to produce this behaviour in some discrete systems, what results is thermodynamically unrealistic. We will investigate the classic example here.

This example of E decreasing while S increases employs a set of charged particles, each of which has quantised spin, and hence a quantised magnetic moment $\boldsymbol{\mu}$. For simplicity, we'll make these particles distinguishable by, say, anchoring each one at a fixed point on a line. The interaction energy E of this magnetic moment with an external magnetic field $\mathbf{B}$ was given in (3.60) by $E = -\boldsymbol{\mu} \cdot \mathbf{B}$ (where we have replaced the U of that equation with E). Define the z direction to be that of the field, making $\mathbf{B} = B\mathbf{u}_z$, where $\mathbf{u}_z$ is the z unit vector. Then $E = -\boldsymbol{\mu} \cdot B\mathbf{u}_z = -\mu_z B$.

Suppose the particles have only two allowed spins, with each z component of magnitude m :

- “spin up” has $\mu_z = m$, and so has magnetic energy $E = -\mu_z B = -mB$;
- “spin down” has $\mu_z = -m$, and so has magnetic energy $E = -\mu_z B = mB$.

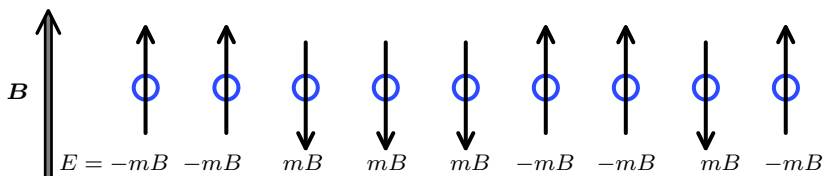


Fig. 3.15 A set of spins, some of which are “up” (their component along the $\mathbf{B}$ field points in the same direction as that field), and the rest “down”. (It’s wise to remember that, when drawn as vectors, spins and their associated magnetic moments don’t really point exactly up or down; drawing the arrows as up/down here is conventional, and merely indicates the sign of the z component of each spin, where the z axis here points in the direction of $\mathbf{B}$. The quantum theory of angular momentum shows that the “spin- $1/2$ ” vector has magnitude $\sqrt{1/2 \times 3/2} \hbar = \sqrt{3}/2 \hbar$ and z component $\pm \hbar/2$. It follows that the vector then actually tilts away from the vertical by the rather large angle of $\cos^{-1} 1/\sqrt{3} \simeq 55^\circ$)

Figure 3.15 shows an example set of spins, some up and some down. We calculate the energy E and entropy S of each spin configuration and then plot E versus S . This plot will not be continuous, and thus has no slope defined, and so finding the temperature of any configuration via “ $T = dE/dS$ ” is not really possible. Instead, we write “ $T \approx \Delta E/\Delta S$ ” and see what results.

A set of N particles, u of which have spin up and $N - u$ spin down, has a total magnetic energy of

$$E = u \times -mB + (N - u) \times mB = (N - 2u)mB. \quad (3.160)$$

The entropy of this configuration is

$$S = k \ln \Omega = k \ln \frac{N!}{u!(N - u)!}. \quad (3.161)$$

For example, only one configuration has $u = 0$ (all spins down), and likewise only one configuration has $u = N$ (all spins up); so, $S = 0$ for both of these cases. But a mixture of spins allows for many configurations, with a maximum number occurring when the mixture is split evenly between up and down. Equations (3.160) and (3.161) parametrise E and S in terms of u , the number of particles with spin up. This allows us to plot energy versus entropy in Figure 3.16. Beginning with the configuration of least energy (all spins up: $u = 100$), adding energy allows particles to flip their spin to down, which initially means more states become available as the number u of spin-up particles decreases. If we hold to the definition of temperature from (3.140) as $T \equiv dE/dS$ here, then it’s clear that the discontinuous “curve” of points in Figure 3.16 doesn’t really have a slope dE/dS ; but, of course, it has an approximate slope, and hence the system’s temperature is at least approximately defined as $T \approx \Delta E/\Delta S$. This temperature is then clearly positive when the total energy E is negative, and it is, in some sense, infinite at

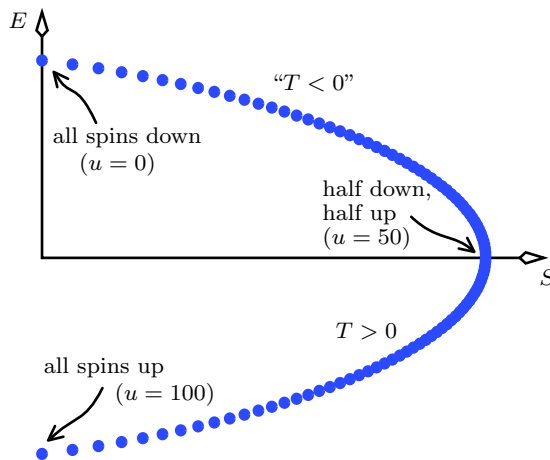


Fig. 3.16 Energy versus entropy for a set of 100 spin-1/2 particles, found from (3.160) and (3.161) for u running from 0 to 100. Each dot represents one value of u , for u running from 0 to 100. If the particles' temperature is defined as the slope of the E -versus- S plot in the way of $T = (\partial E / \partial S)_{V, N}$ in (3.140) (or at least $T \approx \Delta E / \Delta S$ in this discontinuous case), then T is positive only when a majority of spins are up. T is infinite for a half/half mixture of up and down spins, and it is negative when a majority of spins are down

$E = 0$, when the particles are split evenly between up and down. As more energy is added, more particles flip to spin-down, and the approximate slope of the graph becomes *negative*. Does that mean the system now really has a negative temperature?

It seems that a “system 1” with negative temperature (the particles in Figure 3.16, with total energy E positive) would be hotter than a “system 2” with any positive temperature, because if system 1 transferred some of its energy to system 2 (until system 1's spins became half up/half down in the example above), the entropy of both systems would increase—and hence the combined entropy would increase, satisfying the Second Law. In other words, no matter how hot system 2 was, it would become even hotter were it to interact with a set of particles that were mostly spin down. Really?

Thermodynamics and temperature were not designed for such simplistic scenarios. Remember that system 1's particles will usually—perhaps always—have translational quadratic energy terms. When system 1 has positive energy and it gives up some of this energy, its *spin* entropy will increase; but, at the same time, the presence of its translational modes will tend to act to *decrease* its *total* entropy. Thus its energy and entropy decrease together, and so its temperature will be positive, not negative. But even if system 1 has *no* translational quadratic energy terms, it can only transfer energy to system 2 in some way dictated by the way in which its spins interact with system 2; and again, this transfer procedure will involve other energy terms,

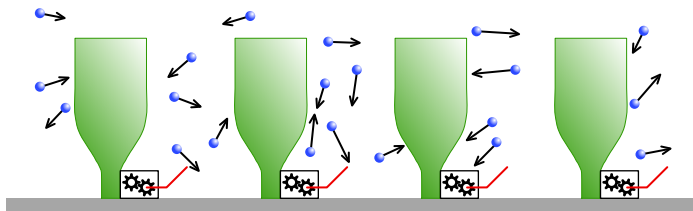


Fig. 3.17 An array of bottles balanced on their noses, surrounded by a gas. Collisions of the gas particles will tend to topple the bottles and add energy to the gas

because systems tend not to be constructed solely of spins. The upshot is that applying the idea of temperature to this highly contrived set of spins is not a useful thing to do. It only produces the absurdity that a box of particles that we can hold in our hand is capable of making a star heat up by a tiny amount.

In fact, we can reduce the above discussion of spin to a very basic level that underlines the essential speciousness of the idea of negative temperature. Consider an array of bottles, each of which is balanced upside down, as shown in Figure 3.17. Being balanced upside down corresponds to spin pointing down in Figure 3.15. There is only one way in which the bottles can stand like this: one microstate, which means they have zero entropy in this configuration (i.e., their entropy is $S = k \ln 1 = 0$). Each bottle is connected by gears to a paddle with a dampening mechanism. The bottles are immersed in a gas, and collisions of the gas particles will tend to topple the bottles. When a bottle topples, the paddle transfers the energy of its toppling to the gas, and the bottle eventually comes to rest. Collisions with the gas particles will topple the bottles, increasing their entropy because the toppled system has many microstates: there are many ways in which the bottles can partly or all be lying down. But, in falling over, the bottles lose gravitational potential energy that is transferred to the gas particles. Here is a situation in which the bottles' energy decreases while their entropy increases. Thus, $\partial E / \partial S < 0$ for the bottles; but to infer from this that the bottles have a negative temperature seems rather perverse, because the concept of temperature was never designed to describe an array of bottles.

In Section 5.1, we'll see that when a system interacts thermally with an environment, the population of each of the system's states with energy E is proportional to $e^{-E/(kT)}$; hence, states with higher energy have exponentially smaller populations. If we could arrange for a "population inversion", in which higher-energy states had *higher* populations, that distribution might be describable using $e^{-E/(kT)}$, where this exponential factor is now given a negative temperature to force it to *increase* with energy. But temperature was defined to describe an equilibrium, and a population inversion is not an equilibrium. This idea of negative temperature is sometimes incorrectly invoked in the context of lasers: although these do involve a population in-

version (see Section 9.9), that inversion only applies to perhaps one or two energy levels, whereas the term $e^{-E/(kT)}$ should apply to *all* energy levels.

A “Negative Length” in Orbital Mechanics

Although negative temperature is irrelevant to the real physical systems for which the subject of thermodynamics was designed, we can certainly find other quantities that at first look nonphysical, but which turn out to be *mathematically* useful. One of these is the “negative semi-major axis length” in orbital mechanics.

Traditionally, the subject of orbital mechanics embarks by calculating the orbit of a planet that is bound gravitationally to its parent star: such a bound orbit is always an ellipse (or possibly a circle, which is just a special case of an ellipse). One of the parameters that characterises this orbit is the ellipse’s semi-major axis length a . The orbits of gravitationally *unbound* objects, such as some fast comets, are open—hyperbolic in shape—and, of course, these don’t have any sort of elliptical axis length associated with them. But the equations describing hyperbolic and elliptical orbits have similarities, and sometimes they differ only in that the hyperbolic case has a negative number in a spot where the analogous equation for the elliptical case has the length a . If we define a kind of “semi-major axis length” for a hyperbolic orbit and allow it to be negative, these equations for hyperbolic orbits will then take on the same form as those for elliptical orbits. This is precisely what is done; yet no one makes the mistake of thinking that we are literally using negative axes lengths. The negative value of a is just a convenience that allows two equations (for hyperbolic and elliptical orbits) to be written as one. It has no physical content; it is purely for mathematical convenience.

3.10 Intensive and Extensive Variables, and the First Law for Quasi-Static Processes

Temperature, pressure, and chemical potential are known as *intensive* variables: they are defined at each point of the system, and at equilibrium, they are all constant throughout the system. They don’t scale with the system: placing two identically prepared gases in contact does not double their common temperature, pressure, or chemical potential. Intensive variables are not defined for a system far from equilibrium.

In contrast, volume and particle number are known as *extensive* variables: they depend on the size of the system, and they scale *proportionally* to that size. That is, placing two identical samples of gas in contact produces a sys-

tem with twice the volume and particle number of each sample. Extensive variables are easy to define for a system far from equilibrium.

Experimentally, these two types of variable—intensive and extensive—appear to be sufficient to quantify all systems encountered in statistical mechanics; for instance, no variables are known that scale as the *square* of the system’s size. This means that an extensive variable can always be converted to a new intensive one: we simply divide the extensive variable by another extensive variable, so that the system’s size cancels out in the division. To demonstrate, consider van der Waals’ equation (3.123). If we define an *intensive* volume variable $v \equiv V/N$, where N is the number of gas particles, the equation becomes

$$\left(P + \frac{a}{v^2}\right)(v - b) = kT. \quad (3.162)$$

This is a slightly simpler form that relates only the intensive variables P, v, T . Indeed, the discussion immediately following (3.123) uses v rather than V (specifically, the $\nu \equiv N/V$ in that discussion equals $1/v$ here). Even so, recognising the existence of *extensive* variables sheds light on the First Law of Thermodynamics. When this law is written with exact differentials in (3.68), intensive and extensive variables form *conjugate pairs* in the terms $-PdV$ and μdN , which have the form

$$\text{intensive} \times d(\text{extensive}). \quad (3.163)$$

Equation (3.68) expresses the First Law as a sum of the mysterious heat transfer $\mathring{d}Q$ and two terms that each have the form of (3.163). These conjugate-pair terms encapsulate mechanical and diffusive interactions. We hope to find a third pair of conjugate variables that will represent thermal interactions, and which will replace the elusive $\mathring{d}Q$ in the First Law with a term of the form “intensive $\times d(\text{extensive})$ ”.

In Section 3.8 we found that *entropy* is extensive: the total entropy of a set of interacting systems is always the sum of the systems’ individual entropies, even though this total entropy grows as the systems evolve toward equilibrium. Indeed, at constant volume and particle number (i.e., for thermal interactions only), the First Law says $dE = \mathring{d}Q$, whereas (3.141) says $dE = TdS$. We infer that TdS is the desired replacement for $\mathring{d}Q$ that makes for an exact-differential-only quasi-static version of the First Law. Replacing the “heat into the system”, $\mathring{d}Q$, with TdS is something that we already saw and used in the discussion around Figure 3.13. We might replace dE with $\mathring{d}Q$ in that discussion and observe that, whereas we certainly can write $\mathring{d}Q_1 = T_1 dS_1$ and $\mathring{d}Q_2 = T_2 dS_2$, we cannot write “ $\mathring{d}Q = TdS$ ” for the entire evolving system. So, the quasi-static version of the First Law using only exact differentials applies individually to each subsystem, but not to the combined system, because it is only the subsystems that are always held very close to equilibrium.

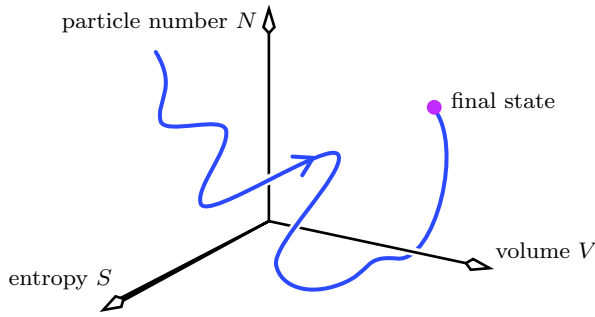


Fig. 3.18 The three extensive variables, one per axis, that describe the quasi-static evolution of a system toward its final state

The First Law of Thermodynamics for Typical Quasi-Static Processes

Representing the mechanical work done on the system by a pressure–volume term alone, the increase in a system’s energy during a quasi-static process is

$$dE = TdS - PdV + \mu dN. \quad (3.164)$$

This restricted version of the First Law depicts a system’s energy as a function of three extensive variables: S takes care of thermal interactions, V is representative of all mechanical interactions, and N allows for diffusive interactions. A system’s evolution is shown in Figure 3.18.

Equation (3.164) is a version of the First Law of Thermodynamics that requires the process being analysed to be quasi-static. Restricting it to such processes renders it *analytical*: open to analysis, in the sense that its differentials have become exact, and so are amenable to the methods of calculus. It is often called the “fundamental equation of thermodynamics”. We will use it extensively in the pages to come.

3.11 A Non-Quasi-static Process

Return now to the free expansion shown in Figure 3.4. On the left in that figure, we start with an ideal gas confined to one part of a thermally insulating box. We remove the partition very quickly, allowing the gas to expand at its own rate to occupy the whole box. The speeds of the particles don’t change in this process; hence, the gas’s temperature remains constant, and thus so does its energy $\nu NkT/2$. Its volume increases, of course. Its pressure decreases, because the particles now bounce less frequently from the walls

that enclose the larger volume. The particles' speeds are well defined during the free expansion, and so the temperature of the gas is also well defined throughout. But its volume and pressure are *not* well defined during the expansion. We ask the question: by how much does this free expansion increase the gas's entropy S ?

In this simple case, we can calculate ΔS from (3.138), where we will approximate the number of accessible states Ω at the gas's energy E by the *total* number of states Ω_{tot} for all energies up to E , as discussed in Sections 2.5 and 3.8. Thus,

$$\begin{aligned}\Delta S &= S_f - S_i = k \ln \Omega_f - k \ln \Omega_i = k \ln \frac{\Omega_f}{\Omega_i} \\ &\simeq k \ln \frac{\Omega_{\text{tot}}^f}{\Omega_{\text{tot}}^i} \stackrel{(2.45)}{=} k \ln \left[\frac{V_f^N (2\pi m E)^{3N/2}}{h^{3N} (3N/2)!} \times \frac{h^{3N} (3N/2)!}{V_i^N (2\pi m E)^{3N/2}} \right] \\ &= Nk \ln \frac{V_f}{V_i},\end{aligned}\tag{3.165}$$

where V_i, V_f are the initial and final volumes of the gas, respectively.

But suppose that no knowledge of Ω or Ω_{tot} is available. We cannot calculate ΔS from (3.164), because that equation requires the system to be very close to equilibrium at all times: that is, all parameters in (3.164) must always be well defined. Nevertheless, S is a function of state, and so ΔS depends only on the initial and final states of the gas. In that case, ΔS will equal the entropy increase of the same type of gas as it undergoes *any other* process having the same initial and final states as those of our freely expanding gas. The most obvious and easiest choice here is to consider the gas to be separated by a piston from an atmosphere that initially is at the same temperature and pressure as the gas. We heat the gas enclosed by the box and piston slowly, keeping it at constant temperature by allowing it to do work against the piston as it expands. Figure 3.19 shows this process on a PV *diagram*, on which a path represents the gas's evolving state. This controlled expansion is quasi-static, and so all parameters in (3.164) are now always well defined. It follows that

$$\Delta S \stackrel{(3.164)}{=} \int dS = \int \frac{dE + P dV - \mu dN}{T}.\tag{3.166}$$

The temperature T is constant, and hence so is the energy: $dE = 0$. The number of gas particles is fixed, and so $dN = 0$. Hence, using $P = NkT/V$ for the ideal gas,

$$\Delta S = \frac{1}{T} \int P dV = \frac{1}{T} \int \frac{NkT}{V} dV = Nk \ln \frac{V_f}{V_i}.\tag{3.167}$$

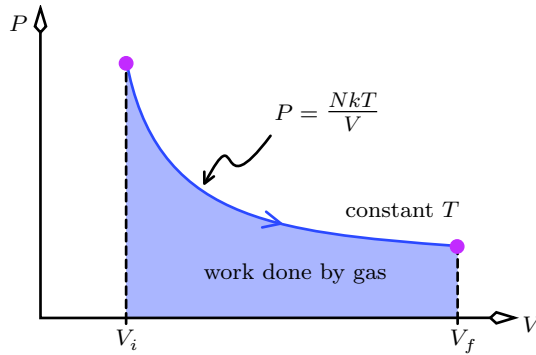


Fig. 3.19 The isothermal expansion of an ideal gas that does work against a piston quasi-statically. The path on the diagram represents the gas's evolution. Recall from the comment just after (3.52) that the gas does work $\int P dV$, which is the area under the path if the evolution heads in the direction of increasing V

This matches (3.165). It demonstrates that although we cannot analyse a non-quasi-static process mathematically using the standard tools above, we can sometimes find another process that has the essential physical features of the non-quasi-static process, and yet *is* quasi-static, and so *can* be analysed mathematically.

3.12 The Ideal-Gas Law from Entropy

Section 3.6 gave us the ideal-gas law from a first-principles study of how the particles in a gas generate pressure by colliding with the walls of their container. This same law also emerges from the quasi-static version of the First Law of Thermodynamics, (3.164). To see how, rearrange the First Law to obtain

$$TdS = dE + PdV - \mu dN. \quad (3.168)$$

Recalling (1.167) and the discussion leading up to it, extract from (3.168) the following partial derivative:

$$P = T \left(\frac{\partial S}{\partial V} \right)_{E,N}. \quad (3.169)$$

Let's use this to calculate P for an ideal gas of, firstly, distinguishable particles. In Section 7.5, we'll explore what distinguishability means in practice, and will show that, in particular, air molecules at room temperature and pressure can indeed be treated as distinguishable: in a sense, it really is as if the molecules were individually numbered. So, recall (2.82) by writing the number of states of such a gas as

$$\Omega \approx \Omega_{\text{tot}} = V^N f(E, N), \quad (3.170)$$

for some function f . This gas's entropy is thus

$$S = k \ln \Omega = Nk \ln V + k \ln f(E, N). \quad (3.171)$$

[This expression is, of course, a generic version of (3.145), since, for an ideal gas, T converts easily to E via the equipartition theorem.] Substitute this entropy into (3.169) and differentiate it at fixed E and N . The number of particles N is fixed, but what about energy E ? The internal energy E of an *ideal* gas resides solely in the motion of its particles; by definition, these particles have no potential energy of separation from each other. If we enlarge the container that contains the gas, the particles simply move around as before—they just collide with the walls less often—so their internal energy E remains the same (but the pressure they exert on the container walls drops due to the less-frequent collisions). Hence E is independent of V for an ideal gas, and (3.169) then yields

$$P = T \frac{\partial}{\partial V} [Nk \ln V + k \ln f(E, N)] = T \frac{Nk}{V}. \quad (3.172)$$

This rearranges to give us the ideal-gas law, $PV = NkT$.

The above procedure barely changes for an ideal gas of identical-classical particles. Here we must divide Ω by $N!$, but doing so only changes the function $f(E, N)$ in the above expression for entropy, $S = Nk \ln V + k \ln f(E, N)$. Thus nothing really changes in the partial derivative (3.172), and the ideal-gas law emerges once more.

The simple procedure of taking a partial derivative that gave us pressure in the last few equations works more generally for any conjugate pair of variables. The quasi-static First Law is written generically as

$$dE = T dS + \sum_n I_n dX_n, \quad (3.173)$$

where I_n and X_n are the n^{th} intensive and extensive conjugate-paired variables, respectively. The above procedure rewrites the First Law as

$$T dS = dE - \sum_n I_n dX_n. \quad (3.174)$$

From this, we see immediately that

$$I_n = -T \left(\frac{\partial S}{\partial X_n} \right)_{E \text{ and all other variables}}. \quad (3.175)$$

For the case of, say, $I_1 = -P$ and $X_1 = V$, (3.175) reproduces (3.169). We'll encounter (3.175) again in Section 5.9.

3.13 Relation of Entropy Increase to Interaction Direction

The Second Law of Thermodynamics is renowned for being the only statement found in physics that provides an approximate *arrow of time*, telling us the direction in which a system will almost certainly evolve. The laws of mechanics are time reversible, meaning we cannot tell whether a movie of colliding billiard balls is being run forward or backward: both motions are mechanically valid. In contrast, when we compare a movie of an egg breaking as it drops onto a floor with that same movie run in reverse, we instinctively know which version shows time running forward. The reversed scenario of the egg fragments re-assembling and rocketing upward is mechanically valid, although we might ask, “But what gives the newly assembled egg the impetus to climb up against gravity?”. The movie probably doesn’t show the vibrations induced in the floor when the egg smashes onto it, with these vibrations quickly running outward and being absorbed by the environment. When running the movie in reverse, these vibrations suddenly appear from outside and converge on the egg. We interpret them as emerging simultaneously from the environment by sheer chance, then running inward to interfere constructively at the precise location of the egg fragments, forcing them to assemble into a complete egg and flicking it into the air. While physically possible, this scenario is so improbable that we instinctively rule it out from ever happening. Instead, the movie version that shows the egg smashing is also showing the egg evolving from occupying one of a small number of microstates in which it is whole, to occupying one of the vastly greater number of microstates in which it is broken. We instinctively interpret this direction of increasing entropy as the forward flow of time.

If we wait long enough, will we ever see a broken egg magically re-assemble in our kitchen? The involvement of probability in “waiting long enough” is subtle.

Probability and “Waiting Long Enough”

Here is a scenario of applying for employment. If you and nine other people all apply for each job advertised, and the selection of a candidate is random (and thus unaffected by past applications), then the chance that you will get the next job you apply for is always $1/10$, no matter how long you have been applying for jobs. Does that mean it’s not worthwhile applying for the next job? Consider that the chance of *not* getting the first job you apply for is 0.9 . The chance of not getting either of the first *two* jobs you apply for is $0.9^2 = 0.81$. It follows that if you are prepared to apply for two jobs, your overall chance of finding employment is $1 - 0.81 = 0.19$. If you are prepared to apply for three jobs, your overall chance of finding employment is $1 - 0.9^3 \simeq 0.27$. And if you are prepared

to apply for twenty jobs, your overall chance of finding employment rises to $1 - 0.9^{20} \simeq 0.88$. If you miss out on the first 19 jobs then you have simply been unlucky; but the past is the past, and your chance of landing the 20th job is still only 1/10. Nevertheless, it's clear that you *will* eventually find a job, even though your chance of finding a job never increases with time.

The question of observing entropy growth is similar to the above example of applying for jobs. The chance that a broken egg magically re-assembles is vanishingly small, and it remains vanishingly small for as long as we wait. But if we are prepared to wait long enough, then we can expect “sometime” to see the egg re-assemble. This process will, of course, take a “similar” length of time as seeing dispersed ink re-assemble into an ink drop in a bath tub. But it's crucial to remember that this decrease in the universe's entropy has *nothing* to do with time running backward. The re-assembly of the egg is still happening with time running forward. It is true that a backward flow of time can be arranged in the special theory of relativity,²² but that subject has nothing to do with the growth of entropy. Just what time actually *is* remains mysterious, and we cannot define “the future” as the situation in which entropy increases. Nonetheless, given two film clips of the same scenario, where one film is the time reverse of the other, we can be as good as certain that the clip on which the universe's entropy increased shows the scenario that really happened.

Let's use the First and Second Laws to check this idea of “entropy growth almost certainly indicates time running forward” against our knowledge of the directions in which various quantities in an interaction will flow. Allow systems 1 and 2 to interact thermally, mechanically, and diffusively, and calculate the immediate increase dS in the entropy of the combined system when systems 1 and 2 are first connected and begin to interact:

$$dS = dS_1 + dS_2, \quad (3.176)$$

where, as usual, dS_1 is the entropy increase of system 1 in this interaction, and similarly for all other variables below. We will express dS solely in terms of dS_1 , dV_1 , dN_1 , and so must eliminate all system 2 infinitesimals in the expressions to follow. Energy is conserved throughout (so the energy dE_2 gained by system 2 equals the energy $-dE_1$ lost by system 1), and suppose that volume and particle number are conserved ($dV_2 = -dV_1$ and $dN_2 = -dN_1$). Begin by writing dS_2 in terms of dS_1 via the First Law:

²² A *uniformly accelerated observer* (one who feels a constant acceleration—which does *not* mean that his acceleration is constant in any inertial frame) will say that time “below his horizon” is running backward. But he can never receive signals from the events that lie below his horizon, and so he can never *watch* any scenario running backward in time.

$$dE_2 = T_2 dS_2 - P_2 dV_2 + \mu_2 dN_2. \quad (3.177)$$

This rearranges to

$$\begin{aligned} T_2 dS_2 &= dE_2 + P_2 dV_2 - \mu_2 dN_2 \\ &= -dE_1 - P_2 dV_1 + \mu_2 dN_1 \quad (\text{and now apply the First Law again}) \\ &= -(T_1 dS_1 - P_1 dV_1 + \mu_1 dN_1) - P_2 dV_1 + \mu_2 dN_1 \\ &= -T_1 dS_1 + (P_1 - P_2) dV_1 + (\mu_2 - \mu_1) dN_1. \end{aligned} \quad (3.178)$$

Incorporate this into (3.176) to arrive at the desired expression for the overall increase in entropy:

$$\begin{aligned} dS &= dS_1 + dS_2 = dS_1 + \frac{1}{T_2} \times (3.178)\text{'s last right-hand expression} \\ &= dS_1 + \frac{-T_1 dS_1 + (P_1 - P_2) dV_1 + (\mu_2 - \mu_1) dN_1}{T_2} \\ &= \frac{T_2 - T_1}{T_2} dS_1 + \frac{P_1 - P_2}{T_2} dV_1 + \frac{\mu_2 - \mu_1}{T_2} dN_1. \end{aligned} \quad (3.179)$$

As the system heads toward equilibrium, $dS > 0$. We have the freedom to control how much of each interaction occurs on the right-hand side of (3.179). It follows that we require the entropy to increase for each of those interactions treated separately:

– **Thermal interaction:** Equation (3.179) reduces to

$$\frac{T_2 - T_1}{T_2} dS_1 > 0. \quad (3.180)$$

If $T_1 < T_2$, then $dS_1 > 0$ (notice that we rule out any idea that temperature can be negative), and so heat flows toward system 1. If $T_1 > T_2$, then $dS_1 < 0$, and so heat flows toward system 2. We see that heat always flows toward the lower temperature, as we expect. Note also that (3.178) says

$$T_2 dS_2 = -T_1 dS_1 + \text{terms involving } dV_1 \text{ and } dN_1. \quad (3.181)$$

So, the heat $T_2 dS_2$ flowing into system 2 will equal the heat $-T_1 dS_1$ flowing out of system 1 *only* when the volume and particle flows are zero; that is, only when the interaction is purely thermal (remembering that “volume” here is really the prototypical term representing all forms of mechanical work). When the interaction is not purely thermal, some of the heat flowing out of one system will be converted into work, resulting in reduced heat flowing into the other system. (As ever, the words “heat flowing” here refer to internal energy being transferred thermally.) We have spoken of this before: some of this energy can be used to perform work, but it is and

remains energy, and there is no real case of something called “heat” being transformed into something called “work”. It’s wise to bear this in mind when analysing heat flow: the thermal transfer of energy generally does *not* define a conserved quantity.

- **Mechanical interaction:** Here, (3.179) reduces to

$$\frac{P_1 - P_2}{T_2} dV_1 > 0. \quad (3.182)$$

If system 1 has a lower pressure than system 2 ($P_1 < P_2$), then $dV_1 < 0$, meaning system 1 loses volume: any mechanical boundary between the two systems moves toward system 1. If system 1 has a higher pressure than system 2 ($P_1 > P_2$), then $dV_1 > 0$, and so system 1 gains volume: the mechanical boundary between the two systems moves toward system 2. In both cases, the mechanical boundary between the two systems moves toward the region of lower pressure, again as we expect intuitively.

- **Diffusive interaction:** Now, (3.179) reduces to

$$\frac{\mu_2 - \mu_1}{T_2} dN_1 > 0. \quad (3.183)$$

If system 1 has a lower chemical potential than system 2 ($\mu_1 < \mu_2$), then $dN_1 > 0$, meaning system 1 gains particles. If system 1 has a higher chemical potential than system 2 ($\mu_1 > \mu_2$), then $dN_1 < 0$, and so system 1 loses particles. In both cases, particles flow toward the region of lower chemical potential.

When the interacting systems above have reached equilibrium, each term in (3.179) becomes zero, meaning that all temperatures, pressures, and chemical potentials have become equal. This process of entropy growth might be viewed as a sort of engine behind phenomena that we loosely interpret as being driven by forces. Take, for example, *osmosis*, discussed in Chapter 4. There, separating two liquids of different chemical potentials by a membrane will cause particles at the higher potential to diffuse across the membrane. This is why we dehydrate after eating salty foods or drinking salty water. This diffusion manifests as a sizeable pressure that “forces the particles across the membrane”; and yet no mechanical pump is present that is “physically pushing” on the molecules.

Describing such an interaction by invoking entropy growth raises a philosophical question of cause and effect. Consider the mechanical interaction: our discussion ties entropy growth to the fact that forcing air into a chamber that terminates in a piston will make that piston move. Even simpler is the idea that entropy growth accounts for the observation that when you push a barrier to overcome any other force present, that barrier will move. Figure 3.20 shows the bending of a flexible barrier that separates gases at high

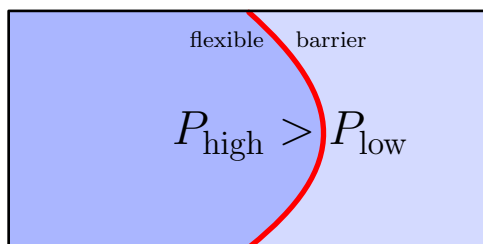


Fig. 3.20 A flexible barrier separating gases at high and low pressure will bend. Was this bending caused by forces of bouncing particles, or did it arise from the operation of the simple probabilistic mechanism of entropy growth?

and low pressure. Should we say that this bending is due to the forces of the particles bouncing around, or is it a simple manifestation of entropy growth? What has caused what here: has the flexing barrier caused the world's entropy to grow, or has entropy growth (the natural operation of the Second Law) caused the barrier to flex? Can we explain the whole of physics as nothing more than the blind growth of entropy?

If entropy growth always appears together with the operation of some force, there is perhaps nothing to be gained by insisting that the entropy growth caused the force. For example, we would be hard pressed to say that the act of opening a door to enter our house is nothing more than the blind operation of entropy growth in the universe. No predictive power results from taking such a view.

3.14 Integrating the Total Energy

The First Law is usually expressed in infinitesimal form, using inexact differentials for a general process (3.50) or exact differentials for a quasi-static process (3.164). The exact differentials in the law's quasi-static form enable dE to be integrated to give an expression for a system's total energy.²³

Irrespective of how the system evolved—how variables such as its temperature, entropy, and pressure varied while the system was being assembled, and which particles entered with what chemical potentials—we know that in its final state, the values of the intensive variables T, P, μ are constant throughout. This homogeneity of those intensive variables allows us to picture the system as partitioned into infinitesimal cells, each with its own infinitesimal values of the extensive variables: entropy, volume, and number of particles.

²³ One of our aims has been to calculate the system's energy in terms of its state variables $T, S, P, \dots$. This is one reason for why we sought to replace the inexact differentials in the most general statement of the First Law, (3.50), with exact differentials, resulting in (3.164).

Next, we imagine re-assembling those infinitesimal parts to create the system, applying the First Law as we go. This approach depends on the distinction between intensive and extensive variables in the First Law. For example, the entropies of two parts can be added, and combining those two parts won't produce an increase in their total entropy, because those parts have the same temperature. Also, any background potential energy is included in the chemical potential μ . In that case, refer to (3.164) to write

$$\begin{aligned} E &= \int dE = \int (T dS - P dV + \mu dN) \\ &= T \int dS - P \int dV + \mu \int dN. \end{aligned} \quad (3.184)$$

This integrates to yield the simple expression

$$E = TS - PV + \mu N. \quad (3.185)$$

The E here denotes all energy described by the First Law; but it might not include *all* of the system's energy. Relativity teaches us that the system will also have a “rest-mass” energy. Quantum mechanics brings in a “zero-point” energy. In the classical realm that most of statistical mechanics is concerned with, these energies play no part in any interactions, and so we can generally ignore them. (After all, classical mechanics studies non-relativistic projectile motion successfully, without the need to mention rest-mass and zero-point energies.) But sometimes, this extra energy cannot be ignored: the behaviour of liquid helium in Chapter 7 is a case in point. If we take the First Law's postulate that its energies can be added, and extend it to *all* types of energy, then we can include a “baseline” energy in (3.185):

$$E = E_{\text{baseline}} + TS - PV + \mu N. \quad (3.186)$$

But this baseline energy will factor out and thus cancel out in the expressions to follow, and so we won't write it explicitly. Equation (3.185) can be taken to give the system's total energy, and will be very useful later when we examine non-isolated systems and again must count microstates. Observe, from (3.139), that the total number of the system's microstates is

$$\Omega \stackrel{(3.139)}{=} e^{S/k} = \exp \frac{ST}{kT} \stackrel{(3.185)}{=} \exp \frac{E + PV - \mu N}{kT}. \quad (3.187)$$

In Section 5.1, this useful result will make quick work of deriving the Boltzmann distribution, a key player in statistical mechanics.

3.14.1 Swapping the Roles of Conjugate Variables

In general, the state of a system is specified by a set of three independent variables, one for each interaction: thermodynamical, mechanical, diffusive. Letting pressure–volume represent the mechanical interaction as usual, complete sets of independent variables are then [referring to (3.185)]

$$(T, P, \mu), (S, P, \mu), (T, V, \mu), (S, V, \mu), \dots, (S, V, N). \quad (3.188)$$

We are used to seeing the last set (S, V, N) in the First Law (3.164). But other choices can be useful, and these can be brought into (3.164) to define new pseudo-energies. To see how, refer to the total energy (3.185), and consider what happens when we define a new variable such as

$$F \equiv E - TS. \quad (3.189)$$

Differentiate this equation, to obtain

$$\begin{aligned} dF &= dE - S dT - T dS \quad [\text{and now invoke (3.164)}] \\ &= \cancel{T dS} - P dV + \mu dN - S dT - \cancel{T dS} \\ &= -S dT - P dV + \mu dN. \end{aligned} \quad (3.190)$$

Compare this to (3.164): the effect has been to swap the roles of S and T by switching from E to F , and, in particular, we have replaced an “intensive times d(extensive)” term with one of the form “extensive times d(intensive)”. Swapping one or more pairs of conjugate variables can be done in seven ways, four of which are well known owing to their long-time use in thermodynamics and chemistry. This technique of defining a new variable by adding or subtracting the product of the relevant conjugate pair is an example of a *Legendre transform*, which is also used in other areas of physics. In this section, we’ll examine each of these four well-known examples of the transform.

Realise that the argument used in Section 3.14 to integrate the energy dE in the First Law required all of the infinitesimals being integrated to be of extensive quantities. That means we cannot use it here to integrate (3.190) and end up inferring, incorrectly, that F equals $-ST - PV + \mu N$.

Helmholtz Energy F

The first example of a well-known Legendre transform is the one we just gave: the *Helmholtz energy* F in (3.189). Its infinitesimal form (3.190) shows that F is useful when analysing *isothermal nondiffusive* processes: these are processes for which temperature and particle number don’t change. Setting $dT = dN = 0$ in (3.189) produces

$$dF = -PdV \quad \text{for } T, N \text{ constant.} \quad (3.191)$$

Recall that $-PdV$ is the work we do when compressing a gas, and so it follows that

$$-dF = \left[\begin{array}{l} \text{mechanical work } PdV \text{ done by an isothermal} \\ \text{nondiffusive compressible system} \end{array} \right]. \quad (3.192)$$

That is, the loss in a system's Helmholtz energy equals the mechanical work done by that system. In general, the work that a system does is accompanied by the exhaust of useless heat, and so F quantifies how much energy in the system is “free” to be fully and usefully converted into mechanical work. This shows its historical relevance to heat engines, and is why the energy is often called the “Helmholtz free energy”.

Now consider two systems 1 and 2 that interact at constant temperature: diffusion is now allowed. Define their total Helmholtz energy to be the sum of their individual Helmholtz energies:

$$F = F_1 + F_2. \quad (3.193)$$

The increase in the combined system's Helmholtz energy is

$$\begin{aligned} dF &= dF_1 + dF_2 \\ &\stackrel{(3.190)}{=} -P_1 dV_1 + \mu_1 dN_1 - P_2 dV_2 + \mu_2 dN_2. \end{aligned} \quad (3.194)$$

During the interaction, the volume and particles gained by system 2 are lost by system 1:

$$dV_2 = -dV_1, \quad dN_2 = -dN_1. \quad (3.195)$$

Equation (3.194) then becomes

$$dF = (P_2 - P_1) dV_1 + (\mu_1 - \mu_2) dN_1. \quad (3.196)$$

But if the two subsystems have unequal pressures ($P_1 \neq P_2$), then, irrespective of which pressure is the greater,

$$(P_2 - P_1) dV_1 < 0. \quad (3.197)$$

And similarly, if the two subsystems have unequal chemical potentials, then it is always true that

$$(\mu_1 - \mu_2) dN_1 < 0. \quad (3.198)$$

From these two inequalities, it's clear that (3.196) says

$$dF < 0. \quad (3.199)$$

Hence, when two systems interact isothermally (i.e., at constant temperature), their total Helmholtz energy decreases as the combined system heads

toward equilibrium. At equilibrium, $P_1 = P_2$, $\mu_1 = \mu_2$, and thus (3.196) says $dF = 0$. So F has reached a minimum value, which it then holds indefinitely.

Enthalpy H

The next example of a Legendre transform seeks to switch variables in the second term of the First Law. Define the *enthalpy* H of a system:

$$H \equiv E + PV. \quad (3.200)$$

The same sort of differentiation procedure as used in (3.190) produces

$$dH = TdS + VdP + \mu dN. \quad (3.201)$$

Enthalpy is used extensively by chemists to analyse *isobaric nondiffusive* processes ($dP = dN = 0$). These are of great relevance to chemical reactions, which are often performed in an open vessel, and so are isobaric. For these,

$$dH = TdS \quad \text{for } P, N \text{ constant}, \quad (3.202)$$

and this is the energy dQ entering the system thermally. In other words,

$$-dH = \left[\begin{array}{l} \text{thermal energy } -TdS \text{ exiting} \\ \text{an isobaric nondiffusive system} \end{array} \right]. \quad (3.203)$$

In such chemical reactions, when the total loss in enthalpy $-\Delta H$ is positive (i.e., $\Delta H < 0$), the reaction is *exothermic*: thermal energy leaves the system and enters the reaction vessel, and this vessel then heats up. When $\Delta H > 0$, thermal energy goes *into* the chemical reaction, and the reaction is *endothermic*: the thermal energy entering the system must come from the reaction vessel, which then cools down.

In Section 4.1, we'll see that $TdS = C_P dT$, where C_P is the system's heat capacity at constant pressure. In other words, for the above chemical reactions,

$$-dH = -TdS = -C_P dT = C_P \times \text{system's temperature drop}. \quad (3.204)$$

It makes intuitive sense that the heat exiting the system should equal its heat capacity times the system's temperature drop.

Now consider two systems 1 and 2 that interact at constant pressure: diffusion between the two is allowed. Define their total enthalpy to be the sum of their individual enthalpies:

$$H = H_1 + H_2. \quad (3.205)$$

Using $dP_1 = dP_2 = 0$, the increase in the combined system's enthalpy is

$$dH = dH_1 + dH_2$$

$$\underline{\underline{(3.200)}} \quad dE_1 + P_1 dV_1 + dE_2 + P_2 dV_2. \quad (3.206)$$

During the interaction, the energy and volume gained by system 2 are lost by system 1:

$$dE_2 = -dE_1, \quad dV_2 = -dV_1. \quad (3.207)$$

Equation (3.206) then becomes

$$dH = (P_1 - P_2) dV_1, \quad (3.208)$$

which must be positive if $P_1 \neq P_2$. So, when two systems interact at constant pressure, their total enthalpy increases as the combined system heads toward equilibrium. At equilibrium it has reached a maximum, and no longer changes. Realise that we are here considering a *diffusive* process, and so the above comments relating the sign of ΔH to exo/endothermicity no longer apply.

You'll often encounter the statement in chemistry books that the change in the enthalpy of a chemical reaction determines its exo/endothermicity. But enthalpy change really only determines exo/endothermicity for *isobaric nondiffusive* processes, because only for these does it equate to “heat” TdS . More generally, exo/endothermicity is fully determined by the sign of TdS .

Gibbs Energy G

The third well-known Legendre transform defines a system's *Gibbs energy* G :

$$G \equiv E - TS + PV \quad \underline{\underline{(3.185)}} \quad \mu N. \quad (3.209)$$

The same sort of differentiation as in (3.190) produces

$$dG = -S dT + V dP + \mu dN. \quad (3.210)$$

G is useful for studying diffusive processes. These are often isothermal and isobaric ($dT = dP = 0$), in which case

$$dG = \mu dN \quad \text{for } T, P \text{ constant.} \quad (3.211)$$

Thus, in this case, dG is the energy brought into the system by particles that enter it, and so

$$-dG = \left[\begin{array}{l} \text{energy } -\mu dN \text{ carried by particles} \\ \text{exiting an isothermal isobaric system} \end{array} \right]. \quad (3.212)$$

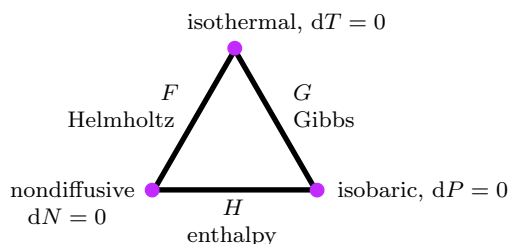


Fig. 3.21 A mnemonic triangle showing the processes (at the vertices) for which the Helmholtz energy, enthalpy, and Gibbs energy are most useful

It follows that the total decrease in a system's Gibbs energy $-\Delta G$ in such a process equals its chemical potential μ (potential energy per particle) times the drop in its number of particles.

When two systems interact at constant temperature and pressure, (3.211) says that

$$dG_1 = \mu_1 dN_1, \quad dG_2 = \mu_2 dN_2. \quad (3.213)$$

Defining the Gibbs energy of the combined system as the sum of the individual Gibbs energies produces

$$dG = dG_1 + dG_2 = \mu_1 dN_1 + \mu_2 dN_2. \quad (3.214)$$

But the number of particles gained by system 2 equals that lost by system 1: $dN_2 = -dN_1$. Equation (3.214) then becomes

$$dG = (\mu_1 - \mu_2) dN_1, \quad (3.215)$$

which must be negative if $\mu_1 \neq \mu_2$. Hence, when two systems interact at constant temperature and pressure, their total Gibbs energy decreases as the combined system heads toward equilibrium, at which point it no longer changes. We'll make use of G in Chapter 4.

Note that (3.200) and (3.209) say that $G = H - TS$. Chemists infer the direction of a reaction from G , and they infer exo/endothermicity from H . These two quantities are linked by entropy S ; hence, a knowledge of S can be useful to physical chemists. But it turns out that S is very difficult to measure, as we'll see ahead in Section 4.1.1. So, in practice, G and H tend to be treated as distinct quantities.

Figure 3.21 has a mnemonic triangle that shows the regimes in which F , G , and H are most useful.

The Gibbs–Duhem Equation

The last well-known Legendre transform of the First Law comes about by differentiating (3.185) for the total internal energy, and cancelling dE internally:

$$dE = S dT + T dS - V dP - P dV + N d\mu + \mu dN. \quad (3.216)$$

What is left is the *Gibbs–Duhem equation*:

$$S dT - V dP + N d\mu = 0. \quad (3.217)$$

[Or rather, (3.217) is just one of many Gibbs–Duhem equations, each one arising from the particular choice of term quantifying the First Law’s mechanical interaction. We have used the customary term $-PdV$ in (3.216).]

Whereas the First Law combines infinitesimal changes in the extensive variables S, V, N , the Gibbs–Duhem equation relates changes in the *intensive* variables T, P, μ . It also shows that if one of these intensive variables changes, then at least one other intensive variable must change for (3.217) to continue to hold.

Note that (3.209) says that $dG = N d\mu + \mu dN$. Where did the $N d\mu$ go in (3.211)? Recall that the relevant process was isothermal and isobaric: $dT = dP = 0$. Gibbs–Duhem (3.217) then says that here, $N d\mu = 0$. Hence we obtain $dG = \mu dN$, as in (3.211).

3.14.2 Maxwell Relations

In equation (1.138) we encountered the *coefficient of isothermal compressibility*,

$$\kappa = \frac{-1}{V} \left(\frac{\partial V}{\partial P} \right)_{T,N} = \left[\begin{array}{l} \text{relative decrease in } V \text{ with } P \\ \text{at constant } T \text{ and } N \end{array} \right]. \quad (3.218)$$

This is a partial derivative of variables present in the First Law. Another case of such a partial derivative is the *coefficient of thermal expansion*,

$$\beta \equiv \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_{P,N} = \left[\begin{array}{l} \text{relative increase in } V \text{ with } T \\ \text{at constant } P \text{ and } N \end{array} \right]. \quad (3.219)$$

These derivatives, and others like them, can be related via a straightforward theory of partial derivatives. In a thermodynamics context, this theory produces *Maxwell relations*, which are equalities of mixed partial derivatives of variables present in the First Law. Maxwell relations enable experimenters to switch focus between the choices of thermodynamical variables in (3.188),

depending on the relevant experimental regime. The independent variables of most use in an experiment tend to be those that are either easily constrained or easily measured; also, one variable each is needed to describe the thermal, mechanical, and diffusive aspects of a thermodynamic system.

To see how Maxwell relations are produced, begin with the infinitesimal increase of a function $f(x, y, z)$:

$$df = X dx + Y dy + Z dz, \quad (3.220)$$

where X, Y, Z are each a function of x, y, z . Partial derivatives of well-behaved functions always commute, by which we mean that for partials with respect to, say, x and y , the following is true:

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}. \quad (3.221)$$

Applying (3.221) to the expression in (3.220) yields

$$\left(\frac{\partial X}{\partial y}\right)_{x,z} = \left(\frac{\partial Y}{\partial x}\right)_{y,z}. \quad (3.222)$$

This is called a Maxwell relation resulting from (3.220). Now consider replacing (3.220) with the First Law:

$$dE = T dS - P dV + \mu dN. \quad (3.223)$$

The analogue to (3.222) is then

$$\left(\frac{\partial T}{\partial V}\right)_{S,N} = -\left(\frac{\partial P}{\partial S}\right)_{V,N}. \quad (3.224)$$

Other Maxwell relations can be produced by applying the same idea to the infinitesimal expressions for the energies F, G, H in the last few pages. For example, it follows from (3.190) that

$$-\left(\frac{\partial P}{\partial N}\right)_{T,V} = \left(\frac{\partial \mu}{\partial V}\right)_{T,N}. \quad (3.225)$$

A tool for manipulating Maxwell relations emerges from the following question. Suppose we have three variables x, y, z , that depend on each other in some way that when plotted, produces a surface $z = z(x, y)$. Now hold z fixed, which equates to slicing the surface to produce a curve of points (x, y) in a plane parallel to the xy plane. Examining this curve of (x, y) points, when we vary y , how does x change—how does it depend on y ?

To answer this question, imagine following this curve at fixed z , and inspecting all points (x, y) as we go. Along the curve, $dz = 0$, so write

$$dz = \left(\frac{\partial z}{\partial x}\right)_y dx + \left(\frac{\partial z}{\partial y}\right)_x dy, \quad (3.226)$$

and set $dz = 0$, to obtain

$$-\left(\frac{\partial z}{\partial x}\right)_y dx = \left(\frac{\partial z}{\partial y}\right)_x dy. \quad (3.227)$$

This rearranges to

$$dx = \frac{-(\partial z/\partial y)_x}{(\partial z/\partial x)_y} dy. \quad (3.228)$$

These increments dx and dy have occurred at constant z , and so it follows that

$$\left(\frac{\partial x}{\partial y}\right)_z = \frac{dx}{dy} \text{ in (3.228)} = \frac{-(\partial z/\partial y)_x}{(\partial z/\partial x)_y}. \quad (3.229)$$

For an example of using Maxwell relations, refer to (3.218) and (3.219) to write (with all expressions at constant N)

$$\frac{\kappa}{\beta} = \frac{-(\partial V/\partial P)_T}{(\partial V/\partial T)_P} \stackrel{(3.229)}{=} \left(\frac{\partial T}{\partial P}\right)_V. \quad (3.230)$$

We see that κ/β is the increase in temperature with pressure at constant volume (and particle number). We'll make good use of a Maxwell relation in Section 4.2.2.

Finally, recall that because $(\partial z/\partial y)_x = 1/(\partial y/\partial z)_x$, we can bring each of the two factors on the right-hand side of (3.229) to its left-hand side. That equation then changes to a symmetrical “cyclic” form:

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1. \quad (3.231)$$

3.15 Excursus: Pressure and Temperature of a Star's Interior

“We can never know the temperature at the core of a star” was stated in the nineteenth century as an example of the idea that knowledge of some things in Nature must be forever beyond our reach. In fact, already at that time, astronomers were using spectroscopy to analyse the light from stars, and this gave them what amounted to a view inside stellar interiors. With the years and the century that followed came further theories about how stars work—such as nuclear physics—and astronomers began to peer ever deeper into the stellar core.

It's a remarkable fact that we *can* form an estimate of the pressure and temperature of a star's interior based on some reasonable assumptions about its physical makeup. This furnishes the discussion that will take place over the next few pages. In Chapter 9, we'll estimate the temperature of our Sun's surface based on the spectrum of light that it emits. That is a rather different calculation, and these very different approaches taken towards studying the Sun's interior and surface underline the richness of the physics of temperature.

Balance Between Pressure and Gravity

We begin with the idea that the Sun's energy output is extremely constant. Not all stars spend their lives this calmly; the light output from commonly found *variable stars* grows and diminishes over periods of days or months, usually in very predictable ways with constant periods. The physics of these variable stars is complex, and any study of them will always begin with a study of the simpler environment that exists inside our very stable Sun.

The Sun owes its stability to an equilibrium between its gravity and its internal pressure generated as it undergoes nuclear fusion. As a star is born from a cloud of dust and gas, nuclear fusion is thought to begin when the material is slowly drawn together under the influence of the combined gravity of all the particles. This is not something that happens overnight. But when the matter has come together sufficiently closely, the inverse-square nature of gravity begins to dominate local particle motions and exerts its compacting pull. The proto-stellar cloud of mostly hydrogen gas starts to contract, converting the gravitational potential energy of its atoms into kinetic energy: their speeds increase, driving up the cloud's temperature. Eventually, these atoms start to interact with sufficient strength that nuclear fusion commences, and the fledgling star begins to shine.

Throughout the star's life, the outward pressure produced by its nuclear furnace balances the immense gravity holding it together. If the furnace were to grow too hot, the star would expand, diminishing the interactions between its particles, which would lead to cooling and contraction. If it were to contract too much, the increased particle interactions would turn its nuclear furnace up and bring the contraction to a halt. This equilibrium between its furnace and gravity persists for several thousand million years, during which the star converts a good part of its hydrogen fuel to helium. Deep in the star's interior, the higher pressure also converts some helium into lithium; and deeper still, similar conversions occur up the chain of atomic numbers. A series of shells of elements of ever-higher atomic number is thought to form, ending with iron in the core. The fusion processes cannot proceed beyond the production of iron, since, for higher atomic numbers, fusion reactions are not energetically favourable to occur. Hence, reaching down to the core of the star, we will find some iron, and outward a mix of lighter elements in shells, ending with the bulk of the star being the lightest of them all, hydrogen.

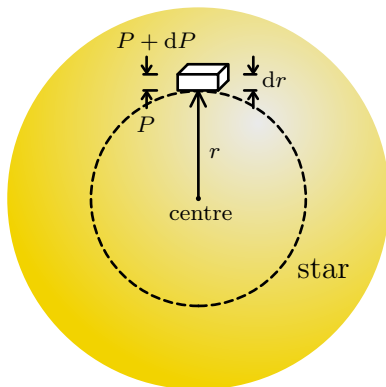


Fig. 3.22 A small oblong element of matter within the star feels different pressures from above and below. The difference in these pressures balances the force of gravity on the element resulting from the rest of the star’s matter. When the star’s density is spherically symmetric, it’s a remarkable fact of gravity’s inverse-square strength that the contributions of gravity forces on the element from all matter farther from the star’s centre than that element will cancel each other out, resulting in no gravity force on the element. The sum of the force contributions from all matter closer to the star’s centre than the element turns out to equal the force on the element due to an imagined particle of the same mass as that “interior mass”, placed at the centre

With the star settled in its main phase of quiet burning, we are in a position to peer into its interior by first noting that everywhere throughout, the outward pressure of its nuclear furnace must act to balance the inward pull of its gravity. Referring to Figure 3.22, the fact that pressure acts in all directions means that the outward pressure force on a small matter element of area $\text{d}A$ and radial thickness $\text{d}r$ at a distance r from the centre is

$$\begin{aligned} \left[\begin{array}{l} \text{outward pressure force} \\ \text{on element} \end{array} \right] &= \left[\begin{array}{l} |\text{force from} \\ \text{pressure below}| \end{array} \right] - \left[\begin{array}{l} |\text{force from} \\ \text{pressure above}| \end{array} \right] \\ &= P(r) \text{d}A - P(r + \text{d}r) \text{d}A. \end{aligned} \quad (3.232)$$

This is positive and (for our equilibrium) equal to the magnitude of the gravitational force, which tries to pull the matter inward. The gravity force on the matter element is given by the usual “ GMm/r^2 ” expression integrated over all point sources of mass in the star. A standard result of this integration for our assumed spherically symmetric mass density is that the total gravity force on the element is identical to that of an imagined particle at $r = 0$ with mass $m(r)$, the mass of the portion of the star *within* radius r of its centre. If the matter has density $\varrho(r)$, equating the above pressure force with gravity produces

$$P(r) \text{d}A - P(r + \text{d}r) \text{d}A = \frac{Gm(r)\varrho(r) \text{d}A \text{d}r}{r^2}. \quad (3.233)$$

This simplifies to

$$\frac{dP}{dr} = \frac{-Gm\rho}{r^2}, \quad (3.234)$$

which is known as the *equation of hydrostatic support*. We can eliminate $\rho(r)$ from this equation in favour of $m(r)$ by noting that an entire shell of matter of radius r has mass $dm = \rho(r) 4\pi r^2 dr$, in which case

$$\rho(r) = \frac{1}{4\pi r^2} \frac{dm}{dr}. \quad (3.235)$$

We substitute this ρ into (3.234), obtaining

$$dP = \frac{-Gm}{r^2} \frac{dm}{4\pi r^2}. \quad (3.236)$$

Now integrate this equation from the centre to the surface, for a star of total mass M :

$$P_{\text{surface}} - P_{\text{centre}} = \int_0^M \frac{-Gm dm}{4\pi r^4}. \quad (3.237)$$

Remember that m is the mass contained in a sphere of radius r ; so, as m increases, then so must r , and hence r is a function of m . This function $r(m)$ is required in order for us to evaluate the integral in (3.237).

We could make an educated guess as to the form of $r(m)$; but alternatively and with very little effort, we can obtain a lower limit on the pressure at the star's centre. If the star has radius R , then $r^4 \leq R^4$. Hence,

$$\frac{-1}{r^4} \leq \frac{-1}{R^4}. \quad (3.238)$$

This allows (3.237) to be written as

$$P_{\text{surface}} - P_{\text{centre}} = \int_0^M \frac{-Gm dm}{4\pi r^4} \leq \int_0^M \frac{-Gm dm}{4\pi R^4} = \frac{-GM^2}{8\pi R^4}. \quad (3.239)$$

Thus,

$$P_{\text{surface}} + \frac{GM^2}{8\pi R^4} \leq P_{\text{centre}}. \quad (3.240)$$

The core pressure will be much larger than the surface pressure, so we simply write

$$P_{\text{centre}} > \frac{GM^2}{8\pi R^4}. \quad (3.241)$$

Our Sun has $M = 2.0 \times 10^{30}$ kg and $R = 7.0 \times 10^8$ m, resulting in

$$P_{\text{centre}} > 450 \text{ million Earth atmospheres}. \quad (3.242)$$

In fact, nothing about this analysis pertains only to stars; the same result could be used to calculate a minimum core pressure for Earth. Here, we use a

mass of $M = 6.0 \times 10^{24}$ kg and $R = 6.4 \times 10^6$ m, resulting in a minimum core pressure of 600,000 atmospheres. The actual figure is thought to be around 3.5 million atmospheres.

Average Temperature of the Star

We can estimate an average temperature of a star through the use of some ideas of ideal gases and thermodynamics. Stellar temperatures are high enough to create a plasma (where electrons are stripped from their nuclei), and this plasma is thought to behave as an ideal gas for the densities considered here, which are around 1400 kg/m^3 for our Sun.

Assume that everywhere in the star, any small region has enough thermodynamic equilibrium to enable a local temperature to be reasonably well defined—but we will allow the star’s temperature distribution to be highly dependent on distance from its core. This assumption allows for a kind of mean temperature $\bar{T}$ to be calculated for the star as a whole. Adding temperatures to compute a mean may seem like an odd thing to do, given that it makes no physical sense to say “1000 K (from region 1) + 1000 K (from region 2) = 2000 K”, or “the mean temperature of a block of ice and a cup of boiling water is 50°C ”. But the numerical mean of a set of temperatures “50, 52, 55, 58, 61°C ” is meaningful as a simple way of representing the entire spread of temperatures by a single number.

In that case, we’ll calculate $\bar{T}$ by summing the temperatures of all the small constituents of the star and dividing that sum by the total number of these constituents. Each of these constituents is modelled mathematically as an infinitesimal number of particles: a very small region indeed!, but at least one that allows us to proceed. In that case, the mean temperature is an integral over the total number of particles N_{total} in the star:

$$\bar{T} \equiv \frac{1}{N_{\text{total}}} \int_0^{N_{\text{total}}} T \, dN. \quad (3.243)$$

We will assume the particles have just three (translational) quadratic energy terms, and so a particle with temperature T has a mean kinetic energy $3kT/2$. We’ll relate this kinetic energy to the potential energy of the star via the equation of hydrostatic support, (3.234), and thence back to the radius of the star via (3.238). Do this by first introducing the star’s total kinetic energy $E_{k \text{ total}}$ into (3.243):

$$\bar{T} = \frac{\int_0^{N_{\text{total}}} \frac{3}{2} kT \, dN}{\frac{3}{2} k N_{\text{total}}} = \frac{E_{k \text{ total}}}{\frac{3}{2} k N_{\text{total}}}. \quad (3.244)$$

We require the particles’ total kinetic energy $E_{k \text{ total}}$:

$$E_{k \text{ total}} = \int_0^{N_{\text{total}}} \frac{3}{2} kT \, dN. \quad (3.245)$$

Now remember that the pressure P of the infinitesimal constituent holding dN particles in volume dV relates to that number of particles via the ideal-gas law:

$$P dV = kT \, dN. \quad (3.246)$$

Equation (3.245) then becomes

$$E_{k \text{ total}} = \frac{3}{2} \int_0^{V_{\text{total}}} P \, dV. \quad (3.247)$$

This integrates by parts from the centre to the surface as:

$$E_{k \text{ total}} = \frac{3}{2} \left[PV \right]_{\text{centre}}^{\text{surface}} - \frac{3}{2} \int_{\text{centre}}^{\text{surface}} V \, dP. \quad (3.248)$$

If the star's surface pressure is comparatively negligible, (3.248) simplifies to

$$E_{k \text{ total}} = -\frac{3}{2} \int_{\text{centre}}^{\text{surface}} V \, dP. \quad (3.249)$$

Now use the fact that $V = \frac{4}{3} \pi r^3$ implicitly involves the mass m —because r relates to m by way of $m(r)$. Write

$$3V \, dP = 4\pi r^3 \, dP \stackrel{(3.236)}{=} \frac{-Gm \, dm}{r}. \quad (3.250)$$

This converts (3.249) into

$$E_{k \text{ total}} = \frac{-1}{2} \int_0^M \frac{-Gm \, dm}{r} = \frac{-1}{2} \times \left[\begin{array}{c} \text{total gravitational} \\ \text{potential energy of} \\ \text{the star} \end{array} \right]. \quad (3.251)$$

This equation is actually an instance of the virial theorem (3.113).²⁴

The virial theorem in a form resembling (3.251) appears in a much simpler context when we examine a small mass m orbiting a large mass M , where the centre of mass of these two particles lies approximately at M . The centripetal acceleration v^2/r of m orbiting M equals the gravity force per unit mass acting on m : thus, $v^2/r = GM/r^2$. Now use this last expression to calculate the kinetic energy of m :

²⁴ Specifically, for the case of two particles with masses m_i and m_j interacting gravitationally, we must set $\mathbf{F}_{ij} = -Gm_i m_j / r_{ij}^2 \hat{\mathbf{r}}_{ij}$ in (3.115), where $\mathbf{r}_{ij}$ points from particle j to particle i , and $\hat{\mathbf{r}}_{ij}$ is the associated unit vector.

$$\begin{aligned} \frac{1}{2}mv^2 &= \frac{1}{2}GMm/r \\ &= -\frac{1}{2} \times \text{gravitational potential energy of the masses.} \end{aligned} \quad (3.252)$$

This expression is a kind of simplified instance of (3.251).

Now use (3.244) and (3.251) to write

$$\bar{T} = \frac{\frac{1}{2} \int_0^M \frac{Gm \, dm}{r}}{\frac{3}{2}kN_{\text{total}}} . \quad (3.253)$$

Here, we face the same difficulty as in (3.237): without knowledge of how r varies with m , we can only find a lower limit for the mean temperature. Use $1/r \geq 1/R$ to write (3.253) as

$$\begin{aligned} \bar{T} &\geq \frac{\int_0^M \frac{Gm \, dm}{R}}{3kN_{\text{total}}} = \frac{GM^2/R}{6kN_{\text{total}}} = \frac{GM}{6Rk} \times \frac{M}{N_{\text{total}}} \\ &= \frac{GM}{6Rk} \times \text{average mass of a particle,} \end{aligned} \quad (3.254)$$

where the “ R ” in (3.254) is the star’s radius, not the gas constant! Given that stars are thought to be mostly hydrogen plasma, the average particle mass is at least one half the mass of a proton m_p . We can thus write

$$\bar{T} \geq \frac{GMm_p}{12Rk} . \quad (3.255)$$

For the Sun, this amounts to a minimum average temperature of 2 million kelvins. The core temperature is believed to be around 15 million K. Thermodynamics has given us the remarkable ability to peer inside a star.



Chapter 4

The First Law in Detail

In which we derive some standard results related to heat, work, and chemical reactions. We follow the mathematics of heat flow, and show what it has in common with diffusion and radar signal processing. We enquire why air escaping a tyre grows cold. We then explore density and pressure in the atmosphere and ocean, determine some melting and boiling points, and examine chemical equilibrium.

The three terms of the First Law of Thermodynamics, (3.50), relate a system's internal energy increase to its thermal, mechanical, and diffusive changes. Historically, the study of thermal and mechanical changes enabled heat processes to be harnessed into engines, from the age of steam through to modern automobiles and jet propulsion. Studies of the law's third term, diffusion, brought chemists an understanding of how to quantify the rates and directions of chemical reactions. In this chapter, we'll study some important aspects of each of these terms that comprise the First Law.

4.1 The First Term: Thermal Interaction

We are all familiar with the fact that some materials are easy to heat: they are more economical to bring to a given temperature than are others. Crystalline materials such as metals fall into this group, whereas more elaborate structures such as porcelain require more energy input for a given rise in temperature. Evidently, porcelain's complex molecules provide a larger number of quadratic energy terms than is the case for crystals. This prompts us to define the *heat capacity* of a material, which is generally a function of temperature, and so requires an *infinitesimal* temperature increase dT in its definition. When a parameter A is held fixed during the heating process, the material's heat capacity C_A denotes the infinitesimal energy dQ added thermally, divided by the resulting infinitesimal temperature increase dT :

$$C_A, \text{ the heat capacity at constant } A, \equiv dQ/dT \text{ for } A \text{ held fixed.} \quad (4.1)$$

This equation applies only when the temperature of the substance being heated *does* indeed increase; it does not apply during a phase change that occurs at constant temperature. For example, the temperature of a block of water ice remains a constant 0°C as it melts, in which case the concept of

heat capacity doesn't apply to this process. Here, we must instead use water's *latent heat of fusion* L_{fusion} , which is the amount of thermal energy absorbed by water ice as it melts at 0°C . We'll encounter latent heat again shortly.

Remember that (4.1) involves an inexact differential, and so is not a derivative! For quasi-static processes, $\text{d}Q = T\text{d}S$, which enables (4.1) to be written as

$$C_A = T \left(\frac{\partial S}{\partial T} \right)_A. \quad (4.2)$$

We are generally interested in the heat capacity of systems with a fixed particle number N and with no background potential energy. For these, the First Law in the form “ $\text{d}E = \text{d}Q + \text{d}W$ ” produces

$$\underbrace{C_A \text{d}T}_{\text{“heat in”}} = \text{d}Q = \underbrace{\text{d}E}_{\substack{\text{increase in} \\ \text{total energy}}} + \underbrace{-\text{d}W}_{\substack{\text{work done} \\ \text{by system}}} \text{ at constant } A. \quad (4.3)$$

For example, a gas undergoing a quasi-static volume change is described by¹ $\text{d}W = -P\text{d}V$:

$$\underbrace{C_A \text{d}T}_{\text{“heat in”}} = \text{d}Q = \underbrace{\text{d}E}_{\substack{\text{increase in} \\ \text{total energy}}} + \underbrace{P\text{d}V}_{\substack{\text{work done} \\ \text{by gas}}} \text{ at constant } A. \quad (4.4)$$

For studies of gases and crystals, A usually denotes either volume or pressure. Write (4.4) as²

$$\text{d}E = C_A \text{d}T - P\text{d}V \text{ at constant } A, \quad (4.5)$$

and consider heating the material to increase its temperature by $\text{d}T$. The internal energy E of an ideal gas or crystal depends only on its temperature via (3.100)'s $E = \nu NkT/2$, and so the value of the energy increase $\text{d}E$ cannot depend on whether we hold pressure or volume fixed while heating the material: $\text{d}E$ must have the same value for both processes. Holding volume (and particle number) fixed simplifies (4.5) to

$$\text{d}E = C_V \text{d}T \text{ at constant } V, N. \quad (4.6)$$

¹ When developing a feel for the energy conservation expressed in these equations involving $P\text{d}V$, always remember that the work done *on* the system is $-P\text{d}V$; the work done *by* the system is $P\text{d}V$.

² Why is the First Law for quasi-static processes not written as (4.5)—wouldn't this be simpler than having to define entropy for use in (3.164)? One reason is that (4.5) must specify “at constant A ”, which makes it too specific for general use. Another reason is that replacing $T\text{d}S$ with $C_A \text{d}T$ is only useful if T is a good variable to describe a state. But the observation that a state can evolve at constant T (think of melting ice) indicates that T is not as useful for describing the evolving state as entropy S , which undoubtedly *will* increase during a phase change such as ice melting. Also, $C_A \text{d}T$ has the form “extensive times d(intensive)”, which is not conducive to being integrated in the way described in Section 3.14.

This is equivalent to writing

$$C_V = \left(\frac{\partial E}{\partial T} \right)_{V,N}. \quad (4.7)$$

Holding pressure (and, as always, particle number) fixed gives the *same* energy increase dE as simply (4.5), now with A set to P :

$$dE = C_P dT - P dV \text{ at constant } P, N. \quad (4.8)$$

In Section 4.2.2, we'll find the following idea useful. Given that the gas's enthalpy is $H = E + PV$ from (3.200), at constant pressure and particle number, we have

$$dH = dE + P dV \text{ at constant } P, N. \quad (4.9)$$

This allows (4.8) to be rearranged to produce

$$dH = C_P dT \text{ at constant } P, N, \quad (4.10)$$

which is equivalent to

$$C_P = \left(\frac{\partial H}{\partial T} \right)_{P,N}. \quad (4.11)$$

How are C_P and C_V related for an ideal gas? Hold pressure fixed, and differentiate $PV = NkT$ to obtain $P dV = Nk dT$. This is then inserted into (4.8) to give us

$$dE = C_P dT - Nk dT \text{ at constant } P, N. \quad (4.12)$$

Equations (4.6) and (4.12) refer to different processes that give identical energy increases dE . Equating their right-hand sides then yields

$$C_V dT = C_P dT - Nk dT; \quad (4.13)$$

or, noting that $Nk = nR$ for n moles of gas, with R the gas constant,

$$\boxed{C_P = C_V + Nk = C_V + nR.} \quad (4.14)$$

It's generally more convenient to work with a heat capacity per unit measure. So, define

$$\begin{aligned} \text{specific heat capacity } C^{\text{sp}} &\equiv C/m \quad (m = \text{total mass of the substance}), \\ \text{molar heat capacity } C^{\text{mol}} &\equiv C/n \quad (n = \text{number of moles of the substance}). \end{aligned} \quad (4.15)$$

(Specific heat capacity is often just called “specific heat”.) Clearly, with M_{mol} being the mass of one mole,

$$C^{\text{sp}} = \frac{C}{m} = \frac{C}{nM_{\text{mol}}} = \frac{C^{\text{mol}}}{M_{\text{mol}}}. \quad (4.16)$$

Equation (4.14) leads to an expression for ideal gases that is useful in chemistry, where moles are ubiquitous:

$$C_P^{\text{mol}} = C_V^{\text{mol}} + R. \quad (4.17)$$

C_V^{mol} has a particularly simple form at temperatures not near absolute zero. This is expressed by the *Dulong–Petit law*, as follows.

The Dulong–Petit Law

We calculate C_V^{mol} for a diatomic ideal gas and for a crystal, at room temperature. Call on (4.15) and (4.7) to write

$$C_V^{\text{mol}} = \frac{1}{n} \left(\frac{\partial E}{\partial T} \right)_{V,N}, \quad (4.18)$$

for the number of moles $n = N/N_A$ and total energy E . The equipartition theorem says that the total energies of both a gas and a crystal whose N particles each have ν quadratic energy terms are $E = \nu NkT/2$. Equation (4.18) then becomes

$$C_V^{\text{mol}} = \frac{N_A}{N} \frac{\nu Nk}{2} = \frac{\nu R}{2}. \quad (4.19)$$

At room temperature, each molecule of the diatomic gas turns out to have $\nu = 5$ quadratic energy terms: 3 translational and 2 rotational, as we'll show in Section 5.6. Thus

$$C_V^{\text{mol}} = \nu R/2 \simeq 5/2 \times 8.314 \text{ JK}^{-1}\text{mol}^{-1} \simeq 20.8 \text{ JK}^{-1}\text{mol}^{-1}. \quad (4.20)$$

This agrees well with tabulated results: for example, the molar heat capacity of O_2 at room temperature is $21.0 \text{ JK}^{-1}\text{mol}^{-1}$.

Each atom of the crystal has $\nu = 6$ quadratic energy terms, because each of its vibrations has 2 terms associated with each of its 3 dimensions of motion, as you'll recall from (2.68) and the discussion around it. For the crystal, then,

$$C_V^{\text{mol}} = \nu R/2 = 3R \simeq 24.9 \text{ JK}^{-1}\text{mol}^{-1}. \quad (4.21)$$

This result for crystals is called the *Dulong–Petit law*. It was found experimentally in the early nineteenth century by Dulong and Petit, while ideas of atomic mass and moles were being developed. The “law” is observed to hold for various crystals, typically at temperatures above 100 K. At lower temperatures, the quantised nature of energy alters the above

discussions; and indeed, as the temperature tends toward zero, C_V is also observed to tend toward zero. At such low temperatures, we must replace the Dulong–Petit law with other analyses that we’ll come to in Chapters 5, 7, and 8.

Now define a slightly temperature-dependent parameter γ , which is usually treated as a constant:

$$\gamma \equiv \frac{C_P}{C_V} = \frac{C_P^{\text{sp}}}{C_V^{\text{sp}}} = \frac{C_P^{\text{mol}}}{C_V^{\text{mol}}}. \quad (4.22)$$

(This useful parameter seems to have no specific name; it is just the ratio of heat capacities.) For an ideal gas with ν quadratic energy terms per particle, write

$$C_V \stackrel{(4.7)}{=} \left(\frac{\partial (\nu N k T / 2)}{\partial T} \right)_{V,N} = \nu N k / 2. \quad (4.23)$$

It follows that

$$\gamma = \frac{C_P}{C_V} \stackrel{(4.14)}{=} \frac{C_V + Nk}{C_V} = 1 + \frac{Nk}{\nu N k / 2} = 1 + \frac{2}{\nu}. \quad (4.24)$$

Measurements of the heat capacities C_P and C_V of real gases that are approximately ideal serve to determine the number of quadratic energy terms of the gas molecules, from (4.24). This number then yields information about the structure of the gas molecules.

Here is an example of the above ideas. The specific heat capacity of helium at constant pressure is $C_P^{\text{sp}} = 5230 \text{ J K}^{-1} \text{ kg}^{-1}$. Given the mass of a helium atom $m_{\text{He}} = 6.7 \times 10^{-27} \text{ kg}$, what conclusion can we draw about whether helium “particles” are single He atoms, or perhaps He_2 , or He_3 , etc.?

Begin to answer this question by assuming that a helium particle is He_n for some number n of helium atoms, with the whole particle having ν quadratic energy terms, and that helium is an ideal gas. We are given a measured value of C_P^{sp} , so trace this back to ν :

$$C_P^{\text{sp}} \stackrel{(4.15)}{\rightarrow} C_P \stackrel{(4.14)}{\rightarrow} C_V \stackrel{(4.23)}{\rightarrow} \nu. \quad (4.25)$$

That is,

$$\begin{aligned} C_P^{\text{sp}} &\stackrel{(4.15)}{=} \frac{C_P \text{ of } N \text{ particles}}{\text{mass of } N \text{ particles}} \stackrel{(4.14)}{=} \frac{C_V + Nk}{N n m_{\text{He}}} \\ &\stackrel{(4.23)}{=} \frac{\nu N k / 2 + Nk}{N n m_{\text{He}}}. \end{aligned} \quad (4.26)$$

Solving this for ν gives us the following, whose middle expression is evaluated using SI units:

$$\nu = 2 \left(\frac{C_P^{\text{sp}} n m_{\text{He}}}{k} - 1 \right) \simeq 2 \left(\frac{5230n \times 6.7 \times 10^{-27}}{1.381 \times 10^{-23}} - 1 \right) \simeq 5.08n - 2. \quad (4.27)$$

Tabulate ν as a function of the first few values of n :

n :	1	2	3	4
ν :	3.1	8.2	13.2	18.3

For $n = 1$, a value of $\nu = 3$ quadratic energy terms makes good sense: a helium atom has only translational kinetic energy in 3 dimensions. For $n = 2$, we are harder pressed to account for $\nu = 8$ energy terms of a He_2 molecule: we might consider 3 translational modes, 3 rotational, and 2 vibrational, but we'll show, in Section 5.6.1, that a diatomic molecule can actually have only 2 modes of rotation. And accounting for 13 energy terms for He_3 is impossible. So, $n = 1$ is a good choice, and we conclude that helium is probably monatomic.

Specific heat capacity tends to be fairly constant over everyday temperature ranges of interest. This enables the $\text{d}Q = C_A \text{d}T$ of (4.1) to be integrated to give the total thermal energy Q that must be absorbed by a substance to increase its temperature by ΔT . For a mass m comprising n moles of particles,

$$Q = \int \text{d}Q = \int C_A \text{d}T \simeq C_A \int \text{d}T = C_A \Delta T. \quad (4.28)$$

And since $C_A = m C_A^{\text{sp}} = n C_A^{\text{mol}}$, we have the standard expression for the thermal energy required to raise the temperature of a substance by ΔT :

$$Q \simeq m C_A^{\text{sp}} \Delta T = n C_A^{\text{mol}} \Delta T. \quad (4.29)$$

A classic example of thermal energy transfer appears in Figure 4.1. Drop a hot block of aluminium into a tub of cold water, which is insulated from its surroundings. What is the eventual temperature of the whole at thermal equilibrium? For liquids and solids, $C_P \simeq C_V$, so we'll make no distinction between these two heat capacities here, and deal simply with the heat capacity C_{Al} of the aluminium and C_{w} of water. Now recall the note after (3.181), which allows us to treat thermal energy transfer as the flow of a conserved quantity, provided no change in volume or exchange of particles occurs. That's approximately the case here, and so we can write

$$\begin{aligned} \overbrace{-Q_{\text{Al}}}^{\text{energy lost by aluminium}} &= \overbrace{Q_{\text{w}}}^{\text{energy gained by water}}. \\ \equiv -Q_{\text{Al}} \stackrel{(4.28)}{=} -C_{\text{Al}} \Delta T_{\text{Al}} &\equiv Q_{\text{w}} \stackrel{(4.28)}{=} C_{\text{w}} \Delta T_{\text{w}} \end{aligned} \quad (4.30)$$

Remember that “delta equals final minus initial”:

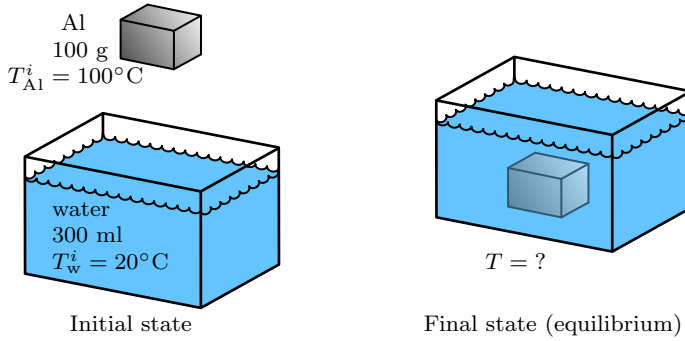


Fig. 4.1 A block of hot aluminium (initial temperature T_{Al}^i) is dropped into cold water (initial temperature T_w^i). What is the equilibrium temperature T of block and water?

$$\Delta T_{Al} = T - T_{Al}^i, \quad \Delta T_w = T - T_w^i. \quad (4.31)$$

Thus, (4.30) becomes

$$-C_{Al}(T - T_{Al}^i) = C_w(T - T_w^i). \quad (4.32)$$

This rearranges to become

$$T = \frac{C_{Al} T_{Al}^i + C_w T_w^i}{C_{Al} + C_w}. \quad (4.33)$$

We see that the final temperature T is a weighted sum of the initial temperatures. You might at first think that we must use *absolute* temperature when evaluating (4.33) numerically, such as kelvins for SI. But, in fact, that is not the case: writing each of T , T_{Al}^i , T_w^i as an offset 273.15 K plus a Celsius temperature, you'll find that the temperatures in (4.33) could just as well be interpreted as being on the Celsius scale.³

We will use the following tabulated *specific* heat capacities:

$$C_{Al}^{\text{sp}} = 900 \text{ J K}^{-1} \text{ kg}^{-1}, \quad C_w^{\text{sp}} = 4186 \text{ J K}^{-1} \text{ kg}^{-1}. \quad (4.34)$$

The required heat capacities are then

$$\begin{aligned} C_{Al} &= m_{Al} C_{Al}^{\text{sp}} = 0.1 \times 900 \text{ J/K} = 90 \text{ J/K}, \\ C_w &= m_w C_w^{\text{sp}} = 0.3 \times 4186 \text{ J/K} = 1256 \text{ J/K}. \end{aligned} \quad (4.35)$$

Insert these and Celsius-scale temperatures into (4.33), to produce the final temperature T :

³ This invariance under an offset holds generally for any linear expression with coefficients summing to one.

$$T = \frac{90 \times 100 + 1256 \times 20}{90 + 1256} \text{ } ^\circ\text{C} \simeq 25.3^\circ\text{C}. \quad (4.36)$$

The final temperature is not much higher than the water's initial temperature. This demonstrates the effect of water's abnormally high specific heat capacity, which enables it to soak up—or release—vast amounts of thermal energy without great changes in its temperature. The oceans' strong buffering to the excesses of heating and cooling lies at the heart of much of the world's weather.

By how much does the entropy of the above system increase as the water and aluminium come to thermal equilibrium? Attack the question head on (this time using a slightly different language to that used above), by writing the total entropy increase ΔS as the integrated infinitesimal increases $dS_{\text{Al}}, dS_{\text{w}}$, relating these to the time-dependent temperatures $T_{\text{Al}}, T_{\text{w}}$:

$$\Delta S = \int dS = \int (dS_{\text{Al}} + dS_{\text{w}}) = \int \left(\frac{T_{\text{Al}} dS_{\text{Al}}}{T_{\text{Al}}} + \frac{T_{\text{w}} dS_{\text{w}}}{T_{\text{w}}} \right), \quad (4.37)$$

where we've written the last fractions to introduce some analysis of energy conservation via the First Law. That is, $dE = dE_{\text{Al}} + dE_{\text{w}} = 0$, or

$$T_{\text{Al}} dS_{\text{Al}} + T_{\text{w}} dS_{\text{w}} = 0. \quad (4.38)$$

This eliminates one infinitesimal in (4.37), giving us

$$\Delta S = \int \left(\frac{T_{\text{Al}} dS_{\text{Al}}}{T_{\text{Al}}} - \frac{T_{\text{Al}} dS_{\text{Al}}}{T_{\text{w}}} \right) = \int \left(\frac{1}{T_{\text{Al}}} - \frac{1}{T_{\text{w}}} \right) T_{\text{Al}} dS_{\text{Al}}. \quad (4.39)$$

Now, (4.2) tells us that when parameter “ A ” is held constant,

$$T dS \text{ at constant } A = C_A dT. \quad (4.40)$$

Ignoring the small difference between A being pressure or volume, we can then make the replacement $T_{\text{Al}} dS_{\text{Al}} = C_{\text{Al}} dT_{\text{Al}}$. We obtain

$$\Delta S = C_{\text{Al}} \int_{T_{\text{Al}}^i}^T \left(\frac{1}{T_{\text{Al}}} - \frac{1}{T_{\text{w}}} \right) dT_{\text{Al}}. \quad (4.41)$$

To evaluate this integral, we require T_{w} as a function of T_{Al} . Consider plotting T_{w} versus T_{Al} as the water and aluminium advance toward their final common temperature T , as shown in Figure 4.2. Why does the state of the combined system follow a straight line in this figure? Suppose that the straight line is really a curve, and calculate this curve's slope $dT_{\text{w}}/dT_{\text{Al}}$ using (4.38):

$$T_{\text{Al}} dS_{\text{Al}} = -T_{\text{w}} dS_{\text{w}}, \quad \text{and so } C_{\text{Al}} dT_{\text{Al}} = -C_{\text{w}} dT_{\text{w}}. \quad (4.42)$$

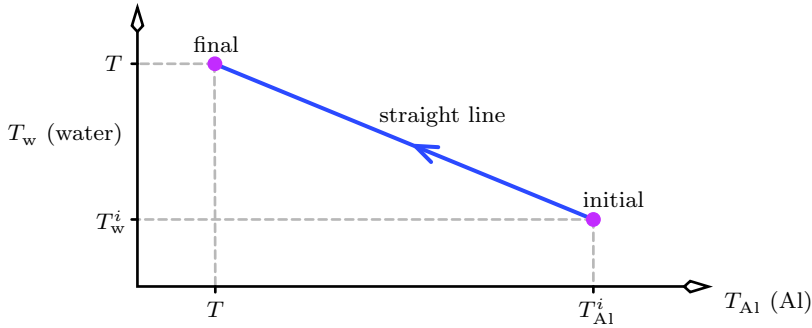


Fig. 4.2 The combined water–aluminium system evolves from temperatures (T_{Al}^i, T_w^i) to (T, T) along a straight line

Hence, with the heat capacities approximately constant over the temperature range used,

$$\text{slope of curve} = \frac{dT_w}{dT_{\text{Al}}} = \frac{-C_{\text{Al}}}{C_w} \simeq \text{constant}. \quad (4.43)$$

The curve is indeed a straight *line*, of slope $-C_{\text{Al}}/C_w$. Hence, along its entire length, we have

$$\text{slope} = \frac{\Delta T_w}{\Delta T_{\text{Al}}} = \frac{T - T_w^i}{T - T_{\text{Al}}^i} = \frac{-C_{\text{Al}}}{C_w}. \quad (4.44)$$

This is clearly equivalent to (4.32), and so rearranges to produce (4.33) once more. This allows the final temperature T to be calculated, which will be needed in (4.47).

The straight line in Figure 4.2 allows T_w to be written as a function of T_{Al} . Hence, (4.41) becomes a function of T_{Al} alone, which enables its integral to be evaluated. The equation of the line in Figure 4.2 is

$$\frac{T_w - T_w^i}{T_{\text{Al}} - T_{\text{Al}}^i} = \frac{-C_{\text{Al}}}{C_w}. \quad (4.45)$$

This rearranges to

$$T_w = \underbrace{\frac{-C_{\text{Al}}}{C_w}}_{\equiv a} T_{\text{Al}} + \underbrace{\frac{C_{\text{Al}}}{C_w} T_{\text{Al}}^i + T_w^i}_{\equiv b}. \quad (4.46)$$

Substitute this T_w into (4.41), to obtain

$$\Delta S = C_{\text{Al}} \int_{T_{\text{Al}}^i}^T \left(\frac{1}{T_{\text{Al}}} - \frac{1}{aT_{\text{Al}} + b} \right) dT_{\text{Al}} = C_{\text{Al}} \left[\ln \frac{T}{T_{\text{Al}}^i} - \frac{1}{a} \ln \frac{aT + b}{aT_{\text{Al}}^i + b} \right]. \quad (4.47)$$

Now call on the known values of a and b (but here we must take care to use kelvins, not degrees Celsius)

$$a = \frac{-C_{\text{Al}}}{C_{\text{w}}} \simeq \frac{-90}{1256} \simeq -0.0717,$$

$$b = \frac{C_{\text{Al}}}{C_{\text{w}}} T_{\text{Al}}^i + T_{\text{w}}^i \simeq \frac{90}{1256} \times 373 \text{ K} + 293 \text{ K} \simeq 319.7 \text{ K}. \quad (4.48)$$

Placing these into (4.47) gives the entropy increase $\Delta S \simeq 2.6 \text{ J/K}$.

Finally, what is the factor by which the number of accessible microstates of the entire system has increased when the aluminium and water have come to equilibrium? This is the final number of accessible microstates Ω_f divided by the initial number Ω_i , where

$$\Delta S = S_f - S_i = k \ln \Omega_f - k \ln \Omega_i = k \ln \frac{\Omega_f}{\Omega_i}. \quad (4.49)$$

Hence,

$$\begin{aligned} \frac{\Omega_f}{\Omega_i} &= e^{\Delta S/k} \simeq 10^{0.4343 \times \Delta S/k} \\ &\simeq 10^{0.4343 \times 2.6 / (1.381 \times 10^{-23})} \simeq 10^{8.2 \times 10^{22}}. \end{aligned} \quad (4.50)$$

This is stupendously large, of course, as we've now come to expect of systems that evolve toward thermal equilibrium.

4.1.1 The Third Law of Thermodynamics

In Section 2.4, we modelled the phase space available to simple isolated systems as a classical continuum over energy, which we quantised into cells whose higher-dimensional volume was set by Planck's constant h . Each cell represented a microstate, and the number Ω of these cells that was accessible to the system turned out to be expressible in terms of the system's energy. This continuum view breaks down when the system's temperature is reduced to zero, and a full quantum-mechanical approach must then be used for any analysis of the system's microstates.

Experiments indicate that the number of microstates available to the system falls as its temperature tends toward zero, and there is often just a single quantum state able to be occupied at zero temperature, or, at most, just a few such states. The system's entropy is thus either zero or slightly greater than zero. Compare this with the case of a system that has *no* quantum states available as its temperature drops to zero: its entropy will decrease to $-\infty$ in the zero-temperature limit. Additionally, experiments suggest that the rate of decrease of entropy with temperature drops to zero in the zero-temperature limit. These two observations make up the *Third Law of Thermodynamics*:

The Third Law of Thermodynamics

Regardless of its makeup or the makeup of its environment, a system's entropy has a lower bound at zero temperature that is zero or close to zero; additionally, $dS/dT \rightarrow 0$ as $T \rightarrow 0$.

Recall the discussion in Section 3.8.4, where we showed that any choice of multiple of Planck's constant h to set the cell size in phase space did not affect calculations of entropy *change*. Defining an exact number of microstates available to a system can be a rather nebulous affair, but if we invoke the Third Law to agree that this number reduces to (at least approximately) $\Omega = 1$ at $T = 0$, making $S = 0$, then we can *measure* the entropy of a non-trivial system at non-zero temperature. This is a stronger statement than simply anchoring the entropy to some finite value at $T = 0$. Consider, for a moment, the gravitational potential “ $-GM/r + \text{constant}$ ” at a distance r from a point mass M . Since only *differences* in the potential matter, we have the mathematical freedom to set the constant to be whatever is useful: so, it is universally chosen to be zero, which then anchors the potential to zero at $r = \infty$. In contrast, the Third Law is the *experimental observation* that the number of states available to a system drops to around $\Omega = 1$ as its temperature drops toward zero.

The fact that entropy decreases to a lower bound as temperature drops has a consequence for heat capacity, which we can see in the following way for the example of water ice. Let's determine ice's entropy at a temperature T_0 that is below its melting point. Recalling that the energy TdS transmitted in the heating process equals $C(T)dT$ when temperature changes [where $C(T)$ is the temperature-dependent heat capacity], the entropy at T_0 is

$$S(T_0) = \int dS = \int_0^{T_0} \frac{C(T) dT}{T}. \quad (4.51)$$

Clearly, $C(T)$ must tend toward zero as $T \rightarrow 0$, since otherwise, the integral in (4.51) would diverge to give an infinite entropy, which we know that ice with its fixed number of states Ω does not have. The Third Law thus predicts that heat capacity tends toward zero in the zero-temperature limit; and this is indeed what is observed experimentally. It means that a system very close to zero temperature needs only the tiniest thermal kick to increase its temperature by some ΔT —and for that same ΔT , the size of this required kick gets smaller and smaller as the temperature approaches absolute zero. This places a practical limit on our ability to cool any system to absolute zero.

A Measurement of Entropy Using Heat Capacity

The Third Law can be seen in action when we calculate, say, the entropy of a mole of water at 25°C and one atmosphere of pressure. The arguments that produce classical expressions for entropy, such as the Sackur–Tetrode equation in (3.146), are heavily tuned to ideal gases, and simply don’t apply to water. But we can calculate the required value by using the integral (4.51), and measuring the energy required to melt the ice.

Picture a block of one mole of ice (18 grams) being heated from absolute zero to just below its melting point, then being melted into liquid water at 0°C, and this mole of water then being heated to 25°C. The ice initially has an entropy set by the Third Law that is (to all intents and purposes) zero. The entropy acquired by the ice just before it melts is

$$\text{entropy increase to “almost-melting”} = \int_0^{273.15 \text{ K}} \frac{C_P(T) dT}{T}. \quad (4.52)$$

The ice’s entropy increases *during* melting by $\Delta S = Q/T$, where Q is the thermal energy that increases the average distance between the ice molecules without increasing their kinetic energy—and therefore without increasing their temperature. This Q is known as the latent heat of fusion L_{fusion} ; we encountered it briefly at the start of this chapter. Also, $T = 273.15 \text{ K}$, and so

$$\text{entropy increase during melting} = \frac{L_{\text{fusion}}}{273.15 \text{ K}}. \quad (4.53)$$

Now heat the icy water at constant pressure to 25°C, resulting in

$$\text{entropy increase of water} = \int_{273.15 \text{ K}}^{298.15 \text{ K}} \frac{C_P(T) dT}{T}. \quad (4.54)$$

Adding the last three equations gives us the sought-after value:

$$S = \int_0^{298.15 \text{ K}} \frac{C_P(T) dT}{T} + \frac{L_{\text{fusion}}}{273.15 \text{ K}}. \quad (4.55)$$

Table 4.1 has approximate values of the molar heat capacities of ice and water; and, of course, because we are dealing with one mole of ice/water, it’s clear that $C_P = C_P^{\text{mol}}$. We also call on the molar latent heat of fusion of ice/water, $L_{\text{fusion}}^{\text{mol}} \simeq 6010 \text{ J/K}$. The latent heat of fusion of our one mole is then, of course, $L_{\text{fusion}} \simeq 6010 \text{ J/K}$.

With these values of C_P and L_{fusion} , we write the integral in (4.55) as

$$S \simeq \int_0^{21} 0 dT + \int_{21}^{85} \frac{11.0 dT}{T} + \int_{85}^{195} \frac{21.5 dT}{T} + \int_{195}^{255} \frac{34.9 dT}{T}$$

Table 4.1 Approximate values of the molar heat capacities of ice and water

T (K)	C_P^{mol} (J/K)
$0 \rightarrow 21$	0
$21 \rightarrow 85$	11.0
$85 \rightarrow 195$	21.5
$195 \rightarrow 255$	34.9
$255 \rightarrow 273$	36.5
$273 \rightarrow 373$	75.3

$$\begin{aligned}
& + \int_{255}^{273} \frac{36.5 \, dT}{T} + \int_{273}^{298.15} \frac{75.3 \, dT}{T} + \frac{6010}{273.15} \quad \text{J/K} \\
& = 11.0 \ln \frac{85}{21} + 21.5 \ln \frac{195}{85} + 34.9 \ln \frac{255}{195} + 36.5 \ln \frac{273}{255} \\
& \quad + 75.3 \ln \frac{298.15}{273} + \frac{6010}{273.15} \quad \text{J/K} \\
& \simeq 73.7 \, \text{J/K}.
\end{aligned} \tag{4.56}$$

A more accurate laboratory value of the molar entropy of water is around 70 J/K. For brevity, our calculation has been coarse, by avoiding the use of a tediously large number of average values of C_P over smaller temperature domains. The sum in (4.56) is sensitive to the value of C_P at low temperatures, because of the $1/T$ in the integrands. But that sum is also sensitive to the value of C_P at high temperatures, since there, the integrals use larger values of C_P as a multiplier.

We could improve our calculation by further sub-dividing the temperature domain and using appropriately finer average values of C_P . Also, we have used a broad-brush value of $C_P = 0$ for all temperatures under 21 K. Care is needed when integrating over this part of the temperature domain: both C_P and T are small, making their ratio C_P/T begin to give a numerical “0/0” problem. It’s clear that a detailed knowledge of the low-temperature behaviour of ice’s heat capacity is crucial to this sort of measurement of its entropy.

On a final note, the number of states Ω available to the above mole of water at 25°C and one atmosphere of pressure is, from $S = k \ln \Omega$,

$$\begin{aligned}
\Omega &= e^{S/k} = 10^{0.4343 \times S/k} \\
&\simeq 10^{0.4343 \times 70 / (1.381 \times 10^{-23})} \simeq 10^{2.20 \times 10^{24}}.
\end{aligned} \tag{4.57}$$

4.1.2 Heat Flow and the Thermal Current Density

The concept of heat capacity is an important part of thermal and statistical physics. But along with any discussion of how well some material can “hold heat” should be a description of how the heat gets into the material in the first place. We begin such a discussion by defining the *current density* $\mathbf{J}$ of heat flow (also known as its *flux density*), where this density refers to a flow per unit *area*, not per unit *volume*. $\mathbf{J}$ is a vector field: a set of vectors defined at all points throughout the material, where each vector points in the direction of the local flow of heat. The length of $\mathbf{J}$ is the power per unit area that is crossing an infinitesimal area perpendicular to the direction in which $\mathbf{J}$ points. (In other words, $J dA$ is the power flowing through area dA that lies perpendicular to $\mathbf{J}$.) Experiments indicate that under reasonable conditions in a three-dimensional material, this heat-current density $\mathbf{J}$ is proportional to the spatial rate of loss of T (that is, $-\nabla T$) throughout the material. This connection of thermal current density to temperature gradient forms the starting point for most discussions of heat flow:

$$\mathbf{J} = -\kappa \nabla T, \quad (4.58)$$

where the proportionality constant $\kappa > 0$ is called the *thermal conductivity* of the material.

To gain a physical understanding of (4.58), we require a geometric view of the gradient of temperature, ∇T . The idea of a gradient is suggested by writing Taylor’s theorem to first order—which becomes an exact expression when we use infinitesimals. We discussed this idea in Section 1.7, but will re-iterate it here due to its importance in wide areas of physics. Given a temperature $T(\mathbf{x})$ defined at each point $\mathbf{x}$ in space, when we take an infinitesimal step $d\mathbf{x}$, the temperature that we feel increases from T to $T + dT$, where

$$dT = \nabla T \cdot d\mathbf{x}. \quad (4.59)$$

Although we have written this for a temperature field, the discussion of the geometrical view of the gradient that follows holds true for any *scalar field*, meaning any quantity that takes on a unique value at every point in space. It’s convenient to use temperature as a concrete example.

As shown in Figure 4.3, the gradient of a function T of space always points in the direction in which T is increasing most rapidly. You can see this by studying the form of $dT = \nabla T \cdot d\mathbf{x}$:

$$dT = \nabla T \cdot d\mathbf{x} = |\nabla T| |\mathbf{x}| \cos(\nabla T, d\mathbf{x}), \quad (4.60)$$

where “ $(\nabla T, d\mathbf{x})$ ” denotes the angle between ∇T and $d\mathbf{x}$. Comparing the values of dT for a set of same-length steps, all taken from the point $\mathbf{x}$ but in different directions, shows that dT is a maximum when $\cos(\nabla T, d\mathbf{x})$ is a

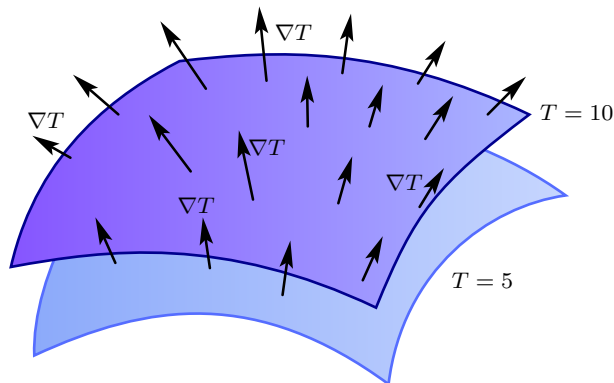


Fig. 4.3 The gradient vector of some function T is always perpendicular to the surfaces of fixed T , pointing in the direction where T is increasing most rapidly

maximum; and, of course, the cosine is a maximum when its argument, the angle between ∇T and $d\mathbf{x}$, is zero—hence, when the step $d\mathbf{x}$ is taken in the direction of ∇T . We conclude that dT is maximal along a step $d\mathbf{x}$ taken in the direction of ∇T .

Also, if we take a step $d\mathbf{x}$ within the constant- T surface, the temperature doesn't change: $dT = 0$. That means $\nabla T \cdot d\mathbf{x} = 0$, which implies that the step is perpendicular to ∇T . We conclude that at each point, ∇T points in the direction in which T increases most rapidly, and ∇T is always perpendicular to surfaces of fixed T .

Because $-\nabla T$ points in the direction in which temperature is *decreasing* most rapidly, it makes sense for $\mathbf{J}$ to point in that same direction: this is precisely what we would expect of thermal energy flow. But the fact that the heat current density $\mathbf{J}$ is actually *proportional* to $-\nabla T$ in (4.58) is an observation about Nature that can only be supplied by experiment.

From the definition of $\mathbf{J}$, the *heat current* across an area A is

$$I = \int \mathbf{J} \cdot \mathbf{n} dA = -\kappa \int \nabla T \cdot \mathbf{n} dA, \quad (4.61)$$

where the unit vector $\mathbf{n}$ is perpendicular to the infinitesimal area element⁴ dA . But remember that the increase in temperature along a small step $\mathbf{n} d\ell$ in space is $dT = \nabla T \cdot \mathbf{n} d\ell$. In other words,

$$\boxed{\frac{dT}{d\ell} \text{ in the } \mathbf{n} \text{ direction} = \nabla T \cdot \mathbf{n}.} \quad (4.62)$$

⁴ Recall that in Section 1.8, we said that dA is perhaps better called dA , but that the notation dA is quite standard.

$\nabla T \cdot \mathbf{n}$ is often called a *directional derivative*, and sometimes written as $\partial T / \partial n$. The heat current in (4.61) can now be written as

$$I = -\kappa \int \left(\frac{dT}{d\ell} \text{ along } \mathbf{n} \right) dA. \quad (4.63)$$

Now realise that the integral of any function f over a given domain is closely tied to the *mean* $\langle f \rangle$ of the function over that domain. For example, in one dimension it's always the case that, for $a < b$,

$$\begin{aligned} \int_a^b f(x) dx &= (\text{mean of } f \text{ in interval } [a, b]) \times (b - a) \\ &= \langle f \rangle \Delta x, \end{aligned} \quad (4.64)$$

where $\Delta x \equiv b - a$ is the length of the interval of integration. Likewise, the heat current in (4.63) can be written as

$$I = -\kappa \left\langle \frac{dT}{d\ell} \text{ normal to surface} \right\rangle A, \quad (4.65)$$

where $\langle \cdot \rangle$ now denotes the mean value over the surface with area A . We now have

$$\left\langle \frac{-dT}{d\ell} \text{ normal to surface} \right\rangle = \frac{I}{\kappa A}. \quad (4.66)$$

This leads to the approximation

$$\left\langle -\Delta T \text{ normal to surface} \right\rangle = \frac{I \Delta \ell}{\kappa A}. \quad (4.67)$$

In other words,

$$\left[\begin{array}{l} \text{mean temperature drop across} \\ \text{boundary of thickness } \Delta \ell \text{ and} \\ \text{area } A \end{array} \right] = I \underbrace{\frac{\Delta \ell}{\kappa A}}_{\equiv \text{thermal resistance } R}. \quad (4.68)$$

This equation is the thermal version of Ohm's rule of electric-circuit theory.⁵ There, an *electric* current I arises along a drop in electric potential Φ across an *electric* resistance R , where

$$-\Delta \Phi = IR, \quad (4.69)$$

where the drop $-\Delta \Phi$ is more usually called the “voltage drop” and written as V . The temperature drop $-\Delta T$ in (4.68) replaces the electric-potential drop $-\Delta \Phi$ in circuit theory: just as a variation in electric potential causes electric current to flow, a variation in temperature causes a thermal current

⁵ Ohm's rule is usually called Ohm's law, but it is not a law; it applies to linear elements only.

to flow. Thus, in analogy to this electric case, the R in (4.68) is called the *thermal resistance* of the material. The reciprocal of the thermal conductivity κ appears in (4.68), leading to the following terms that apply to both thermal and electrical theory:⁶

$$\text{resistivity} = \frac{1}{\text{conductivity}}; \quad \text{resistance} = \frac{1}{\text{conductance}}. \quad (4.70)$$

(In more complicated materials, these quantities become tensors, with the elements of each being written as a matrix. The resistivity matrix is then the inverse of the conductivity matrix, and ditto for the resistance and conductance matrices. See the grey box at the start of Section 8.2.) In particular, with resistivity $1/\kappa$ usually written as ϱ , the following applies to both thermal and electrical theory:

$$R = \frac{\varrho \Delta \ell}{A}. \quad (4.71)$$

We have here a correspondence between the ideas of thermal current flow and electrical current flow. Not surprisingly, when we connect thermal resistors in series or parallel to model heat flow through complex objects, we combine their resistances in the same way that we combine electrical resistances. The result that the resistance R is proportional to the resistor's length $\Delta \ell$ in (4.71) embodies the rule that “series resistances add to give the total resistance”. And R 's inverse proportionality to the cross-sectional area A embodies the rule that “parallel conductances add to give the total conductance”.

In the building trade, materials sold by thickness $\Delta \ell$ and area A are rated by their thermal conductivity κ , leaving us to calculate the resulting thermal resistance $R = \Delta \ell / (\kappa A)$. Other materials, such as slate tiles, are sold with a pre-set thickness that the customer has no control over, and these are rated by the fixed quantity $\Delta \ell / \kappa$, known as their *R-factor* R_f . Referring to (4.68), their thermal resistance is then $R = R_f / A$.

Heat Loss through a Roof

A 20 m $\times$ 10 m roof is made of 25 mm-thick pine board with thermal conductivity $\kappa = 0.11 \text{ W m}^{-1} \text{ K}^{-1}$, covered with asphalt shingles of *R-factor* $R_f = 0.0776 \text{ K m}^2 \text{ W}^{-1}$. Neglecting the overlap of the shingles, how much heat is conducted through the roof when the inside temperature is 20°C and the outside temperature is 5°C?

By “conducted heat”, we mean the heat current I , calculated in (4.68). The pine board and shingle resistors are being connected in series, so their resistances are added to give a total thermal resistance of

⁶ A *resistor* is a material object, a circuit element. It has a *resistance* determined by its size, and a *resistivity* determined by its physical make-up, which is independent of its size.

$$R = R_{\text{pine}} + R_{\text{asph}} = R_{\text{pine}} + R_{f \text{ asph}}/A. \quad (4.72)$$

Equation (4.68) becomes

$$\begin{aligned} I &= \frac{\text{mean temperature drop}}{R} = \frac{\text{mean temp. drop}}{R_{\text{pine}} + R_{f \text{ asph}}/A} \\ &= \frac{\text{mean temp. drop} \times A}{AR_{\text{pine}} + R_f(\text{asph})} = \frac{\text{mean temp. drop} \times A}{\Delta\ell/\kappa(\text{pine}) + R_f(\text{asph})} \\ &= \frac{(20 - 5) \times 20 \times 10}{0.025/0.11 + 0.0776} \text{ W} \simeq 9.8 \text{ kW}. \end{aligned}$$

We see that 9.8 kilowatts of power is being continuously lost through the roof to the cold air outside the house.

4.1.3 The Continuity Equation

Any substance whose quantity is conserved over time will satisfy the *continuity equation*. This equation describes the *local conservation* of that substance, and it appears often in the physics of flow. In our current context of heat flow, we'll illustrate the continuity equation and local conservation by using energy as the conserved substance.

In Figure 4.4, we picture a closed volume V that holds some amount of energy, and whose volumetric energy density (energy content per unit *volume*) is ϱ_E , which can vary over space. This energy is able to pass through the walls of the volume, giving rise to an areal current density (current flow per unit *area*) $\mathbf{J}$, that carries energy out of the volume across its surface, which has total area A . In a time dt then, the volume loses an amount of energy equal to $-d \int \varrho_E dV$. This energy lost is precisely that which flowed out through the closed surface—and that amount is $\oint dt \mathbf{J} \cdot \mathbf{n} dA$, where the circled integral sign $\oint$ reinforces that we are integrating over a *closed* surface:

$$\underbrace{-d \int_{\text{volume}} \varrho_E dV}_{\text{energy lost from volume in } dt} = \underbrace{\oint_{\text{surface}} dt \mathbf{J} \cdot \mathbf{n} dA}_{\text{energy flowing out through surface in } dt}. \quad (4.73)$$

At this point we call on Gauss's theorem, also known as the divergence theorem. This concerns sources and sinks: it says that the amount of current coming out of an infinitesimal volume, per unit volume, equals the diver-

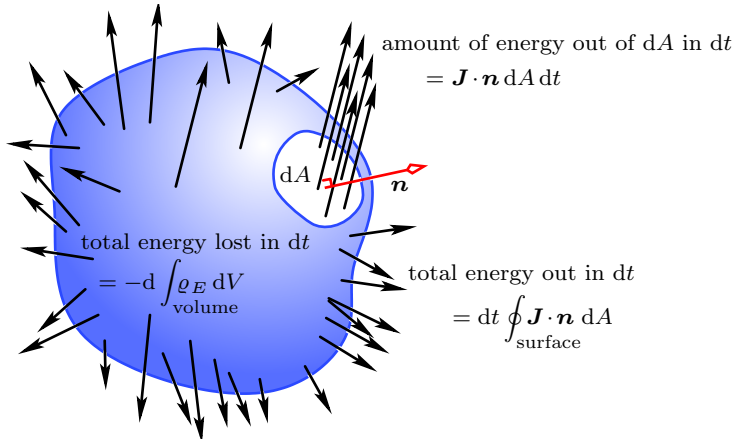


Fig. 4.4 Interpreting (4.73). Some of the energy residing within the closed volume is lost in a time dt , because it flows out through the surface. Note that ρ_E is an energy density per unit *volume*, while $\mathbf{J}$ is an energy-flow density per unit *area*, also known as a flux density or current density, being the flow of energy per unit area per unit time. These two densities form a natural pair. The energy flow out of the infinitesimal area dA is determined by the component of $\mathbf{J}$ that is parallel to $\mathbf{n}$, the unit normal vector to the surface dA

gence of the areal current density:

$$\frac{\text{total current out through surface}}{\text{infinitesimal volume enclosed}} = \nabla \cdot (\text{areal current density}). \quad (4.74)$$

The theorem converts the right-hand integral in (4.73) into an integral over volume:

$$-d \int \rho_E dV = \oint dt \mathbf{J} \cdot \mathbf{n} dA = dt \int \nabla \cdot \mathbf{J} dV. \quad (4.75)$$

Collecting terms in dV yields

$$\int dV \left[\frac{\partial \rho_E}{\partial t} + \nabla \cdot \mathbf{J} \right] = 0 \quad \text{for all volumes } V. \quad (4.76)$$

Because we are integrating over an arbitrary volume V , the bracketed term in (4.76) must be zero; and this gives rise to the *continuity equation*:

$$\boxed{\frac{\partial \rho_E}{\partial t} + \nabla \cdot \mathbf{J} = 0.} \quad (4.77)$$

This is *local* conservation of energy, meaning that the energy that disappears from within any given volume must pass through the walls of that volume. Contrast local conservation with *global conservation*, in which a substance vanishes at one point and re-appears at another, without necessarily having

crossed the space in between. Although the amount of substance in this case might well have been conserved globally, there may have been no flow across any surface in between the two points of vanishing and emergence. Global conservation is a weak type of conservation; local conservation is a much stronger concept, because it requires something to flow. Experimentally, everything conserved in the physical world is always found to be conserved locally.

It's important to remember that there are two densities present in (4.77). The energy density ϱ_E is a density over *volume* (a volumetric density), whereas the current density $\mathbf{J}$ is a density over *area* (an areal density).

On a side note, equation (4.77) adds time and space derivatives of four quantities, and, in so doing, places time and space on an equal mathematical footing. This is completely compatible with the ideas of relativity, and indeed, that subject defines a *four-current* $\vec{J}$ with the four cartesian components $(\varrho_E, J_x, J_y, J_z)$. Defining $J_t \equiv \varrho_E$ allows the four-current to be written as (J_t, J_x, J_y, J_z) , which then gives (4.77) a very symmetrical form:

$$\frac{\partial J_t}{\partial t} + \frac{\partial J_x}{\partial x} + \frac{\partial J_y}{\partial y} + \frac{\partial J_z}{\partial z} = 0. \quad (4.78)$$

Examining such *four-vectors* as $\vec{J}$ is a core topic in relativity theory, because their components transform between the frames of relatively moving observers identically to the way in which time and space coordinates transform between such observers.⁷ We'll touch on this subject again in Section 6.10.

4.1.4 The Heat Equation, or Diffusion Equation

The above discussion of heat flow allows us to determine how the temperature distribution in a hot material evolves over time. We know this temperature distribution in the material at some initial time, but we have no knowledge of the energy density ϱ_E and the current density $\mathbf{J}$ in the continuity equation (4.77). So, we wish to replace ϱ_E and $\mathbf{J}$ with appropriate expressions that involve the temperature T .

Begin with (4.58)'s empirical expression $\mathbf{J} = -\kappa \nabla T$. Place this into the continuity equation (4.77), to produce

$$\frac{\partial \varrho_E}{\partial t} - \kappa \nabla^2 T = 0. \quad (4.79)$$

⁷ In relativity, the components of four-vectors can be written with the coordinates placed either as subscripts or superscripts, depending on the choice of basis vectors. In particular, the subscripts t, x, y, z in (4.78) are normally written as superscripts.

To eliminate ϱ_E , refer to (4.29) to write the increase in thermal energy dE in the volume as $mC^{\text{sp}} dT$, where m is the mass contained within the volume. Then divide $dE = mC^{\text{sp}} dT$ by the volume, yielding

$$d\varrho_E = \varrho_m C^{\text{sp}} dT, \quad (4.80)$$

where ϱ_m is the mass density throughout the volume. Thus,

$$\frac{\partial \varrho_E}{\partial t} = \varrho_m C^{\text{sp}} \frac{\partial T}{\partial t}. \quad (4.81)$$

Now substitute this expression for $\partial \varrho_E / \partial t$ into (4.79), arriving at

$$\boxed{\nabla^2 T = \frac{\varrho_m C^{\text{sp}}}{\kappa} \frac{\partial T}{\partial t}}. \quad (4.82)$$

This is the *heat equation*, or *diffusion equation*. It has been produced by combining the continuity equation (a general principle of physics) with (4.58): the experimental observation that the heat current density is proportional to the spatial rate of temperature loss. Besides its application to temperature, the heat equation describes the diffusion of particles more generally.

Solving the Heat Equation

Let's bundle the various constants in the heat equation into one positive constant K called the *diffusion constant*, to write (4.82) as

$$\nabla^2 T = \frac{1}{K} \frac{\partial T}{\partial t} \quad (K > 0). \quad (4.83)$$

It's worth noting here that the dimensions of K are length²/time, and that the dimensions of T are not important to (4.83): they can be anything at all. We'll make use of this information shortly.

Aside from the fact that we derived (4.83) by appealing to concepts of heat flow, that equation might well be expected to model the flow of heat even if we have never seen that derivation. The reason for this rests on the idea that $\nabla^2 T$ is a sum of second spatial derivatives:

$$\nabla^2 T = \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}. \quad (4.84)$$

Consider for a moment, a single spatial dimension x , and note that a function $T(x)$ with a negative second derivative $T''(x)$ in some region is shaped concave down in that region; a classic example is $T(x) = -x^2$, whose second derivative is -2 everywhere. The same is true in three spatial dimensions for a function $T(x, y, z)$: if all of its second spatial derivatives are negative in some region

of space (so that $\nabla^2 T$ is negative there too), then T is peaked in that region: if T is temperature, then that region of space is a “hot spot”. So, when no heat sources are present and T is peaked around a hot spot, we know that $\nabla^2 T < 0$ in that region; and thus (4.83) implies that $\partial T/\partial t$ is negative there too. In other words, the temperature in and around a hot spot decreases with time, just as you would expect.

Similarly, when T is a trough (a cold spot), its second spatial derivatives are all positive. This implies that $\nabla^2 T$ is positive, and thus so is $\partial T/\partial t$: hence, the temperature in and around a cold spot increases with time. Again, this behaviour is just what we expect of temperature.

To set about solving the heat equation (4.83), we begin with the basic observation that it is linear, by which is meant that any linear combination of its solutions is also a solution. This is easy to prove, using the fact that ∇^2 and $\partial/\partial t$ are themselves linear operators. To do so, suppose that T_1 and T_2 are solutions to (4.83). Then, substituting a linear combination of them, $T = aT_1 + bT_2$, into the left-hand side of (4.83) results in

$$\begin{aligned}\nabla^2(aT_1 + bT_2) &= a\nabla^2 T_1 + b\nabla^2 T_2 \\ &= \frac{a}{K} \frac{\partial T_1}{\partial t} + \frac{b}{K} \frac{\partial T_2}{\partial t} = \frac{1}{K} \frac{\partial}{\partial t}(aT_1 + bT_2),\end{aligned}\quad (4.85)$$

and this last expression is the right-hand side of (4.83). So, $T = aT_1 + bT_2$ satisfies (4.83) too. Linearity plays an important role in the study of differential equations, due to its ability to generate new solutions from known solutions.

A huge amount of literature is devoted to solving partial differential equations in which the *laplacian operator*, ∇^2 , acting on some function is set proportional either to that function, or else to its first or second partial derivative with respect to time. The topic is normally covered in detail in applied maths courses; and so we’ll be content to consider here just one approach to calculating the flow of heat over a particularly simple domain. After being set in motion by the relevant initial conditions, the flow of heat throughout any domain is influenced by what is happening on the domain’s boundary. Because that boundary complicates the analysis of the heat equation beyond where we want to go here, we’ll treat the simple case of an infinite domain: no boundary conditions need then be considered. On this domain, one solution to the heat equation (4.83) is

$$T_{\mathbf{x}'}(t, \mathbf{x}) = t^{-3/2} \exp \frac{-|\mathbf{x} - \mathbf{x}'|^2}{4Kt}, \quad (4.86)$$

where $\mathbf{x} \equiv (x, y, z)$, and $\mathbf{x}' \equiv (x', y', z')$ is some arbitrary point in space. Different values of $\mathbf{x}'$ give different solutions, and so we have singled $\mathbf{x}'$ out by making it a subscript in (4.86). The proof that $T_{\mathbf{x}'}(t, \mathbf{x})$ satisfies (4.83) is easy to construct (and so is omitted here), by calculating its second space

derivatives and first time derivative, making use of

$$|\mathbf{x} - \mathbf{x}'|^2 = (x - x')^2 + (y - y')^2 + (z - z')^2. \quad (4.87)$$

Observe that $T_{\mathbf{x}'}(t, \mathbf{x})$ does not have dimensions of temperature; but that's okay: see the comment just after (4.83) above. It is a kind of template for a solution, and we'll use it shortly to construct a more realistic solution—one that does have dimensions of temperature.

$T_{\mathbf{x}'}(t, \mathbf{x})$ is a gaussian, peaking at $\mathbf{x}'$ and symmetrical about that point in each of the x, y, z directions. It corresponds to a hot spot, a localisation of high temperature around $\mathbf{x}'$. Comparing it to (1.119), we see that its characteristic widths on each of the axes are $\sigma_x, \sigma_y, \sigma_z$, where

$$2\sigma_x^2 = 2\sigma_y^2 = 2\sigma_z^2 = 4Kt. \quad (4.88)$$

It follows that these widths are $\sigma_x = \sigma_y = \sigma_z = \sqrt{2Kt}$. It's clear from this that the hot spot spreads out as time passes, as expected. As the spot spreads, its strength diminishes as per the $t^{-3/2}$ factor in (4.86). Its infinite extent⁸ reflects the lack of a boundary in this simple scenario.

One fact about our template solution $T_{\mathbf{x}'}(t, \mathbf{x})$ is very important to note: its integral over all space is constant throughout time. To see this, refer to (4.86), and use the discussion in Section 1.5 to show that

$$\iiint_{-\infty}^{\infty} t^{-3/2} \exp \frac{-|\mathbf{x} - \mathbf{x}'|^2}{4Kt} dx dy dz = (4\pi K)^{3/2}. \quad (4.89)$$

[Alternatively, you can apply (1.123).] Because K has no time dependence, it follows that as this gaussian hot spot (4.86) spreads out over time, its integral over all space is conserved. We can use this fact to answer the following question: what happens when we attempt to trace this template solution (4.86) back in time to $t = 0$? Its height grows without bound and its width shrinks to zero; in other words, it becomes a spike, proportional to a delta function $\delta(\mathbf{x} - \mathbf{x}')$. Now recall that any multiple of the template solution is also a solution to the heat equation (4.83), because that equation is linear! So, normalise $T_{\mathbf{x}'}(t, \mathbf{x})$ by dividing the right-hand side of (4.86) by $(4\pi K)^{3/2}$. Then, when evolved back to $t = 0$, the normalised $T_{\mathbf{x}'}(t, \mathbf{x})$ becomes *exactly* $\delta(\mathbf{x} - \mathbf{x}')$. We'll recycle the notation $T_{\mathbf{x}'}(t, \mathbf{x})$, to write the normalised solution as

$$T_{\mathbf{x}'}(t, \mathbf{x}) = \frac{1}{(4\pi Kt)^{3/2}} \exp \frac{-|\mathbf{x} - \mathbf{x}'|^2}{4Kt}, \quad (4.90)$$

⁸ The part of a function's domain on which it is non-zero is often called the function's "support".

because the normalised solution is what is really important here, since it evolves backward in time to become the very simple $\delta(\mathbf{x} - \mathbf{x}')$.

We can now use the above discussion to build more general solutions to the heat equation. Suppose we start out with an infinitesimally small hot spot at $\mathbf{x}'$, meaning a temperature distribution given by

$$T_{\mathbf{x}'}(0, \mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}'), \quad (4.91)$$

where, for now, we are not concerned about the dimensions of temperature. Nonetheless, this hot spot has infinite temperature; nonphysical for sure, but mathematically useful for the reasoning that follows. Over time, the spot will spread out and lose strength until, at time t , it has the form of (4.90).

Next, consider a more realistic scenario in which the initial temperature distribution $T(0, \mathbf{x})$ [no subscript $\mathbf{x}'$ appears here] is not necessarily a delta function, but is certainly known—and *does* have the correct dimensions of temperature! This initial temperature distribution can *always* be written as a linear combination of delta functions:

$$T(0, \mathbf{x}) = \int_{-\infty}^{\infty} T(0, \mathbf{x}') \delta(\mathbf{x} - \mathbf{x}') d^3x'. \quad (4.92)$$

Now allow each of these delta functions to evolve from 0 to t , meaning

$$\delta(\mathbf{x} - \mathbf{x}') \text{ evolves to } \frac{1}{(4\pi Kt)^{3/2}} \exp \frac{-|\mathbf{x} - \mathbf{x}'|^2}{4Kt}. \quad (4.93)$$

Then, the initial temperature distribution (4.92) will evolve as the *same* linear combination of those evolving functions, to become

$$T(t, \mathbf{x}) = \int_{-\infty}^{\infty} T(0, \mathbf{x}') \underbrace{\frac{1}{(4\pi Kt)^{3/2}} \exp \frac{-|\mathbf{x} - \mathbf{x}'|^2}{4Kt}}_{\text{Green function for heat equation}} d^3x'. \quad (4.94)$$

This is the general solution of the heat equation for our no-boundary scenario. Provided we can perform this integral (perhaps numerically), any given temperature distribution can be propagated forward in time. The evolved version of the delta function in (4.93) and (4.94) is called the *Green function* for the heat equation.⁹ Notice that because this Green function has dimensions of $1/\text{length}^3$, the dimensions of $T(t, \mathbf{x})$ are the same as the dimensions of $T(0, \mathbf{x}')$; these are, say, kelvins in the SI system.

Equation (4.94) is an example of the *convolution* of two functions. This concept is explored more easily in one dimension, where the convolution “ $*$ ” of two functions $f(x)$ and $g(x)$ is defined as

⁹ It is very often called “a Green’s function” by practitioners who would *never* say “a Mozart’s concerto”.

$$f(x) * g(x) \equiv \int_{-\infty}^{\infty} f(x') g(x - x') dx' = g(x) * f(x). \quad (4.95)$$

It can be shown with little effort (but that subject lies outside this text) that convolving two functions is nothing more than the procedure of using one as a “weighted moving mean” to smoothen the other: $f(x) * g(x)$ is the result of smoothening $f(x)$ with $g(x)$, or smoothening $g(x)$ with $f(x)$, since convolution is commutative.¹⁰ In the language of convolution, the evolution of a given temperature distribution (4.94) is written as

$$T(t, \mathbf{x}) = T(0, \mathbf{x}) * \frac{1}{(4\pi Kt)^{3/2}} \exp \frac{-|\mathbf{x}|^2}{4Kt}. \quad (4.96)$$

Observe that all reference to $\mathbf{x}'$ has now disappeared: it is the dummy variable inside the convolution integral (4.94). At each moment in time, the convolution acts to spread the initial temperature distribution $T(0, \mathbf{x})$ out, with the (gaussian) spread becoming wider and wider as time goes on. The temperature throughout space is slowly evening out.

Let’s give an example of these ideas in one dimension. Suppose the initial temperature distribution is

$$T(0, x) = 1 \text{ K}, \quad \text{for } 0 \leq x \leq 1, \quad (4.97)$$

and zero elsewhere. This is plotted as the “top hat” in Figure 4.5. A one-dimensional version of the analysis above produces the one-dimensional version of (4.96), where the exponent $3/2$ that arose from three space dimensions becomes $1/2$ for one space dimension:

$$\begin{aligned} T(t, x) &= T(0, x) * \frac{1}{\sqrt{4\pi Kt}} \exp \frac{-x^2}{4Kt} \quad [\text{the 1-dim. version of (4.96)}] \\ &= \int_{-\infty}^{\infty} T(0, x') \frac{1}{\sqrt{4\pi Kt}} \exp \frac{-(x - x')^2}{4Kt} dx' \\ &= \frac{1}{\sqrt{4\pi Kt}} \int_0^1 \exp \frac{-(x - x')^2}{4Kt} dx'. \end{aligned} \quad (4.98)$$

Use a change of variables $u = x - x'$ here to simplify the integral; as x' advances from 0 to 1, u advances from x to $x - 1$:

¹⁰ I use the verbs “smoothen” and “smoothening” instead of the more commonly used “smooth” and “smoothing”. Consider that we all speak of whitening and softening a fabric, straightening a cloth, lengthening a speech, sharpening a pencil, a darkening sky, a reddening sunset, and many other similar words. If we whiten a wall with paint (make it white), then perhaps we’ll smoothen it first (make it smooth). Despite this, the use of “smoothing” stubbornly persists in the field of signal processing.

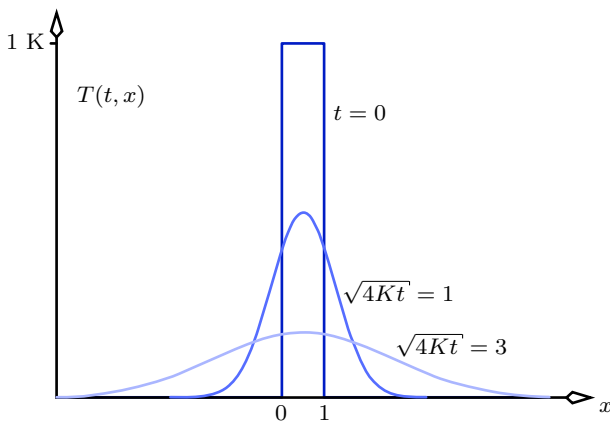


Fig. 4.5 Plots of $T(t, x)$ in (4.99) at three representative times

$$\begin{aligned}
 T(t, x) &= \frac{1}{\sqrt{4\pi Kt}} \int_x^{x-1} \exp \frac{-u^2}{4Kt} \times -du = \frac{1}{\sqrt{4\pi Kt}} \int_{x-1}^x \exp \frac{-u^2}{4Kt} du \\
 &\stackrel{(1.89)}{=} \frac{1}{2} \left[\operatorname{erf} \frac{x}{\sqrt{4Kt}} - \operatorname{erf} \frac{x-1}{\sqrt{4Kt}} \right]. \tag{4.99}
 \end{aligned}$$

Figure 4.5 plots $T(t, x)$ at three representative times. At $t = 0$, the hot area is confined and spread uniformly within the interval from $x = 0$ to 1. At later times, the heat spreads out and that initially hot area cools. This behaviour is, of course, exactly what we expect of a real hot spot.

We have convolved continuous functions here via integration, and the result has been a “movie” of the time evolution of a hot spot. In practice, evolving highly complicated temperature distributions in the real world becomes a fully numerical task, in which the convolutions that were integrals of products in (4.95) are approximated by sums of products on a grid that we set up in space. Evaluating these discrete convolutions lies, in fact, at the very core of modern *signal processing*, without which such diverse subjects as radar, image recognition, and modern approaches to high-speed exact computing would be impossible. Summing thousands of products possibly a great many times per second would tax any computer, were it not for a theorem that relates convolution (or, more precisely, a variety of it called *circular convolution*) to the discrete Fourier transform. The necessary discrete Fourier transforms would be no faster to perform; but the situation is rescued by modern algorithms (that were actually first discovered by Gauss!), collectively called the *fast Fourier transform*, that enable the calculations of the transform to be done very quickly and efficiently. That is another subject entirely; but we see here how the core mathematical ideas of our modern digital world can be related back to the simple notion of an evolving hot spot.

4.2 The Second Term: Mechanical Interaction

We turn next to a discussion of the mechanical interaction, the second term in the First Law. This is a subject treated in great depth in books on thermodynamics, and so we'll not spend a great deal of time on it.

4.2.1 Heat Engines and Reversibility

Much of the mechanical energy that drives mechanical processes is eventually dissipated thermally, although indeed, some is radiated away by the process discussed in Chapter 9. An early Holy Grail of thermodynamics was the idea that a process might perhaps be found that converts this discarded “heat energy” into mechanical work with 100% efficiency. Early work in this area focussed on the pressure–volume term of the First Law, and gave us the *heat engine*. A core requirement of a heat engine is that it should be re-usable (unlike, say, a rocket engine). This suggests that it should operate in a cycle.

Harnessing the power of heat engines started the Industrial Revolution in the eighteenth century, and later studies of the efficiencies that could be expected of such engines formed the core of the young subject of thermodynamics. The key concept here is *reversibility*, discussed previously in Section 3.8.3. We call a process reversible if it can (at least in principle) be made to run in reverse by our making only an infinitesimal change to it. For a process to be reversible:

1. No work must be done by dissipatory processes such as friction, since attempting to run the system in reverse would then require the dissipatory processes to work in reverse, which they are never observed to do. (That is, if sliding a brick forward on a concrete surface creates friction, then sliding it backward on that same surface will also create friction; at no time does the process of friction run in reverse.)
2. There must be no heat conduction across a non-infinitesimal temperature difference, since running the system in reverse would then require energy to flow thermally from a cold reservoir to a hot one, which—while consistent with the laws of physics—is “ruled out” by the Second Law of Thermodynamics.¹¹
3. The process must be quasi-static, which ensures the system is always arbitrarily close to equilibrium. In fact, it's not at all clear how to take a non-equilibrium system and run it in reverse. But this requirement isn't

¹¹ Recall Section 3.8, where we said that the Second Law is not really a “law”: energy is perfectly allowed and able to flow spontaneously from a cold body to a hot body. But, as we showed when introducing temperature in Section 3.5, the process is so improbable as to be *completely* discounted in practice, which is as good as a law.

onerous: we saw, in Section 2.1, that even the burning of fuel in a car's piston engine is as good as quasi-static.

The central theorem for the work performed by heat engines and refrigerators (which can be viewed as heat engines running in reverse) was established in the early nineteenth century by the French engineer Sadi Carnot:

A reversible engine working in a cycle between two heat reservoirs is always more efficient at doing work than an irreversible engine that works between those same two heat reservoirs.

A “proof by contradiction” of Carnot’s theorem is a standard exercise found in thermodynamics textbooks. We start by assuming the theorem to be false: that we have an irreversible engine whose efficiency is greater than that of a given reversible engine. We then connect the reversible engine (running it in reverse) to the irreversible one, and show that after one cycle, the only transfer of energy that has occurred has been from the cold reservoir to the hot reservoir. But this reduction in entropy of the universe is, to all intents and purposes, inconsistent with the Second Law; and *that* is the sought-after contradiction that proves Carnot’s theorem.

Such discussions of heat engines traditionally make strong use of the Second Law. The statement of the Second Law as we know it, “the entropy of a closed system will, to all intents and purposes, never decrease”, is specific to statistical mechanics, but the law was originally couched in the language of thermodynamics. The 1851 version credited to Kelvin states:

No heat engine (reversible or irreversible) working in a cycle can take in heat from its surroundings and convert *all* of that heat into work.

The 1850 version, credited to Clausius, is concerned with “heat pumps”, meaning refrigerators:

No heat pump (reversible or irreversible) working in a cycle can transfer heat from a cold reservoir to a hot reservoir without external work being done.

Of course, the refrigerator that keeps our food from spoiling performs external work to transfer heat from a cold reservoir to a hot reservoir: it has a motor that must exhaust some heat to the environment.

The Carnot Cycle

Recall from Section 3.4.1, that the work we do *on* a gas by changing its volume¹² is $-PdV$, and so the work done *by* the gas is $+PdV$. Thus, when

¹² Note, with respect to the discussion of Section 1.6, that the use of the word “changing” is completely appropriate here.

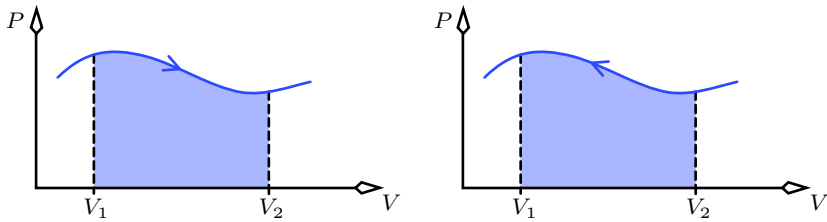


Fig. 4.6 **Left:** When a gas expands from volume V_1 to V_2 , it does work equal to the blue area, as shown in (4.101). **Right:** When a gas contracts from volume V_2 to V_1 , it does work equal to minus the blue area, as shown in (4.102)

a gas's volume is changed from some initial volume V_i to some final volume V_f , it does an amount of work:

$$\text{work done by gas} = \int_{V_i}^{V_f} P dV, \quad (4.100)$$

irrespective of whether the gas expands or contracts (that is, the integral could be positive or negative). With this integral in mind, refer to the pressure–volume diagram in Figure 4.6. The left-hand picture shows a gas expanding from volume V_1 to volume V_2 . The work done *by* the gas is

$$\text{work done by gas} = \int_{V_1}^{V_2} P dV = \text{blue area under curve}. \quad (4.101)$$

The right-hand picture shows the gas contracting from volume V_2 to volume V_1 . The work done *by* the gas is

$$\text{work done by gas} = \int_{V_2}^{V_1} P dV = - \int_{V_1}^{V_2} P dV = \text{minus blue area under curve}. \quad (4.102)$$

Now combine two such paths, in Figure 4.7. The work done by the gas in

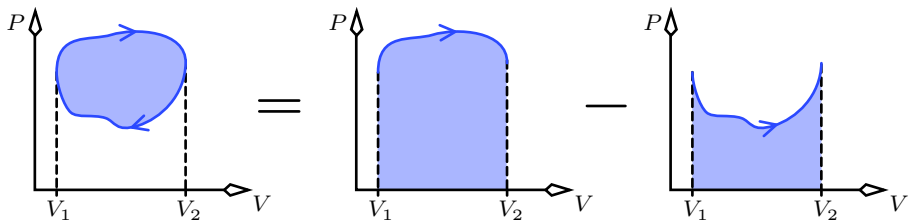


Fig. 4.7 These pictures refer to (4.103). The left-hand blue area equals the middle blue area minus the right-hand blue area, and each of these areas can be expressed in terms of the work done by the gas as it follows the blue curves. Note the reversed arrow in the right-hand picture

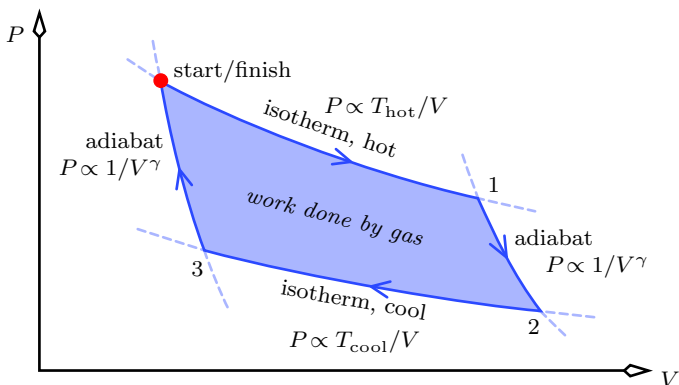


Fig. 4.8 The Carnot cycle follows the pressure of an ideal gas as a function of its volume. It starts and finishes at the red disk. First, it follows an isotherm, then an adiabat, then another isotherm, and finally another adiabat

going around the closed cycle *clockwise* in the left-hand picture in the figure can be written as

$$\begin{aligned}
 \left[\begin{array}{c} \text{work done by} \\ \text{gas in cycle} \end{array} \right] &= \underbrace{\int_{V_1}^{V_2} P dV}_{\text{upper curve}} + \underbrace{\int_{V_2}^{V_1} P dV}_{\text{lower curve}} = \underbrace{\int_{V_1}^{V_2} P dV}_{\text{upper curve}} - \underbrace{\int_{V_1}^{V_2} P dV}_{\text{lower curve}} \\
 &= \left[\begin{array}{c} \text{blue area in} \\ \text{middle picture} \end{array} \right] - \left[\begin{array}{c} \text{blue area in} \\ \text{right-hand picture} \end{array} \right] \\
 &= \text{blue area in left-hand picture.}
 \end{aligned} \tag{4.103}$$

Similarly, when the gas follows a closed cycle *counter-clockwise*, it does work equal to *minus* the area enclosed by the cycle on the PV diagram.

In particular, the *Carnot cycle* is a *reversible* process that constructs this closed curve in the following way, as shown in Figure 4.8. Its working substance is a container of ideal gas that is brought into contact with two heat sources in succession, at pressures and volumes determined partly by the gas's current temperature and its container size. The first heat source, the hotter of the two, makes the gas expand at almost the same temperature T_{hot} as the heat source: the gas must be infinitesimally cooler for heat to be transferred to it, but not more than infinitesimally cooler, for otherwise, the heat conduction at such a non-infinitesimal temperature difference would be irreversible. The expansion of the gas at constant temperature T_{hot} follows the top *isotherm* (a constant-temperature curve, $P \propto 1/V$) in Figure 4.8.

To continue building what will eventually be a closed curve on the PV diagram in Figure 4.8 (so that the gas will end up doing useful work), we construct another isotherm at a cooler temperature, formed by placing the gas in contact with a cooler reservoir—or what is better termed a *heat sink*.

We require the state of the gas to move from right to left along this bottom isotherm in Figure 4.8. For this to happen, we must first arrange for this state to go from corner 1 to corner 2 in the figure.

We do this by allowing the gas to expand further along the right-hand curve that connects corner 1 to corner 2. This curve is called an *adiabat*, and is the path of successive states occupied by the gas as it expands further *adiabatically*, meaning without thermal transfer of energy: it exchanges no heat with its environment.¹³ So the gas cools, but without exhausting any heat, and certainly without exhausting any heat at a non-infinitesimal temperature difference to the cooler reservoir on the bottom isotherm—meaning that, again, this process of moving along the adiabat is reversible. We will show shortly that an adiabat is expressed as $P \propto 1/V^\gamma$, where $\gamma \equiv C_P/C_V$ was defined in (4.22). Since $\gamma > 1$ [see (4.24)], adiabats do indeed slope down more steeply than isotherms, as drawn in Figure 4.8.

Because a system undergoing an adiabatic process exchanges no heat with its environment ($\delta Q = T dS = 0$), no entropy change occurs during this process. We are usually interested in systems with a fixed particle number N , and so for these, the First Law becomes $dE + P dV = 0$. In particular for an ideal gas, we can replace dE with $C_V dT$, as we saw in (4.6): although that equation was derived holding volume fixed, it must hold for *any* process involving an ideal gas, because the energy E of such a gas is a fixed function of its temperature T . So, write

$$C_V dT + P dV = 0. \quad (4.104)$$

Even under extreme conditions, most gases can be treated as ideal; so bring in $T = PV/(Nk)$, to write

$$C_V \frac{d(PV)}{Nk} + P dV = 0. \quad (4.105)$$

Multiply both sides by Nk , and use the product rule of differentiation to obtain

$$C_V V dP + (C_V + Nk) P dV = 0. \quad (4.106)$$

Now apply (4.14), to find

$$C_V V dP + C_P P dV = 0. \quad (4.107)$$

Dividing through by PVC_V gives us

$$\frac{dP}{P} + \frac{C_P}{C_V} \frac{dV}{V} = 0. \quad (4.108)$$

¹³ Because no heat transfer occurs on an adiabat, and everything is quasi-static, the relation “ $\delta Q = T dS$ ” says that entropy is constant. Hence, an adiabat is also called an *isentrope*.

Keeping the comments of Section 1.9.2 in mind, and noting that the parameter $\gamma = C_P/C_V$ is approximately independent of pressure, we integrate (4.108) to obtain

$$\ln P + \gamma \ln V = \text{constant}. \quad (4.109)$$

In other words,

$PV^\gamma = \text{constant for adiabatic processes at fixed } N.$

(4.110)

This last expression describes the adiabats in Figure 4.8. After the expanding gas has reached the bottom of the right-hand adiabat at corner 2 and starts moving left along the bottom isotherm, it exhausts heat to the cool reservoir at temperature T_{cool} across an infinitesimal temperature difference (and thus reversibly), causing the gas to contract as it follows the isotherm. Finally, it reaches corner 3 and gets compressed as it moves up the left adiabat, heating up as it does so until its temperature equals the heat source's T_{hot} . It has now completed the cycle to arrive at its starting values of pressure and volume, after which the process begins anew. The work done by the gas equals the area within the closed curve in Figure 4.8.

As the gas followed the top isotherm, it absorbed energy thermally and expanded ($T_{\text{hot}} dS = \text{“heat into gas”} > 0$), and so the entropy of the world grew. That entropy remained constant along the right-hand adiabat when no heat was exchanged with the heat sources ($T dS = 0$). Then, that entropy decreased along the bottom isotherm as the gas expelled heat while contracting ($T_{\text{cool}} dS = \text{“heat into gas”} < 0$); and finally, that entropy remained constant along the left-hand adiabat. The overall entropy change turns out to be zero for *any* reversible cyclic process.

The efficiency of a heat engine is defined as the ratio of the work done to the heat absorbed. The Carnot cycle is a reversible cyclic process, and is the gold standard against which all other (irreversible) heat engines are compared to calculate their efficiency, which can never be larger than that of the Carnot cycle. The efficiency of the Carnot cycle turns out to be $1 - T_{\text{cool}}/T_{\text{hot}}$, and so it can never be 100% efficient. But it's as efficient as any heat engine can ever be that works between these two heat reservoirs: all real heat engines working between the same two heat reservoirs will have a lower efficiency than this. Although the 100% efficient Holy Grail of heat engines cannot be realised, the Carnot cycle tells us what *is* attainable in practice.

4.2.2 The Joule–Thomson Process

Why does the air escaping from a bicycle tyre cool as it emerges? The cooling implies that the gas molecules lose kinetic energy as they escape; but where has this energy gone? In particular, as air escapes from a bicycle tyre, the

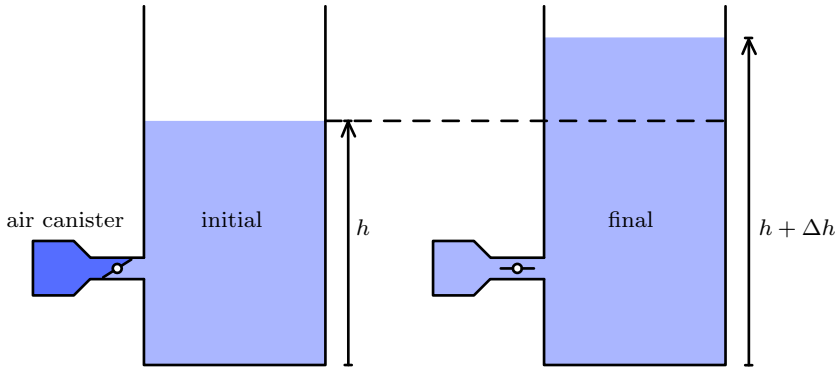


Fig. 4.9 Left: A canister of compressed air is held closed, connected to a vertical air column of height h that is held together under its own weight. **Right:** The can is opened and releases its air into the air column. After the can's interior has come down to atmospheric pressure, the vertical air column has been “pumped up” to a new height $h + \Delta h$

valve grows cold. This suggests that the temperature drop occurs before the escaping air has a real chance to mix with the larger volume of air in the room.

But suppose something different for a moment: that the valve doesn't grow cold, and that instead, some cooling occurs based on the idea that the emerging air does work by “pumping Earth's atmosphere up” ever so slightly. To analyse such a process, refer to Figure 4.9, which shows a canister of compressed air before and after being emptied into a vertical air column of initial height h . The column's height increases by Δh as a result. The air emerging from the canister must do work to increase the height of the large air column. Does doing such work account for the observed cooling of the air that escapes from the canister?

Suppose N molecules of the canister air escape into the height- h air column, increasing its height by Δh . Since we are really modelling our atmosphere here, we'll make the simplifying assumption that Δh is so small as to be ignorable, which means the following analysis does not apply to a very narrow column of air. We'll also assume that the air column's density is independent of its height. Comparing the left- and right-hand sides of Figure 4.9, we see that the escaping air does work to lift the N molecules up to a height h :

$$\begin{aligned} \left[\begin{array}{l} \text{work done by } N \text{ molecules} \\ \text{of escaping air} \end{array} \right] &= \left[\begin{array}{l} \text{potential energy given to } N \\ \text{molecules (each of mass } m) \\ \text{lifted through height } h \end{array} \right] \\ &= Nmgh, \end{aligned} \quad (4.111)$$

where g is the acceleration due to Earth's gravity. As the escaping air molecules mix with the air in the column, some kinetic energy of the col-

umn's air must be converted to this potential energy $Nmgh$. We'll model this process by saying that some larger number N_0 of air molecules lose this kinetic energy. For example, if the escaping air mixes thoroughly with the air column so that ten times as many molecules end up being disturbed, then $N_0 = 10N$. Modelling air as an ideal gas with ν quadratic energy terms per particle, the N_0 air molecules will have kinetic energy $\nu N_0 kT/2$. Their loss in kinetic energy is then

$$-\Delta(\nu N_0 kT/2) = \nu N_0 k/2 \times -\Delta T. \quad (4.112)$$

Equating this to $Nmgh$ gives the drop in air temperature as

$$-\Delta T = \frac{Nmgh}{\nu N_0 k/2} = \frac{NN_A mgh}{\nu N_0 N_A k/2} = \frac{2NM_{\text{mol}} gh}{\nu N_0 R}, \quad (4.113)$$

where M_{mol} is air's molar mass and R is the gas constant. The task is to estimate N_0 , the total number of air molecules that lose kinetic energy. Set $N_0 = \alpha N$ for some constant α . Assuming $h = 8$ km and $\nu = 5$ for air with $M_{\text{mol}} = 29.0$ grams, (4.113) becomes, using SI units,

$$-\Delta T = \frac{2M_{\text{mol}} gh}{\nu \alpha R} = \frac{2 \times 0.0290 \times 9.8 \times 8000}{5\alpha \times 8.314} \text{ K} \simeq \frac{109}{\alpha} \text{ K}. \quad (4.114)$$

If the escaping air produces a total of ten times as many molecules being disturbed, then $\alpha = 10$ and we might expect a drop of around 10 K—assuming our model of the atmosphere is valid. If the escaping air produces 100 times as many disturbed air molecules, then $\alpha = 100$ and we can expect a drop of around 1 K in this model. Our atmosphere's density is not, in fact, constant with height: a much better model has it dropping exponentially with height. If we add some molecules to it, the whole atmosphere becomes heavier and settles a little, and modelling the N added molecules as effectively being lifted through a height h becomes too simplistic; they must really be modelled as being lifted through varying heights. So, a first attempt at a better analysis might use a smaller value of h that embodies the exponential drop in density. We might replace the $h = 8$ km in (4.114) with, say, $h = 5$ km, changing (4.114) into $-\Delta T \simeq 68/\alpha$. Now, we are predicting a temperature drop of, at most, a few kelvins. And in practice, since it takes time for a change in density to propagate upward, perhaps an effective value of h is only tens or hundreds of metres, which points to a temperature drop of a fraction of a kelvin.

Nevertheless, these predicted values express a temperature drop of a certain *volume* of air, whereas in practice, it is the *valve* that is observed to cool. Perhaps kinetic energy is being lost just as the air expands through the valve, suggesting that our model of “pumping up the atmosphere” is deficient in some way.

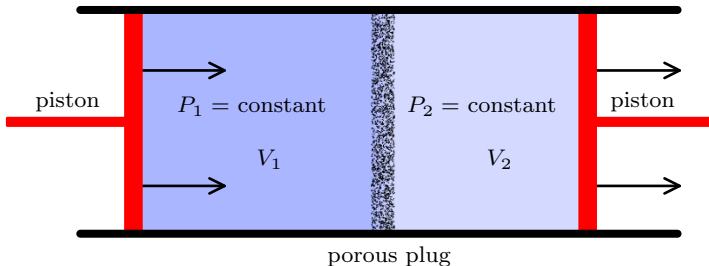


Fig. 4.10 The Joule–Thomson process. A piston forces gas slowly at constant pressure P_1 from chamber 1 through a porous plug into chamber 2. The piston in chamber 2 is pulled out slowly to maintain a constant pressure P_2 in chamber 2

This suggests that we model the escaping air as a *Joule–Thomson process*. This idea is shown in Figure 4.10. The left-hand piston moves slowly to the right, forcing the gas in chamber 1 through a porous plug into chamber 2 while maintaining a constant pressure P_1 in chamber 1 throughout the process. The piston in chamber 2 is pulled slowly to the right to maintain a constant pressure P_2 in chamber 2. Consider that the work we do on the pistons equals the increase in the system’s total energy E :

$$-P_1 dV_1 - P_2 dV_2 = dE_1 + dE_2. \quad (4.115)$$

Rearranging gives us

$$-dE_1 - P_1 dV_1 = dE_2 + P_2 dV_2. \quad (4.116)$$

Now recall the enthalpy $H = E + PV$ from (3.200); so, at the constant pressure in each chamber,

$$dH = dE + PdV. \quad (4.117)$$

Equation (4.116) then becomes

$$-dH_1 = dH_2. \quad (4.118)$$

We conclude that the enthalpy lost from chamber 1 equals the enthalpy gained by chamber 2: the Joule–Thomson process occurs at constant enthalpy.

Figure 4.10 has a non-infinitesimal drop in pressure across the plug. Now suppose we make P_1 and P_2 differ only infinitesimally: $P_2 = P_1 + dP$. Envisage assembling a quasi-static version of the scenario by chaining together an infinite number of chambers and porous plugs with an infinitesimal pressure drop across each, resulting in a non-infinitesimal pressure drop from start to end. (In that case, each intermediate piston will need to act from the side of its chamber instead of from its end.) The *Joule–Thomson coefficient* μ_{JT} specifies the way in which temperature changes with pressure across each infinitesimal porous plug:

$$\mu_{JT} \equiv \left(\frac{\partial T}{\partial P} \right)_H \stackrel{(3.229)}{=} \frac{-(\partial H / \partial P)_T}{(\partial H / \partial T)_P} \stackrel{(4.11)}{=} \frac{-(\partial H / \partial P)_T}{C_P}. \quad (4.119)$$

Next, what is $(\partial H / \partial P)_T$ in terms of quantities that can easily be measured? Return to the enthalpy $H = E + PV$ to write, for the pressure drop across each plug,

$$\begin{aligned} dH &= dE + VdP + PdV \\ &= TdS - \cancel{PdV} + VdP + \cancel{PdV} \\ &= TdS + VdP. \end{aligned} \quad (4.120)$$

This, of course, differs from (4.117), because pressure is not held fixed across the plugs. It follows from (4.120) that

$$\left(\frac{\partial H}{\partial P} \right)_T = T \left(\frac{\partial S}{\partial P} \right)_T + V. \quad (4.121)$$

Now call in the Gibbs energy increase across each plug, (3.210) with $dN = 0$ (because each particle that enters a plug also exits the plug):

$$dG = -SdT + VdP. \quad (4.122)$$

This says, on recalling (3.222), that

$$-\left(\frac{\partial S}{\partial P} \right)_T = \left(\frac{\partial V}{\partial T} \right)_P. \quad (4.123)$$

Now substitute this into (4.121), to find

$$\left(\frac{\partial H}{\partial P} \right)_T = -T \left(\frac{\partial V}{\partial T} \right)_P + V. \quad (4.124)$$

Finally, this last expression enables (4.119) to be written in terms of easily measured quantities:

$$\mu_{JT} = \frac{T(\partial V / \partial T)_P - V}{C_P}. \quad (4.125)$$

We will focus on calculating $(\partial V / \partial T)_P$ for a *real* gas—say, one described by van der Waals' equation (3.123). Implicitly differentiate that equation with respect to T at constant P and N (and, as usual, we'll omit the subscript N from our partial derivatives):

$$\frac{-2N^2a}{V^3} \left(\frac{\partial V}{\partial T} \right)_P (V - Nb) + \left(P + \frac{N^2a}{V^2} \right) \left(\frac{\partial V}{\partial T} \right)_P = Nk. \quad (4.126)$$

Now solve this for $(\partial V / \partial T)_P$:

$$\left(\frac{\partial V}{\partial T}\right)_P = \frac{Nk}{-N^2a/V^2 + 2N^3ab/V^3 + P}. \quad (4.127)$$

Equation (4.125) takes this to become

$$\mu_{JT}C_P = \frac{V(NkTV^2 + N^2aV - 2N^3ab - PV^3)}{-N^2aV + 2N^3ab + PV^3}. \quad (4.128)$$

None of the terms in (4.128) are necessarily vastly smaller than the others for all regimes of temperature and pressure, and so we must retain them all. What does (4.128) say for an ideal gas? Such a gas has $a = b = 0$, and (4.127) becomes

$$\left(\frac{\partial V}{\partial T}\right)_P = \frac{V}{T}. \quad (4.129)$$

This, of course, also follows more easily from the ideal-gas law—but using (4.127) to calculate it was a good check on that equation. We can now say that for an ideal gas,

$$\mu_{JT} \stackrel{(4.125)}{=} \frac{TV/T - V}{C_P} = 0. \quad (4.130)$$

Thus, the Joule–Thomson process predicts *no* temperature drop in an ideal gas being forced across the porous plug.

When the Joule–Thomson coefficient $\mu_{JT} \equiv (\partial T/\partial P)_H$ is positive, then because the pressure drops as the gas exits the porous plug, its temperature must drop too. Likewise, when μ_{JT} is negative, the gas must be *warmer* as it exits the plug. If we could determine μ_{JT} for each possible pair of temperature–pressure values, we would find a set of (T, P) pairs for which $\mu_{JT} = 0$: this set would be a boundary of points on the TP plane, dividing that plane into regions for which the exiting gas warmed and for which it cooled. We wish to find this set of (T, P) points.

So, set $\mu_{JT} = 0$: that is, set the numerator of (4.128) to zero, and assume the denominator does not also go to zero at the same time:

$$NkTV^2 + N^2aV - 2N^3ab - PV^3 = 0. \quad (4.131)$$

Now incorporate van der Waals' equation (3.123). Multiplying the terms in (3.123) together produces the expanded form

$$PV = \frac{-N^2a}{V} + PNb + \frac{N^3ab}{V^2} + NkT. \quad (4.132)$$

Substitute this into the last term of (4.131). A little simplification yields

$$2NaV - 3N^2ab - PbV^2 = 0. \quad (4.133)$$

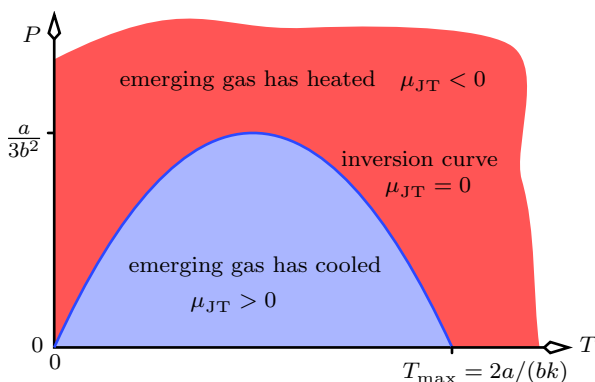


Fig. 4.11 The dark blue parabola is a plot of (4.134). As shown in the text, this “inversion curve” separates the TP plane into two regions: blue, where μ_{JT} is positive (gas emerging from the porous plug has cooled down) and red (the rest of the plane), where μ_{JT} is negative (gas emerging from the porous plug has heated up)

We’ve accounted for our gas not being ideal by incorporating van der Waals’ equation; but we are required to work with T rather than V . We could eliminate V from (4.133) by noting that (4.132) is a cubic polynomial in V , and so can be solved exactly for V ; but a far easier approximation uses the ideal-gas law to eliminate V from (4.133). In that case, substitute $V = NkT/P$ into (4.133). The result is

$$P = \frac{kT(2a - bkT)}{3ab}. \quad (4.134)$$

That is, $\mu_{JT} = 0$ holds at the (T, P) values given by (4.134). A plot of P versus T is the inverted parabola in Figure 4.11. The parabola divides the TP plane into two regions, in each of which μ_{JT} is either everywhere positive or everywhere negative. Find this sign in the region below the parabola by calculating μ_{JT} at, say, the point $(T, P) = (a/(bk), 0)$. This is a limit point, since we can’t really arrange for the pressure to be zero. In this limit, the ideal-gas expression $V = NkT/P$ says that $V \rightarrow \infty$. Referring to (4.128), we see that with $PV \approx NkT$,

$$\mu_{JT}C_P \rightarrow \frac{V(NkTV^2 + N^2aV - NkTV^2)}{-N^2aV + NkTV^2} \simeq \frac{Na}{kT} > 0. \quad (4.135)$$

Hence, $\mu_{JT} > 0$ holds below the parabola. For the region above the parabola, choose the test point $(T, P) = (a/(bk), \infty)$. In this regime of infinite pressure, van der Waals’ equation (3.123) says $V \rightarrow Nb$. Now substitute these values $P \rightarrow \infty$, $V = Nb$, $kT = a/b$ into (4.128):

$$\mu_{JT}C_P \rightarrow \frac{V \times -PV^3}{PV^3} < 0. \quad (4.136)$$

Thus, $\mu_{JT} < 0$ holds above the parabola. Figure 4.11 shows these regions of temperature–pressure where μ_{JT} is positive, zero, and negative. The fact that μ_{JT} changes sign across the parabola has given rise to the name “inversion curve” for the parabola.

The idea here is that when the gas is forced through the porous plug with a given initial pair of values of temperature and pressure, these values define a point in the TP plane in Figure 4.11. When that point lies in the red region (where $\mu_{JT} < 0$), the gas will be hotter as it exits the plug; when the point lies on the parabola itself ($\mu_{JT} = 0$), the gas’s temperature doesn’t change; and when the point lies in the blue region ($\mu_{JT} > 0$), the emerging gas will have cooled. For any given temperature between the two roots of the parabola, the gas will emerge heated if the pressure is high enough; else it will cool.

Focus on the right-hand root of the parabola, $2a/(bk)$: on entering the plug, if the gas is hotter than this *maximum inversion temperature* $T_{\max} = 2a/(bk)$, it defines a point in the red region in Figure 4.11 irrespective of the pressure, and so must emerge hotter than it was on entry. On the other hand, if, on entering the plug, the gas is cooler than the maximum inversion temperature *and* its pressure is low enough [as determined by (4.134)], it will emerge cooler than it was on entry. This means that the Joule–Thomson process can be used to liquefy gases.

How big is the blue region in Figure 4.11? Its extent is set by the maximum inversion temperature $2a/(bk)$, and the pressure corresponding to the peak of the parabola, $a/(3b^2)$. We will calculate these values for some representative gases.

Consider first CO_2 . We require the van der Waals parameters a and b , which are often tabulated in their molar form that is specific to the molar form of van der Waals’ equation, (3.133). Refer to (3.134):

$$a = a_{\text{mol}}/N_A^2, \quad b = b_{\text{mol}}/N_A. \quad (4.137)$$

For CO_2 , standard tabulated values are

$$a_{\text{mol}} \simeq 3.592 \ell^2 \text{ atm/mol}^2, \quad b_{\text{mol}} \simeq 0.04267 \ell/\text{mol}, \quad (4.138)$$

where these units of litres (ℓ) and atmospheres (atm) are commonly used:

$$\begin{aligned} 1 \ell &\equiv 1000 \text{ cm}^3 = 1000 \times (10^{-2} \text{ m})^3 = 10^{-3} \text{ m}^3, \\ 1 \text{ atm} &\equiv 101,325 \text{ Pa}. \end{aligned} \quad (4.139)$$

Recall that we can omit the “mol” from (4.138), following the discussion in Section 1.9.1. The maximum inversion temperature $T_{\max}$ is

$$T_{\max}(\text{CO}_2) = \frac{2a}{bk} \stackrel{(4.137)}{=} \frac{2a_{\text{mol}}}{b_{\text{mol}}R} \stackrel{(4.138)}{=} \frac{2 \times 3.592 \ell^2 \text{ atm}}{0.04267 \ell \times R}$$

$$= \frac{2 \times 3.592 \ell \text{ atm}}{0.04267 R} = \frac{2 \times 3.592 \times 10^{-3} \times 101,325}{0.04267 \times 8.314} \text{ K} \simeq 2050 \text{ K}. \quad (4.140)$$

Similarly, the pressure at the parabola's peak is

$$\text{pressure at peak (CO}_2\text{)} = \frac{a}{3b^2} \stackrel{(4.137)}{=} \frac{a_{\text{mol}}}{3b_{\text{mol}}^2} = \frac{3.592}{3 \times 0.04267^2} \text{ atm} \simeq 658 \text{ atm}. \quad (4.141)$$

The blue region in Figure 4.11 clearly has a considerable extent. For temperatures only a little lower than $T_{\text{max}} = 2050 \text{ K}$, cooling will be possible but difficult—unless the pressure is relatively low; whereas, at a temperature of $T_{\text{max}}/2 = 1025 \text{ K}$, cooling will occur for any pressure up to 658 atmospheres.

Values of T_{max} and the pressure at the parabola's peak for various gases are listed in Table 4.2. The agreement between predicted and measured value of T_{max} is satisfactory, and we might suppose that significant departures (such as for CO_2) show that the van der Waals model and perhaps the Joule–Thomson analysis are not necessarily suitable for *all* gases.

Of particular interest is helium's very low maximum inversion temperature of around 40 kelvins, due to a very low value of a_{mol} . This low value of a_{mol} expresses the exceptionally weak long-range attraction that helium atoms have for each other, as discussed in Section 3.7. It follows that helium is difficult to liquefy; one approach is to pre-cool it using liquid hydrogen. Hydrogen's maximum inversion temperature is also comparatively low. In the open air at room temperature, forcing hydrogen gas through a porous plug can cause the gas to combust spontaneously.

What gives gases their ability to be either heated *or* cooled in a Joule–Thomson process? When pressures are high (specifically, above the parabola in Figure 4.11), the gas molecules are essentially being forced through the porous plug close enough together that they feel a mutual hard-core repulsion, like a gas of ball bearings. When they emerge from the plug, this repulsion forces them to fly apart like a jack-in-the-box, and this increase in their speeds manifests as a temperature increase. On the other hand, when pressures are

Table 4.2 Values of the molar van der Waals parameters a_{mol} , b_{mol} for various gases, with predicted and measured values of the maximum inversion temperature T_{max} , and predicted values of the pressure at the peak of the inversion curve

	CO ₂	N ₂	He	H ₂
a_{mol} ($\ell^2 \text{ atm}$)	3.592	1.390	0.0346	0.2444
b_{mol} (ℓ)	0.04267	0.03913	0.0238	0.02611
T_{max} (predicted)	2050 K	866 K	35 K	228 K
T_{max} (measured)	1500 K	620 K	~ 40 K	200 K
pressure at peak (predicted)	658 atm	303 atm	20 atm	119 atm

low, the gas molecules' mutual interactions are dominated by their long-range attraction to each other. As they emerge from the plug, they move farther apart, converting some of their kinetic energy into the potential energy of this long-range attractive force. This means they slow down, which manifests as a drop in temperature.

4.3 The Third Term: Diffusive Interaction

The diffusive interactions of particles that are ubiquitous to chemistry can be usefully treated and understood using the chemical potential μ . In the examples that follow, we'll stay with the physics language as used in previous chapters; but note that chemists tend to use a somewhat different language of statistical mechanics than that of physicists.

Chemical reactions involve particles moving toward lower chemical potentials, so we wish to derive an expression for μ that can be used in what follows. But before doing that, an easy entry into the subject begins by examining the behaviours of pressure and density in simplified models of the atmosphere and ocean. We will then rederive the same core equations by developing an expression for the chemical potential μ of an ideal gas. The success of this approach then suggests invoking the chemical potential in other areas of chemistry.

4.3.1 Pressure and Density of the Atmosphere

The variations of pressure and density in our atmosphere and ocean are commonly analysed with the same approach that we used in Section 3.15 to peer into a star's interior. We focus on a small element of, say, air at height z , and apply Newton's second law to the forces on it.

The scenario is shown in Figure 4.12. The tiny element of matter in the figure calls in the language of infinitesimals. Whenever you see an infinitesimal, remember that, like Δ , it denotes "final minus initial", and so you should be aware of what is "initial" and what is "final". The forces on the sides of the small element are equal and opposite, and thus cancel each other out. Hence, it's sufficient to analyse a small element of air by examining only the forces on its top and bottom faces, mentally beginning with one and moving to the other. Which face we choose to start with is immaterial, but our arbitrary choice defines this face as "initial". To make this point, both choices of initial face are coloured red in the figure, and are allocated height z . The final face defines height $z + dz$. Either choice in the figure is equally valid, so it's a worthwhile exercise to run the calculation for each.

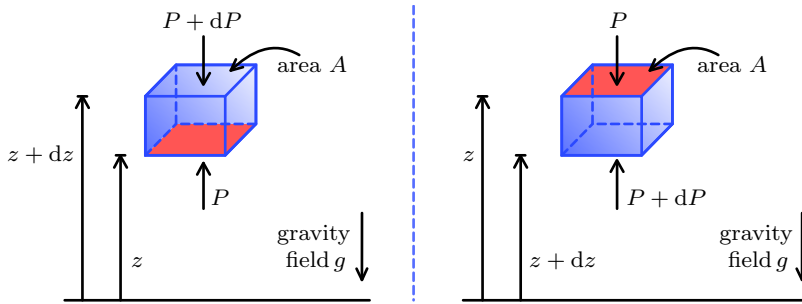


Fig. 4.12 Examining the vertical forces on an element of air. Remember that both of the above conventions for labelling the height z and pressure P are valid; we need only ensure that “ z ” and “ P ” are allocated to the “initial” choice of face. This initial choice is the red bottom face at left in the figure, and it is the red top face at right

Consider, then, a small element of air with flat top and bottom, and begin with the left-hand choice in Figure 4.12. The pressure difference between top and bottom is due to gravity’s pull on the element. The element’s mass is the product of its mass density ϱ and its volume. Its volume equals its top area times its infinitesimal height (meaning its vertical extent now, not its height above $z = 0$), and is, of course, a positive number. From the figure, we see that this infinitesimal height is $z + dz - z = dz$; so, the element’s volume is $A dz$. The upward and downward forces on the bottom plane must balance, since the element doesn’t accelerate. Thus, with g being the local gravity field,

$$\underbrace{(P + dP)A + \varrho A dz g}_{\text{downward force}} = \underbrace{PA}_{\text{upward force}}. \quad (4.142)$$

Equation (4.142) simplifies to

$$dP + \varrho g dz = 0. \quad (4.143)$$

We must now relate the pressure P to the mass density ϱ , which itself relates to the *particle number density* ν via

$$\varrho = \frac{\text{particle mass} \times \text{number of particles}}{\text{volume}} = m\nu, \quad (4.144)$$

where m is the mass of one particle. We are modelling air as an ideal gas, in which case

$$P = \frac{NkT}{V} = \nu kT. \quad (4.145)$$

Equation (4.144) then becomes

$$\varrho = \frac{mP}{kT}, \quad (4.146)$$

and so (4.143) becomes

$$\frac{dP}{P} = \frac{-mg dz}{kT}. \quad (4.147)$$

We will assume a constant temperature T in the atmosphere. (This is very roughly true, but we'll relax that assumption in Section 6.5.) This simplification enables (4.147) to integrate to

$$P(z) = P_0 \exp \frac{-mgz}{kT}, \quad (4.148)$$

where, from now on, we denote a variable's value at $z = 0$ (say, at sea level) by a subscript 0 for conciseness [e.g., $P(0)$ is written as P_0]. It follows from (4.144)–(4.148) that

$$\frac{\varrho(z)}{\varrho_0} = \frac{\nu(z)}{\nu_0} = \frac{P(z)}{P_0} = \exp \frac{-mgz}{kT}. \quad (4.149)$$

Chemists prefer using the molar mass $M_{\text{mol}} \equiv N_A m$ rather than the particle mass m , and the gas constant $R \equiv N_A k$ instead of Boltzmann's constant k . The physicist's common expression m/k then becomes M_{mol}/R , and (4.149) is written as

$$\frac{\varrho(z)}{\varrho_0} = \frac{\nu(z)}{\nu_0} = \frac{P(z)}{P_0} = \exp \frac{-M_{\text{mol}} gz}{RT}. \quad (4.150)$$

Suppose that we prefer the right-hand convention for “initial” and “final” heights in Figure 4.12. The infinitesimal height of the element is then $z - (z + dz) = -dz$, making its volume $-A dz$. Equation (4.142) is replaced by

$$\underbrace{PA + \varrho \times -A dz g}_{\text{downward force}} = \underbrace{(P + dP)A}_{\text{upward force}}. \quad (4.151)$$

This reduces to (4.143); so, with the rest of the analysis unchanged, we arrive at (4.149) once more.

We can gain insight into this exponential decrease-with-height of mass and particle densities and pressure, by determining the common height $z_{1/2}$ at which each of these quantities has dropped to half of its sea-level ($z = 0$) value. Using the particle number density ν , (4.149) becomes

$$\frac{\nu(z_{1/2})}{\nu_0} = \exp \frac{-mgz_{1/2}}{kT} \equiv \frac{1}{2}. \quad (4.152)$$

Solve this for $z_{1/2}$, obtaining

$$z_{1/2} = \frac{kT \ln 2}{mg} = \frac{RT \ln 2}{M_{\text{mol}} g}. \quad (4.153)$$

Suppose the atmosphere has a constant temperature of 5°C and a molar mass of $M_{\text{mol}} = 29.0 \text{ g}$. Equation (4.153) becomes

$$z_{1/2} = \frac{8.314 \times 278 \times 0.693}{0.0290 \times 9.8} \text{ m} \simeq 5.6 \text{ km.} \quad (4.154)$$

Note too that

$$\begin{aligned} \exp \frac{-mgz}{kT} &= \exp \frac{-mgz \ln 2}{kT \ln 2} = 2^{\frac{-mgz}{kT \ln 2}} \\ &= 2^{\frac{-z}{kT \ln 2 / (mg)}} \stackrel{(4.153)}{=} 2^{-z/z_{1/2}}. \end{aligned} \quad (4.155)$$

This enables (4.149) to be written as

$$\frac{\varrho(z)}{\varrho_0} = \frac{\nu(z)}{\nu_0} = \frac{P(z)}{P_0} \stackrel{(4.154)}{=} 2^{-z/(5.6 \text{ km})}. \quad (4.156)$$

In reality, the temperature of our atmosphere is not constant with height, because great masses of air are churned with a Coriolis acceleration arising from Earth's rotation. Even so, our estimate of the characteristic “half height” as 5.6 km agrees reasonably well with measurements: for all heights up to about 9 km (which just includes the height of Mount Everest), the pressure of the real atmosphere departs, at most, by about 3% from an exponential fall-off with a half height of 5.35 km. Above 15 km, the density and pressure begin to depart appreciably from an exponential fall-off. But most of the atmosphere lies below this height anyway.

How much does this atmosphere weigh? The standard value of air pressure at sea level tells us the answer. This is 101,325 N/m², which means that a column of atmosphere with a footprint of one square metre weighs 101,325 N. Dividing this weight by gravity's $g = 9.81 \text{ m/s}^2$ yields a mass of the column of about 10,330 kg, or a little over 10 tonnes. We don't feel this weight because our bodies are composed mostly of water, and thus are relatively incompressible; and also because air is a fluid, so its pressure evens out to push from all sides, including from below and within our lungs. If that air pressure were to be removed from one side of an object, that extreme weight would make itself felt on the other side. This, of course, is the principle behind the suction cup. When we pull on a suction cup, we are trying to lift a section of atmosphere of cross section equal to the cup's area, and that is generally a huge weight. If the rubber edge of the cup bends a little and a small amount of air sneaks underneath it, that air exerts its own pressure in the same direction that we are pulling on the cup, helping us to lift the cup further with great ease. (That is, the cup loses its suction.) It follows that the heaviest object liftable by a perfectly manufactured suction cup of area n square metres is about $10n$ tonnes.

The appearance of the gravitational potential energy of one particle, mgz , in the exponential decrease (4.149) is an example of the appearance of potential energy for *any* conservative force in the atmosphere. Without loss of generality, focus on the left-hand picture in Figure 4.12, and replace the

gravity field with a general conservative force field that has a potential energy per particle of U . (We discount non-conservative forces. If the force were not conservative, particles would gain energy each time they followed, say, a closed loop in the atmosphere. The atmosphere would then never reach equilibrium.) The force on a particle is thus $-dU/dz$. Now repeat the analysis we applied around that picture. The total force on the block of air is zero; hence

$$PA - (P + dP)A + \left[\frac{\text{force per}}{\text{particle}} \right] \times \left[\frac{\text{number of particles}}{\text{in block}} \right] = 0. \quad (4.157)$$

With air modelled as an ideal gas, we have

$$\text{number of particles in block} = \frac{PV}{kT} = \frac{PA dz}{kT}. \quad (4.158)$$

Equation (4.157) becomes

$$-dPA + \frac{-dU}{dz} \times \frac{PA dz}{kT} = 0. \quad (4.159)$$

This rearranges to

$$\frac{dP}{P} = \frac{-dU}{kT}, \quad (4.160)$$

which integrates to yield

$$P(z) = P_0 \exp \frac{-U(z)}{kT}. \quad (4.161)$$

For the case of gravity (where $U = mgz$), this matches (4.149). We see that (4.161) is a more general expression than (4.149), which applies only to a gravity field.

4.3.2 Pressure and Density of the Ocean

The above analysis of the forces on an infinitesimal element of air applies equally well to the analysis of pressure and density in the ocean. Here, it's more natural to measure depth z positive downward, and an argument analogous to that which produced (4.143) yields (unsurprisingly, since z is really just “minus height”)

$$dP - \varrho g dz = 0. \quad (4.162)$$

As with the discussion of air just after (4.143), we must relate pressure P to mass density ϱ . Water is very difficult to compress—we'll see how difficult soon—so in a first analysis, its density ϱ can be taken as independent of depth z , and thus equal to the density ϱ_0 at the ocean surface. Equa-

tion (4.162) then integrates to

$$P(z) = P_0 + \varrho_0 g z. \quad (4.163)$$

We see, from either (4.162) or (4.163), that underwater pressure increases uniformly with depth; and indeed, for every 10 metres we descend, the pressure increases by

$$\Delta P = \varrho_0 g \Delta z = 1000 \times 9.8 \times 10 \text{ pascals} \simeq 1 \text{ atmosphere}. \quad (4.164)$$

This is a rule of thumb known to all divers, for as they descend into the sea, their breathing apparatus must pump more air into each breath to keep their lungs inflated. This higher demand on their air tank reduces the total time they can spend under water. Also, the pressure of the large amount of air that they breathe in forces some of its nitrogen into their blood. When they ascend to lower pressures, this nitrogen begins to bubble out of the blood in a process akin to the fizzing that occurs when we take the top off a lemonade bottle, causing a dangerous situation known as *the bends*.

Scuba divers have their lungs inflated “for free” by the pressure of compressed air that the scuba apparatus provides. In contrast, for snorkellers without scuba gear, the increased pressure due to the weight of water above the diver makes it difficult to breathe through a snorkel at any real depth. Snorkellers effectively must lift the water above them by expanding their chest muscles, and chest muscles are unable to support much weight. But even a diver with very strong chest muscles will have problems breathing when more than half a metre or so down, because the necessarily longer snorkel retains exhausted CO_2 that never quite gets the chance to escape, and so ends up being inhaled again.

An interesting exercise is to calculate the slight increase in water density due to the extreme pressure at depth, although we will simplify matters by assuming that both the ocean’s temperature and the strength of gravity remain constant with depth—which are not entirely accurate assumptions! So, take ϱ to be a function of depth z , and now assume water’s bulk modulus B to be independent of depth. Equation (1.138) says

$$B \equiv \frac{dP}{-dV/V} = \frac{\text{infinitesimal pressure increase}}{\text{infinitesimal relative volume loss}}. \quad (4.165)$$

Experiments show that an infinitesimal pressure increase causes an infinitesimal fractional volume loss in a ratio that is constant over a wide range of pressures. Water’s bulk modulus is $B = 2 \text{ GPa}$. To gain a feel for this number, note that (4.165) says that to compress the volume of a block of water by 1% (which is thus the relative volume loss, or $-\Delta V/V$), whatever pressure the block is currently under must be increased by $\Delta P \simeq B \times -\Delta V/V = B \times 1\%$,

or 20 MPa. This is around 200 atmospheres of pressure increase, equivalent to applying the weight of a 200 kg mass to each square centimetre.¹⁴

To apply the concept of bulk modulus to calculate ϱ , imagine taking some water from a depth z , where it occupies volume V , to a new depth $z + dz$, where it occupies $V + dV$. The mass density at depth z is ϱ , and at the new depth, it is

$$\begin{aligned} \varrho + d\varrho &= \frac{\text{mass}}{\text{new volume}} = \frac{\varrho V}{V + dV} = \frac{\varrho}{1 + dV/V} = \varrho \left(1 - \frac{dV}{V} \right) \\ &\stackrel{(4.165)}{=} \varrho \left(1 + \frac{dP}{B} \right). \end{aligned} \quad (4.166)$$

It follows that $d\varrho/\varrho = dP/B$. If we assume that the bulk modulus is independent of depth, then this integrates to

$$\varrho = \varrho_0 e^{(P-P_0)/B}, \quad (4.167)$$

where, as usual, ϱ_0 and P_0 are the density and pressure at the surface ($z = 0$). Now substitute this expression for density into (4.162), and integrate the result to arrive at

$$P(z) = P_0 - B \ln \left(1 - \frac{\varrho_0 g z}{B} \right). \quad (4.168)$$

Substitute this $P(z)$ into (4.167), to obtain

$$\varrho(z) = \varrho_0 \left(1 - \frac{\varrho_0 g z}{B} \right)^{-1}. \quad (4.169)$$

Limiting Cases of Pressure and Density at Titanic's Depth

As a check on (4.168), note that when $\varrho_0 g z / B \ll 1$ (the “small- z /large- B limit”), (4.168) becomes

$$P(z) \simeq P_0 - B \times \frac{-\varrho_0 g z}{B} = P_0 + \varrho_0 g z. \quad (4.170)$$

This agrees with the constant-density approximation (4.163). Similarly, the small- z /large- B limit of (4.169) is

$$\varrho(z) \simeq \varrho_0 \left(1 + \frac{\varrho_0 g z}{B} \right) = \varrho_0 + \frac{\varrho_0^2 g z}{B}. \quad (4.171)$$

¹⁴ Gases, of course, are much more compressible than liquids: it is relatively easy to compress the air in a cylinder by a large amount. But when the air isn't confined in this way, it is not easily compressed. Even in the design of large commercial jet aircraft, the air flowing over the aircraft is, in fact, modelled as an incompressible fluid.

Let's calculate the pressure and water density at the wreck of Titanic, about $z = 4.0$ km below the ocean surface. This is the small- z /large- B limit, because

$$\frac{\varrho_0 g z}{B} \simeq \frac{1000 \times 9.8 \times 4000}{2 \times 10^9} \simeq 0.02 \ll 1. \quad (4.172)$$

Equation (4.170) then produces (with $P_0 = 1$ atmosphere)

$$\begin{aligned} P(4.0 \text{ km}) &= \left(1 + \frac{\varrho_0 g z}{1 \text{ atm}}\right) \text{ atm} = \left(1 + \frac{1000 \times 9.8 \times 4000}{101,325}\right) \text{ atm} \\ &\simeq 388 \text{ atm}. \end{aligned} \quad (4.173)$$

This matches what we expect from simply calculating the weight of a cylinder of water 4 kilometres high with a cross section of 1 m^2 . This column has a volume of 4000 m^3 , and thus a mass of 4,000,000 kg. Multiplying this by $g \approx 10 \text{ m/s}^2$ gives a weight of about 40 meganewtons. This weight (along with that of the atmosphere) presses down on the column's bottom area of 1 m^2 at Titanic's depth. The pressure of the water alone is thus 40 MPa, while that of the atmosphere is a negligible 0.1 MPa. Hence, the total pressure is about 400 atmospheres, which agrees with (4.173) up to the level of accuracy used here.

Equation (4.171) gives us (with $\varrho_0 = 1000 \text{ kg/m}^3$)

$$\begin{aligned} \varrho(4.0 \text{ km}) &\simeq \left(1000 + \frac{1000^2 \times 9.8 \times 4000}{2 \times 10^9}\right) \text{ kg/m}^3 \\ &\simeq 1020 \text{ kg/m}^3. \end{aligned} \quad (4.174)$$

The value of g is actually slightly less than 9.8 m/s^2 at this depth; but the point is that water is extremely incompressible, and it supports a column of water stretching four kilometres above it with little change to its density.

The expressions for density and pressure calculated in the last few pages are summarised in Table 4.3. Because we refer to both height and depth in the table, and wish to use positive numbers for each to agree with our everyday use of those words, we replace “ z for height” with H in that table. That is, setting $H = 5 \text{ m}$ in the table means a height of 5 metres in the atmosphere. We also replace “ z for depth” with D , so that $D = 5 \text{ m}$ in the table means a depth of 5 metres in the ocean.

Table 4.3 Expressions for mass density and pressure as functions of height H in the atmosphere and depth D in the ocean. H and D are always positive. These expressions are present in the text as (4.146), (4.148), (4.163), (4.167)–(4.171)

	Mass Density ϱ	Pressure P
Ideal gas:	$\frac{mP}{kT}$	$P_0 \exp \frac{-mgH}{kT}$
Incompressible water:	constant ϱ_0	$P_0 + \varrho_0 gD$
Compressible water:	$\left\{ \begin{array}{l} \varrho_0 \left(1 - \frac{\varrho_0 gD}{B} \right)^{-1} \\ = \varrho_0 e^{(P-P_0)/B} \end{array} \right\}$	$P_0 - B \ln \left(1 - \frac{\varrho_0 gD}{B} \right)$
Compressible water (small- D /large- B):	$\varrho_0 + \frac{\varrho_0^2 gD}{B}$	$P_0 + \varrho_0 gD$

4.3.3 Pressure and Density from the Chemical Potential

The above pressure/density analysis of Figure 4.12 is standard, and requires no knowledge of statistical mechanics. A far less well-known alternative approach makes use of the chemical potential of an ideal gas. We'll investigate that here as a first use of the chemical potential—a quantity that remains somewhat obscure to most physicists. Start by rewriting the First Law for quasi-static processes in (3.185) as

$$\mu = \frac{E + PV - TS}{N}. \quad (4.175)$$

We'll rederive the dependence of the number density ν on height z in an atmosphere of an ideal gas of point particles, (4.149). The following discussion is known for its unusual demand that we work with identical-classical particles and not distinguishable particles. Just why this should be so is currently an open problem in the subject.

The entropy of an ideal gas of identical-classical point particles, each of mass m , is given by the Sackur–Tetrode equation (3.146):

$$S = Nk \left[\ln \frac{V}{N} + \frac{5}{2} + \frac{3}{2} \ln \frac{2\pi mkT}{h^2} \right]. \quad (4.176)$$

We wish to investigate how an atmosphere's particle density $\nu(z) \equiv N(z)/V$ changes with height $z > 0$ above ground. For brevity, write $N(z)$ as N , and recall that the total energy of an ideal gas of point particles at height z is

$$E = 3NkT/2 + Nmgz. \quad (4.177)$$

Apply (4.175) with $PV = NkT$, and incorporate (4.176), with a switch from V/N to $\nu = N/V$ in the third line of the following:

$$\begin{aligned} \mu &= E/N + PV/N - TS/N \\ &= \frac{3kT}{2} + mgz + kT - kT \left[\ln \frac{V}{N} + \frac{5}{2} + \frac{3}{2} \ln \frac{2\pi mkT}{h^2} \right] \\ &= mgz + kT \ln \frac{N}{V} - \frac{3kT}{2} \ln \frac{2\pi mkT}{h^2} \\ &\equiv mgz + kT \ln \nu(z) + f(m, T), \end{aligned} \quad (4.178)$$

where we write the last term in the third line as $f(m, T)$, because its precise form isn't required here. Note that the “ N ” in $kT \ln(N/V)$ on the third line of (4.178) would have been absent had we used the entropy for *distinguishable* particles in (3.145), which would have prevented any further discussion about particles at all.

The central expression here is

$$\mu(z) = mgz + kT \ln \nu(z) + f(m, T), \quad (4.179)$$

with $f(m, T)$ defined in (4.178). The chemical potential does not vary with height in an atmosphere at equilibrium: $\mu(z) = \mu(0)$. (This was discussed in Section 3.4.2, and is also able to be inferred from the discussion of flow directions in Section 3.13. Remember that an atmosphere's height-dependent density at equilibrium is associated with the chemical potential always quickly evolving to be independent of height.) So, write this equality $\mu(z) = \mu(0)$ using (4.179). Assume that the atmosphere's temperature T is constant with height z , which enables $f(m, T)$ to be cancelled out in the following expression:

$$\begin{aligned} \mu(z) &= \mu(0), \text{ and so:} \\ mgz + kT \ln \nu(z) + f(m, T) &= kT \ln \nu(0) + f(m, T). \end{aligned} \quad (4.180)$$

Equation (4.180) rearranges to produce the particle-number density $\nu(z)$:

$$\nu(z) = \nu(0) \exp \frac{-mgz}{kT}. \quad (4.181)$$

This agrees with (4.149), and demonstrates the chemical potential's role in situations that involve macroscopic particle movement, such as in our atmosphere.

It's worthwhile to compare the above expressions with those of the discussion of the chemical potential in Section 3.4.2. The discussion in that section

used a slightly different notation so as to be very explicit but still concise. Revisit that notation by writing the above $\nu(z)$ as ν_z , and the above $\mu(z)$ as $\mu(z, \nu_z)$. The expression of the chemical potential not varying with height in an atmosphere at equilibrium above was $\mu(z) = \mu(0)$, which is now written as

$$\mu(z, \nu_z) = \mu(0, \nu_0). \quad (4.182)$$

This matches the right-hand part in Figure 3.8 and (3.65). Note also that

$$\begin{aligned} \mu(z, \nu_z) - \mu(0, \nu_z) &\stackrel{(4.179)}{=} mgz + kT \ln \nu_z + f(m, T) \\ &\quad - [kT \ln \nu_z + f(m, T)] \\ &= mgz, \end{aligned} \quad (4.183)$$

which matches (3.66). [Take careful note that the above expression and (3.66) refer to $\mu(0, \nu_z)$, not $\mu(0, \nu_0)$!] Also, writing (4.179) as

$$\mu(z, \nu_z) = mgz + kT \ln \nu_z + f(m, T) \quad (4.184)$$

enables the second line of (4.180) to be written as

$$mgz + \mu(0, \nu_z) = \mu(0, \nu_0). \quad (4.185)$$

This is (3.67) again.

The chemical potential finds a major use in the study of *phase transitions*, which we study next.

4.3.4 Phase Transitions and the Clausius–Clapeyron Equation

Recall that the chemical potential μ is the potential energy lost by a single particle as it enters a system, and it determines how particles move—just as the gravitational potential energy determines how particles fall. The chemical potential is thus well suited to describing phase transitions, such as how liquids boil, and how particles react to their immediate environment when they are in a liquid solution.

In particular, based on knowledge of how μ evolves in various situations, the Gibbs–Duhem equation (3.217) can be manipulated to produce the *Clausius–Clapeyron equation*, which enables us to probe phase changes of solids and liquids. The basic idea is shown at the left in Figure 4.13. This diagram helps us to study how a substance might move between its liquid and vapour forms as a function of pressure and temperature. Each point on the diagram defines particular values of the chemical potentials of the liquid and vapour phases μ_{liq} and μ_{vap} , respectively.

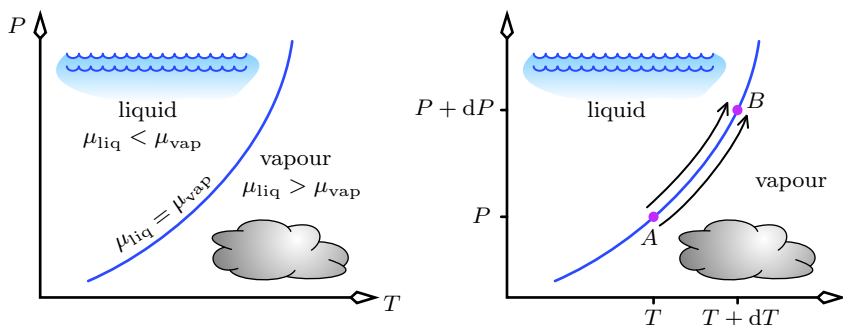


Fig. 4.13 **Left:** Generic pressure–temperature diagram, showing regions where a substance is in liquid and vapour forms. Recall that particles tend to move from higher to lower chemical potentials. **Right:** Just to the left of the dividing line between the phases, the liquid can follow a path from state A to state B . The vapour can do the same, just to the right of the dividing line

At higher pressures and/or lower temperatures, the substance prefers to be in liquid form: here we have $\mu_{\text{liq}} < \mu_{\text{vap}}$, since particles tend to advance from states of higher to lower chemical potentials: vapour to liquid here. The situation is reversed at lower pressures and higher temperatures, for which the vapour form is preferred. Separating these phases on the diagram is a boundary on which $\mu_{\text{liq}} = \mu_{\text{vap}}$.

Suppose the pressure and temperature are such that the liquid and vapour are in equilibrium, which places the substance somewhere on the boundary curve shown in the figure, where both liquid and vapour have the same chemical potential μ . A value of pressure P on this boundary is called the *vapour pressure* for its corresponding temperature T . Follow N particles of the liquid as they traverse from state A , in which their entropy and volume are S_{liq} and V_{liq} , respectively, to state B , following the arrow drawn just to the left of the phase boundary shown in the right-hand picture in Figure 4.13. Apply the Gibbs–Duhem equation (3.217), writing

$$N d\mu = -S_{\text{liq}} dT + V_{\text{liq}} dP. \quad (4.186)$$

Similarly, follow N particles of the *vapour* as they traverse from state A (entropy and volume S_{vap} and V_{vap} respectively) to state B , following the arrow drawn just to the right of the phase boundary in Figure 4.13. Here, the Gibbs–Duhem equation (3.217) is

$$N d\mu = -S_{\text{vap}} dT + V_{\text{vap}} dP. \quad (4.187)$$

Equations (4.186) and (4.187) combine to yield

$$(V_{\text{vap}} - V_{\text{liq}}) dP = (S_{\text{vap}} - S_{\text{liq}}) dT. \quad (4.188)$$

From this, the slope of the boundary curve on the pressure–temperature diagram in Figure 4.13 is seen to be

$$\frac{dP}{dT} = \frac{S_{\text{vap}} - S_{\text{liq}}}{V_{\text{vap}} - V_{\text{liq}}}. \quad (4.189)$$

This is one form of the Clausius–Clapeyron equation. Although (4.189) *does* find use, it presents a difficulty, since entropies of substances tend not to be easy to measure. Alternatively, we can re-express its right-hand side using parameters that are easily measured.¹⁵

First, $S_{\text{vap}} - S_{\text{liq}}$ is the entropy gained by the N particles as they jump across the phase boundary, changing from “liquid form” to “vapour form” (so to speak). Follow them as they leave the liquid: to escape the binding forces of the liquid’s surface, they have been given the liquid’s *latent heat of vaporisation*, energy that increases the average distance between the particles without increasing their kinetic energy. (This embodies the observation that when we boil a liquid, it changes to vapour without changing its temperature. The same idea applies to heating ice: the ice melts with no temperature change.) Chemists tend to work with the latent heat per mole, or *molar latent heat of vaporisation* $L_{\text{vap}}^{\text{mol}}$. (They also tend to call it the *enthalpy of vaporisation*, and write it as ΔH_{vap} .) We’ll use $L_{\text{vap}}^{\text{mol}}$ too, because this is usually tabulated for reference. Use the idea that “heat Q ”, or energy supplied thermally, increases entropy by¹⁶ $\Delta S = Q/T$, and switch to speaking of the $n = N/N_A$ moles of particles present:

$$S_{\text{vap}} - S_{\text{liq}} = \frac{\text{latent heat of } n \text{ moles of particles}}{T} = \frac{L_{\text{vap}}^{\text{mol}} n}{T}. \quad (4.190)$$

Next, we require $V_{\text{vap}} - V_{\text{liq}}$. The volume V_{vap} occupied by the particles in vapour phase is much larger than their volume V_{liq} in liquid phase, so apply this idea, along with the ideal-gas law, to write

$$V_{\text{vap}} - V_{\text{liq}} \simeq V_{\text{vap}} = \frac{nRT}{P}, \quad (4.191)$$

where R is the gas constant. Equation (4.189) now becomes

$$\frac{dP}{dT} = \frac{L_{\text{vap}}^{\text{mol}} n}{T} \times \frac{P}{nRT} = \frac{L_{\text{vap}}^{\text{mol}} P}{RT^2}. \quad (4.192)$$

Rearrange this equation to obtain

¹⁵ Although the numerator and denominator of the right-hand side of (4.189) refer to some arbitrary number N of particles, we could just as well divide the entropies and volumes by N , to re-interpret the right-hand quantities as entropy and volume *per particle*; or we could multiply them by N_A/N , to re-interpret them as entropy and volume *per mole*.

¹⁶ Recall that this is the non-infinitesimal form of (3.141), “ $dS = dQ/T$ ”.

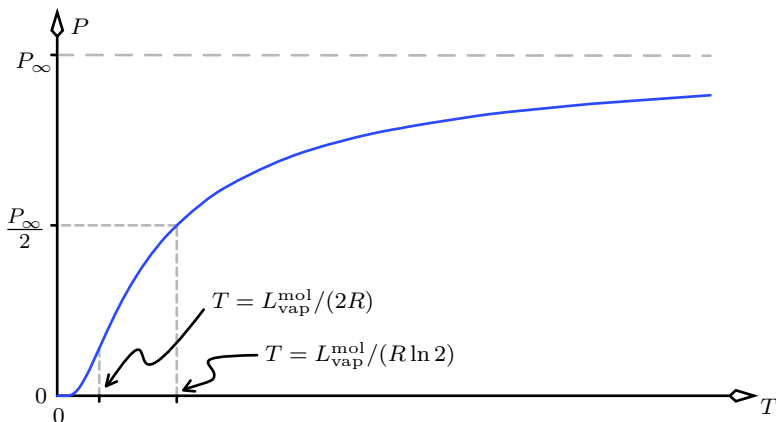


Fig. 4.14 Plot of the Clausius–Clapeyron equation in (4.194). The inflection point at $T = L_{\text{vap}}^{\text{mol}}/(2R)$ is found by solving $d^2P/dT^2 = 0$

$$\frac{dP}{P} = \frac{L_{\text{vap}}^{\text{mol}} dT}{RT^2}. \quad (4.193)$$

This integrates to give the vapour pressure as a function of temperature:¹⁷

$$P = P_{\infty} \exp \frac{-L_{\text{vap}}^{\text{mol}}}{RT}, \quad \text{where } P_{\infty} \equiv P(T \rightarrow \infty). \quad (4.194)$$

Like (4.189), equation (4.194) is often given the name Clausius–Clapeyron, and its form is shown in Figure 4.14.

Equation (4.194) can be put into a more practical form by writing it for pressures P_1, P_2 and their respective temperatures T_1, T_2 . Start with

$$P_1 = P_{\infty} \exp \frac{-L_{\text{vap}}^{\text{mol}}}{RT_1}, \quad P_2 = P_{\infty} \exp \frac{-L_{\text{vap}}^{\text{mol}}}{RT_2}. \quad (4.195)$$

The ratio of the two expressions in (4.195) is

$$\frac{P_1}{P_2} = \exp \left[\frac{L_{\text{vap}}^{\text{mol}}}{R} \left(\frac{1}{T_2} - \frac{1}{T_1} \right) \right]. \quad (4.196)$$

This form of the Clausius–Clapeyron equation is widely used by chemists to calculate pressures and boiling points. Here are two examples of such calculations.

¹⁷ We evaluated this integral in the earlier discussion starting with (1.234).

Pressure and Boiling Point of Water

Water's molar latent heat of vaporisation is $L_{\text{vap}}^{\text{mol}} = 40.7 \text{ kJ/mol}$. Calculate its vapour pressures at 50°C , 100°C , and 150°C .

Water boils at 100°C at one atmosphere (101,325 Pa), meaning its vapour pressure at $T = 100^\circ\text{C}$ is precisely $P = \text{one atmosphere}$, since, at this temperature, the water molecules are just able to counteract atmospheric pressure and begin to leave the surface. In (4.196), set

$$P_2 = 1 \text{ atm (that is, 1 atmosphere) and } T_2 = 100^\circ\text{C}. \quad (4.197)$$

T_1 is set to the required temperature, and we solve (4.196) for P_1 , using SI units where appropriate. For $T_1 = 50^\circ\text{C}$, (4.196) gives us

$$\frac{P_1}{1 \text{ atm}} = \exp \left[\frac{40,700}{8.314} \left(\frac{1}{373} - \frac{1}{323} \right) \right] = 0.13, \quad (4.198)$$

so that $P_1 = 0.13$ atmospheres. For $T_1 = 150^\circ\text{C}$, we have

$$\frac{P_1}{1 \text{ atm}} = \exp \left[\frac{40,700}{8.314} \left(\frac{1}{373} - \frac{1}{423} \right) \right] = 4.7, \quad (4.199)$$

so $P_1 = 4.7$ atmospheres. This high pressure becomes the driving force in a steam engine.

What is water's boiling point at the top of Mount Everest?

Recall from the discussion just after (4.156) that to the height of Everest (8848 m) and somewhat beyond, our atmosphere has an exponential decrease in pressure with height, with a half fall-off distance of about 5.35 km. We estimate the pressure at the top of Everest then to be

$$1 \text{ atmosphere} \times 2^{-8.848/5.35} \simeq 0.32 \text{ atmospheres}. \quad (4.200)$$

This compares well with the standard more accurate value of 0.31 atmospheres. Now call on (4.196) with “1” = top of Everest and “2” = sea level:

$$\frac{0.31 \text{ atm}}{1 \text{ atm}} = \exp \left[\frac{40,700}{8.314} \left(\frac{1}{373} - \frac{1}{T_1} \right) \right]. \quad (4.201)$$

This leads to $T_1 = 342.4 \text{ K} = 69.3^\circ\text{C}$. Water is easy to boil on mountain tops, because much less atmospheric pressure exists there to prevent slightly heated water molecules from leaving the water's surface. But this comparatively cool boiling water might not be hot enough to boil an egg. A team of mountaineers ascending to extreme altitudes can carry a small pressure cooker for their meals. This is a strong pot with a tightly sealed

lid, which allows the internal pressure to rise to a high level. This high pressure discourages water molecules from leaving the water's surface, and thus forces the water to boil at a higher temperature.

Gibbs–Duhem and Osmotic Pressure

In the next few pages, we will apply the chemical potential to describe *osmotic pressure*, and predict how the melting and boiling points of water change when a small amount of substance, such as common salt, is dissolved in the water.

Consider first pure water. This consists mostly of H_2O molecules, with a vanishingly small number of H_3O^+ and OH^- ions mixed in.¹⁸ We will refer collectively to all of these molecules as “water particles”. Now add a *small* amount of common salt, NaCl , to the water. (Adding a small amount of salt allows us to discuss salt water, as opposed to wet salt!) The salt dissociates into Na^+ and Cl^- ions, which we will collectively call “salt particles”. This number of salt particles is a fraction ϕ of the total number of particles of salt and water present:

$$\phi \equiv \frac{\text{no. of salt particles}}{\text{no. of salt particles} + \text{no. of water particles}}. \quad (4.202)$$

How does this addition of salt change the chemical potential of what was originally pure water? Refer to the central equation (4.179) with $g = 0$ (we are not concerned with a variable gravity field here). Note that although we derived that equation for a monatomic ideal gas, it also holds well for liquids; and so we can use it to calculate the water's chemical potential before and after adding the salt. We require the particle density $\nu \equiv N/V$. Suppose that one water particle occupies some small volume v_{water} in the solution. The initial particle density ν_i is one water particle per volume v_{water} , or just

$$\nu_i = 1/v_{\text{water}}. \quad (4.203)$$

Alternatively, to match the equations that follow, we could write

$$\nu_i = \frac{\text{no. of water particles}}{\text{volume they occupy}} = \frac{\text{no. of water particles}}{\text{no. of water particles} \times v_{\text{water}}} = \frac{1}{v_{\text{water}}}. \quad (4.204)$$

¹⁸ How many of these H_3O^+ and OH^- ions are present? Recall that pure water has a pH of 7. The pH is minus the base-10 logarithm of the concentration of H_3O^+ ions when measured in moles per litre. A litre of pure water then contains 10^{-7} moles of H_3O^+ ions; and for each of these ions, there exists one ion of OH^- . This litre also contains about 56 moles of H_2O molecules. The H_3O^+ and OH^- are thus present in only a tiny proportion.

Now add the salt, and suppose too that any one salt particle also occupies volume v_{water} in the solution. The final particle density of the *water* is

$$\begin{aligned}\nu_f &= \frac{\text{no. of water particles}}{(\text{no. of salt particles} + \text{no. of water particles}) \times v_{\text{water}}} \\ &= \left(1 - \frac{\text{no. of salt particles}}{\text{no. of salt particles} + \text{no. of water particles}}\right) \frac{1}{v_{\text{water}}} \\ &= (1 - \phi)\nu_i.\end{aligned}\tag{4.205}$$

Referring to (4.179), the process of adding salt increases the water's chemical potential by (realising that the “ m ” in this latter function is an average mass of all particles present)

$$\Delta\mu = \mu_f - \mu_i \simeq kT \ln \nu_f + f(m, T_f) - kT \ln \nu_i - f(m, T_i).\tag{4.206}$$

Any temperature change that occurs is very minor (you don't notice it when you dissolve salt into water), so $T_f = T_i$, and

$$\Delta\mu \simeq kT \ln \frac{\nu_f}{\nu_i} = kT \ln(1 - \phi).\tag{4.207}$$

Simplify the last term of (4.207) in the usual way by noting that when $|x| \ll 1$, taking the logarithm of (1.24) produces¹⁹ $x \simeq \ln(1 + x)$. Equation (4.207) thus becomes

$$\Delta\mu \simeq -\phi kT.\tag{4.209}$$

So, with the addition of the salt, the water's chemical potential reduces from μ_i to $\mu_i - \phi kT$.

This reduction in chemical potential explains the well-known fact that drinking sea water when you are thirsty will only increase your thirst. To see how this comes about, refer to the left-hand picture in Figure 4.15. This shows two containers of pure water at the same temperature, pressure, and chemical potential μ_i , that are separated by a membrane through which water particles can pass, but not salt particles. Add a fraction ϕ of salt to the left container, which lowers its chemical potential to $\mu_i - \phi kT$, as shown in the middle picture in the figure. Next, allow some small time to pass in which the system evolves to the right-hand picture in Figure 4.15. Now apply the Gibbs–Duhem equation (3.217) to this process of pure water transforming to brine, which is accompanied by changes in temperature, pressure, and chemical potential: that is, we are essentially rewriting (4.186), but now as

¹⁹ This expression is recognisable as the truncation to first order of the Taylor series

$$\ln(1 + x) = x - x^2/2 + x^3/3 - \dots, \quad -1 < x \leq 1.\tag{4.208}$$

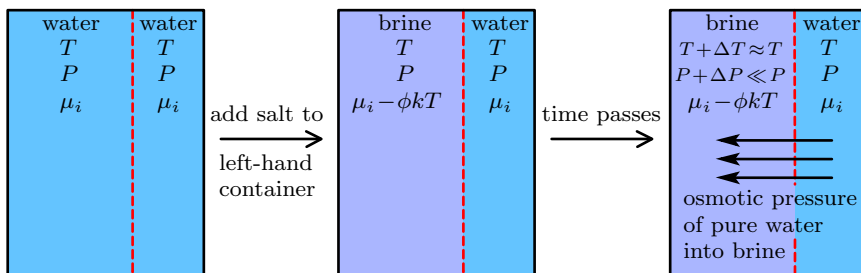


Fig. 4.15 Left: Initially, we have two containers of pure water separated by a membrane (shown in red) through which water particles can pass, but not salt particles. **Middle:** Salt is now added to the left-hand container. **Right:** The Gibbs–Duhem equation predicts that the temperature, pressure, and chemical potential in the left-hand container will change

an approximation for non-infinitesimal changes.²⁰ We obtain

$$-S_{\text{brine}}\Delta T + V_{\text{brine}}\Delta P \simeq N_{\text{brine}}\Delta\mu = -N_{\text{brine}}\phi kT. \quad (4.210)$$

Although the temperature change on adding the salt is negligible, we could just as well maintain the salt water at its original temperature, thus ensuring that $\Delta T = 0$ in (4.210). But the pressure change is marked: the pressure in the pure water is still P , but the pressure in the brine is

$$P + \Delta P \simeq P - N_{\text{brine}}\phi kT/V_{\text{brine}}. \quad (4.211)$$

This manifests as an *osmotic pressure* that forces water molecules to diffuse from high to low potentials: μ_i to $\mu_i - \phi kT$, or from the pure water to the brine through the membrane:

$$\text{osmotic pressure} \equiv |\Delta P| \simeq \frac{N_{\text{brine}}}{V_{\text{brine}}}\phi kT. \quad (4.212)$$

(We might prefer to be careful here by writing

$$\begin{aligned} \text{osmotic pressure} &\equiv \text{pressure from water to brine} \\ &= P_{\text{water}} - P_{\text{brine}} = P - (P + \Delta P) \\ &= -\Delta P \simeq \frac{N_{\text{brine}}}{V_{\text{brine}}}\phi kT. \end{aligned} \quad (4.213)$$

It’s always wise to check that minus signs make sense!)

²⁰ Omitting the subscript “brine” for brevity in this footnote, it’s worth saying that we could include the flow of particles by replacing the N in (4.210) with $N + \Delta N$, but that would only add second-order terms $\Delta N \Delta T$ and $\Delta N \Delta P$ to that equation. Hence, we ignore the ΔN here.

The number density $N_{\text{brine}}/V_{\text{brine}}$ is the number of salt particles per unit volume of brine. When only a small amount of salt is present, this density approximately equals the number of water particles per unit volume of water ($N_{\text{water}}/V_{\text{water}}$), which is more easily calculated. So, write (4.212) as

$$\text{osmotic pressure} \simeq \frac{N_{\text{water}}}{V_{\text{water}}} \phi kT. \quad (4.214)$$

This expression no longer refers to the salt, and thus must hold quite generally, as long as ϕ is small and each salt particle occupies a volume similar to that of a water particle. Indeed, it holds independently of “salt” and “water”; we could replace the salt with a generic *solute*, and the water with a generic *solvent*.

These chemical calculations are usually written more easily using molar quantities, since everything can then easily be calculated using molar masses. Use the idea that the N_{water} particles of pure water that occupy volume V_{water} are equivalent to n_{water} moles of water particles:

$$\begin{aligned} \text{osmotic pressure} &\simeq \frac{N_{\text{water}} \phi kT}{V_{\text{water}}} = \frac{n_{\text{water}} N_A \phi kT}{V_{\text{water}}} \\ &= \frac{\phi RT}{V_{\text{water}}/n_{\text{water}}} = \frac{\phi RT}{\text{molar volume of water}}. \end{aligned} \quad (4.215)$$

Drinking Sea Water

What is the osmotic pressure of pure water diffusing into sea water at 25°C? Sea water typically has a salt content by weight of about 3.5%. We require ϕ , where sea water has ϕ salt particles (whether sodium or chlorine is immaterial) for every $1-\phi$ water particles. That is, model sea water to have $\phi/2$ sodium atoms (really ions, but the exchange of an electron between the sodium and chlorine doesn't affect the numbers), $\phi/2$ chlorine atoms, and $1-\phi$ H_2O molecules (a tiny fraction of which are dissociated into H_3O^+ and OH^- , but again, that doesn't affect the numbers). Use the following molar masses:

$$\text{Na: } 23 \text{ g}, \quad \text{Cl: } 35.5 \text{ g}, \quad \text{H: } 1 \text{ g}, \quad \text{O: } 16 \text{ g}.$$

It follows that the relative mass of salt particles (each of which is either a sodium or chlorine ion) in the water is

$$\frac{\phi/2 \times 23 + \phi/2 \times 35.5}{\phi/2 \times 23 + \phi/2 \times 35.5 + (1-\phi) \times 18}. \quad (4.216)$$

Given that sea water's salt content by weight is 3.5%, the expression in (4.216) must equal 0.035. It follows that $\phi \simeq 0.022$. Equation (4.215)

also requires the molar volume of pure water. One mole of pure water has mass 18 g, and so has volume of 18 cm^3 , or $18 \times 10^{-6} \text{ m}^3$. The osmotic pressure now follows from (4.215):

$$\begin{aligned} \text{osmotic pressure} &= \frac{0.022 \times 8.314 \times 298}{18 \times 10^{-6}} \text{ Pa} \\ &= 3.0 \text{ MPa} \simeq 30 \text{ atmospheres.} \end{aligned} \quad (4.217)$$

This is very large, showing that pure water will push its way into salty water with considerable force. The salinity of humans lies between pure water and sea water:

$$\begin{array}{ccccc} \mu_{\text{sea water}} & < & \mu_{\text{human}} & < & \mu_{\text{pure water}} \\ & \longleftarrow & & \longleftarrow & \\ & \text{osmotic pressure} & & \text{osmotic pressure} & \end{array}$$

Following the direction of the osmotic pressure, we see that when we drink pure water, it diffuses into our organs, which is a good thing and is how we absorb water in everyday life. But when we drink sea water, pure water diffuses out of our organs into the sea water, causing us to dehydrate.

What has caused what here: has the force of the diffusing water simply manifested as a pressure, or has a new phenomenon called “osmotic pressure” caused the water to diffuse? This question was discussed earlier at the end of Section 3.13. In this case, it seems clear that the simple operation of entropy growth causes pure water to move into salty water, which manifests as a pressure on the pure water—even though conventionally, we say the reverse: that an “osmotic pressure” has *pushed* pure water into the salty water. But more generally, even when entropy growth can be tied to the operation of some mechanism, it’s not clear at all whether invoking the blind growth of entropy as a sort of primal force is entirely useful for discussing that mechanism from a predictive point of view.

Clausius–Clapeyron and the Melting and Boiling Points of a Solution

The perturbation of the chemical potential that we have discussed in the last few pages explains why the addition of, say, salt alters the melting and boiling points of water. Figure 4.16 shows the scenario. At its left is a container of pure water that is in equilibrium with its vapour: they share the same temperature, pressure, and chemical potential μ_i . A proportion ϕ of salt is then added to the water, reducing the water’s chemical potential to $\mu_i - \phi kT$

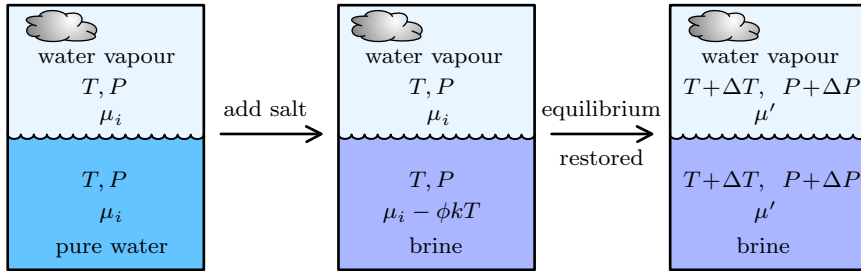


Fig. 4.16 **Left:** Initially, pure water is in particle equilibrium with its vapour, with all at chemical potential μ_i . **Middle:** Salt is now added to the water, which lowers its chemical potential below that of the vapour to $\mu_i - \phi kT$, as in (4.209). **Right:** The vapour and brine interact to restore a balance based on a new chemical potential μ'

[recall (4.209)], and so destroying the water–vapour equilibrium, as shown in the middle picture. Finally, at right in the figure, the brine–vapour equilibrium has been restored with a new temperature, pressure, and chemical potential μ' .

Apply the Gibbs–Duhem equation (4.187) in non-infinitesimal form in the way that we did in Figure 4.13, by following N particles of the vapour as it acts to restore equilibrium from the second to the third picture in Figure 4.16:

$$-S_{\text{vap}}\Delta T + V_{\text{vap}}\Delta P = N\Delta\mu_{\text{vap}} = N(\mu' - \mu_i). \quad (4.218)$$

The brine and vapour are finally in thermal and pressure equilibrium, and so the brine’s temperature and pressure also increase by the same amounts $(\Delta T, \Delta P)$ as equilibrium is restored. Again in the spirit of Figure 4.13, follow the behaviour of N particles on the liquid side of the liquid–vapour boundary curve: Gibbs–Duhem for the brine is then

$$\begin{aligned} -S_{\text{liq}}\Delta T + V_{\text{liq}}\Delta P &= N\Delta\mu_{\text{liq}} = N[\mu' - (\mu_i - \phi kT)] \\ &= N(\mu' - \mu_i) + N\phi kT \quad [\text{and now recall (4.218)}] \\ &= -S_{\text{vap}}\Delta T + V_{\text{vap}}\Delta P + N\phi kT. \end{aligned} \quad (4.219)$$

This rearranges to

$$(S_{\text{vap}} - S_{\text{liq}})\Delta T - (V_{\text{vap}} - V_{\text{liq}})\Delta P = N\phi kT. \quad (4.220)$$

This is another instance of the Clausius–Clapeyron equation: apart from the “ Δ ” versus “ d ”, it is identical to (4.188) when no salt is present ($\phi = 0$). Examine it as follows for the two cases of fixed temperature and fixed pressure.

– **When the temperature is held constant** ($\Delta T = 0$), (4.220) becomes

$$-(V_{\text{vap}} - V_{\text{liq}})\Delta P = N\phi kT. \quad (4.221)$$

But the volume of N vapour particles is much greater than the volume of N liquid particles: $V_{\text{vap}} \gg V_{\text{liq}}$; also, treat the vapour as an ideal gas following $PV_{\text{vap}} = NkT$. We arrive at

$$V_{\text{vap}} - V_{\text{liq}} \simeq V_{\text{vap}} = NkT/P. \quad (4.222)$$

Equation (4.221) then becomes

$$\frac{-NkT}{P} \Delta P = N\phi kT, \quad \text{or} \quad \frac{-\Delta P}{P} = \phi. \quad (4.223)$$

That is, the relative drop in vapour pressure is approximately ϕ , the fraction of salt particles in the solution.

- **When the pressure is held constant** ($\Delta P = 0$), (4.220) becomes

$$(S_{\text{vap}} - S_{\text{liq}}) \Delta T = N\phi kT. \quad (4.224)$$

We calculated $S_{\text{vap}} - S_{\text{liq}}$ in (4.190), so write

$$\Delta T = \frac{N\phi kT}{S_{\text{vap}} - S_{\text{liq}}} \stackrel{(4.190)}{=} N\phi kT \times \frac{T}{L_{\text{vap}}^{\text{mol}} n} = \frac{N\phi kT^2}{L_{\text{vap}}^{\text{mol}} n} = \frac{\phi RT^2}{L_{\text{vap}}^{\text{mol}}}. \quad (4.225)$$

Picture adding salt to a pot of pure water that is boiling at 100°C . This water is in equilibrium with its vapour. Adding the salt lowers the chemical potential of the water by ϕkT below that of its vapour. Diffusion always occurs in the direction of higher to lower chemical potential; hence vapour particles will start to enter the boiling brine, which then causes a lowering of the vapour pressure, as we saw in (4.223). If we wish to restore the liquid–vapour equilibrium at constant pressure, we must give these new water particles that are entering the brine the energy to escape again—meaning we must increase the temperature by the ΔT of (4.225): this ΔT is just the temperature increase that must occur when pressure is held fixed. Hence, the boiling point of salt water is higher than that of pure water.

In particular, suppose we add a mole of salt to a litre of water: what will be the temperature at which the salt water boils? Refer to (4.225), which requires ϕ , defined in (4.202). A mole of salt particles contains N_A atoms of sodium and N_A atoms of chlorine, totalling $2N_A$ atoms (that is, $2N_A$ salt particles). Also, a litre of water has mass 1000 g, and water’s molar mass is 18 g. Thus,

$$\begin{aligned} \phi &= \frac{\text{no. of salt particles}}{\text{no. of salt particles} + \text{no. of water particles}} \\ &= \frac{2N_A}{2N_A + 1000N_A/18} \simeq 0.035. \end{aligned} \quad (4.226)$$

The mixture's molar latent heat of vaporisation has the tabulated value of $L_{\text{vap}}^{\text{mol}} = 40,700 \text{ J/mol}$. Equation (4.225) now yields

$$\Delta T = \frac{\phi RT^2}{L_{\text{vap}}^{\text{mol}}} = \frac{0.035 \times 8.314 \times 373^2}{40,700} \text{ K} \simeq 1.0 \text{ K}. \quad (4.227)$$

The boiling point of the brine is thus about 101°C . Note that (4.225) says that the temperature increase is proportional to ϕ ; but because the number of salt particles is much less than the number of water particles, ϕ is roughly proportional to the amount of salt added. Hence, the temperature increase is also proportional to the amount of salt added. A temperature increase of 1 K per mole of salt²¹ added is indeed confirmed experimentally.

A similar analysis shows that the melting point of water ice is lowered when salt is added to it. Do this by replacing the water vapour in the above discussion with ice. That is to say, where just after (4.225) we added salt to boiling water in equilibrium with its vapour at 100°C , now instead, we add salt to cold water in equilibrium with water ice at 0°C . Again, the salt lowers the chemical potential of the water; but in place of vapour particles entering the newly created brine, ice particles now enter the brine. So, when salt is added to a water/ice mixture, the ice begins to melt. Spreading salt over an icy road to melt the ice is, of course, well known to those who drive in countries whose roads ice up in winter. The salt melts the ice, but we can also view this as the salt having lowered the ice's melting point.

This lowering of ice's melting point is predicted by (4.225). Consider that with the ice in the salt–ice–water mixture melting, to restore equilibrium at constant pressure, we must now *remove* thermal energy. This is equivalent to using a negative value of the latent heat in (4.225). The relevant parameter is called the *molar latent heat of fusion* $L_{\text{fusion}}^{\text{mol}}$ of water, which is conventionally taken as positive, and whose value for water is about 6010 J/mol . So, we must replace $L_{\text{vap}}^{\text{mol}}$ in (4.225) with $-L_{\text{fusion}}^{\text{mol}}$. The new melting point of the ice is $0^\circ\text{C} + \Delta T$, where

$$\Delta T = \frac{\phi RT^2}{-L_{\text{fusion}}^{\text{mol}}} = \frac{0.035 \times 8.314 \times 273^2}{-6010} \text{ K} \simeq -3.6 \text{ K}. \quad (4.228)$$

The new melting point—or freezing point, if you prefer—of the brine is -3.6°C . As before, ΔT is approximately proportional to the amount of salt added. Although sprinkling greater amounts of salt on an icy road might mean fast-icing de-icing, in practice, we must weigh this against the cost of the salt and the rust damage that it does to cars.

²¹ For reference, NaCl has molar mass $M_{\text{mol}}(\text{Na}) + M_{\text{mol}}(\text{Cl}) = 23 \text{ g} + 35.5 \text{ g} = 58.5 \text{ g}$.

4.3.5 Chemical Equilibrium

We conclude this chapter with a short discussion of the role of the chemical potential in equilibrium chemical reactions.

Recall the discussion in Section 3.14.1, where the Gibbs energy was defined to be $G = E - TS + PV = \mu N$, with μN understood here to denote $\sum \mu_i N_i$ when several species of particle are present. As we saw in that section, when several systems interact, their total Gibbs energy decreases as the whole heads toward equilibrium, at which point G is a minimum.

This idea allows us to determine the direction in which a chemical reaction will naturally proceed. Consider molecules A, B, C that can react in either direction:



for some numbers n_A, n_B, n_C . Suppose that at some point in time, we have measured the species to have chemical potentials μ_A, μ_B, μ_C . These potentials will determine the direction in which the reaction proceeds, by way of the fact that $G = \sum \mu_i N_i$ always decreases on the way to equilibrium. We need only calculate ΔG for each direction of the reaction: the direction for which ΔG is negative is where the reaction proceeds. In the common case of fixed temperature and pressure, refer to (3.211) to write $dG = \mu dN$; but when more than one particle species is present, we write $dG = \sum_i \mu_i dN_i$, or

$$\Delta G \simeq \sum_i \mu_i \Delta N_i. \quad (4.230)$$

If you make a point of remembering that Δ always refers to an *increase*, you will always have the correct signs in analyses like the ones immediately below. Recall the comments in Section 1.6 to ensure that each ΔN_i has the correct sign.

Let's apply (4.230) to calculate ΔG for each direction in (4.229). We require ΔN_i for the three species in (4.229):

1. **The reaction proceeds left to right:** Here, the mixture loses n_A molecules of A and n_B molecules of B , and gains n_C molecules of C :

$$-\Delta N_A = n_A, \quad -\Delta N_B = n_B, \quad \Delta N_C = n_C. \quad (4.231)$$

It follows that ΔG for the left-to-right reaction is

$$\Delta G_{L \rightarrow R} = \sum_i \mu_i \Delta N_i = -n_A \mu_A - n_B \mu_B + n_C \mu_C. \quad (4.232)$$

2. **The reaction proceeds right to left:** Now, everything is reversed: the mixture gains n_A molecules of A (so $\Delta N_A = n_A$) and so on. Hence, all the signs in the calculation of ΔG are reversed from above, and

$$\Delta G_{R \rightarrow L} = -\Delta G_{L \rightarrow R}. \quad (4.233)$$

At equilibrium, both $\Delta G_{L \rightarrow R}$ and $\Delta G_{R \rightarrow L}$ are zero; but on the way to equilibrium, one of them must be negative (meaning G is decreasing), and this one tells us the direction in which the reaction proceeds.

As a convention, the increases in particle numbers for the reaction proceeding *left to right*, as in (4.231), are called its *stoichiometric coefficients*. Consider a general reaction with stoichiometric coefficients $b_1, b_2, b_3, \dots$. The increase in the Gibbs energy for the left-to-right direction is

$$\Delta G_{L \rightarrow R} = \sum_i \mu_i \Delta N_i = \sum_i \mu_i b_i. \quad (4.234)$$

But we saw in (4.179), that the i^{th} chemical potential can be written in terms of the particle density ν_i as (and here gravity is not relevant)

$$\mu_i = kT \ln \nu_i + f(m_i, T). \quad (4.235)$$

For notational convenience, define a ζ_i and use it to rewrite $f(m_i, T)$ in (4.235) as $-kT \ln \zeta_i$. In that case,

$$\mu_i = kT [\ln \nu_i - \ln \zeta_i] = kT \ln \frac{\nu_i}{\zeta_i}. \quad (4.236)$$

Equation (4.234) becomes

$$\Delta G_{L \rightarrow R} = \sum_i b_i kT \ln \frac{\nu_i}{\zeta_i}. \quad (4.237)$$

It follows that

$$\begin{aligned} \exp \frac{\Delta G_{L \rightarrow R}}{kT} &= \exp \sum_i b_i \ln \frac{\nu_i}{\zeta_i} = \exp \sum_i \ln \left(\frac{\nu_i}{\zeta_i} \right)^{b_i} \\ &= \prod_i \left(\frac{\nu_i}{\zeta_i} \right)^{b_i} = \frac{\prod_i \nu_i^{b_i}}{\prod_i \zeta_i^{b_i}}. \end{aligned} \quad (4.238)$$

Chemists write the concentration ν_i of molecule i as “[i]”, and we’ll follow suit. The denominator in the last expression of (4.238) is called the reaction’s *equilibrium constant*. Generally, in studies of the directions in which two-way reactions can proceed, only the temperature is made to vary. In that case, we’ll write the equilibrium constant as a function only of temperature:

$$\text{equilibrium constant } A(T) \equiv \zeta_1^{b_1} \zeta_2^{b_2} \zeta_3^{b_3} \dots \quad (4.239)$$

Examining (4.238) leads to the following:

$$\begin{aligned} \prod_i [i]^{b_i} < A(T) &\iff \Delta G_{L \rightarrow R} < 0 \iff \text{reaction goes from left to right,} \\ \prod_i [i]^{b_i} > A(T) &\iff \Delta G_{L \rightarrow R} > 0 \iff \text{reaction goes from right to left,} \\ \prod_i [i]^{b_i} = A(T) &\iff \Delta G_{L \rightarrow R} = 0 \iff \text{reaction is at equilibrium.} \end{aligned} \quad (4.240)$$

The last line above is known as the *law of mass action*: it specifies the densities of the various species present at equilibrium. In practice, the concentrations $[i]$ are usually expressed as molar densities, such as moles per litre, and the law of mass action becomes

$$\prod_i [i]^{b_i} = A(T) \text{ at equilibrium.} \quad (4.241)$$

To demonstrate, examine the following reaction that involves molecules A, B, C, D :



For the temperature at which this reaction occurs, the equilibrium constant happens to be tabulated as “100 M⁵”, meaning 100 mol⁵/ℓ⁵ (recall that “ℓ” stands for “litre”). Suppose the equilibrium concentrations are

$$\begin{aligned} [A] &= 1 \text{ M (that is, 1 mol/ℓ),} \\ [B] &= 2 \text{ M,} \quad [C] = 3 \text{ M.} \end{aligned} \quad (4.243)$$

What is the concentration of D ? The law of mass action (4.241) says that at equilibrium,

$$[A]^{-2} [B]^{-1} [C]^5 [D]^3 = 100 \text{ M}^5. \quad (4.244)$$

We can rewrite this as

$$\frac{[C]^5 [D]^3}{[A]^2 [B]} = 100 \text{ M}^5. \quad (4.245)$$

This equation is a no-fuss way to remember how to apply (4.241). Clearly, the left-hand side of (4.245) is simply a denominator–numerator grouping of the terms in the reaction (4.242), with all coefficients written as powers. The right-hand side of (4.245) is the equilibrium constant. Finally, solve (4.245) for $[D]$:

$$[D]^3 = \frac{100 \text{ M}^5 [A]^2 [B]}{[C]^5} = \frac{100 \text{ M}^5 \times 1 \text{ M}^2 \times 2 \text{ M}}{3^5 \text{ M}^5} = 0.82 \text{ M}^3. \quad (4.246)$$

We arrive at $[D] = 0.94 \text{ M}$.

The interactions that make up the First Law of Thermodynamics appear widely throughout physics and chemistry. Up until now, we have treated all

manner of systems as if they were isolated; and indeed, much useful work can be accomplished in that way by applying the First and Second Laws as we have done here. Nonetheless, no system can ever be truly isolated from the rest of the world. In the next chapter, we make this dependence on the world very clear, by asking what can be said about a system that interacts with an environment about which we might know very little.



Chapter 5

The Non-Isolated System: the Boltzmann Distribution

In which we take the wider world into consideration, where a system of interest is never truly isolated. We derive the Boltzmann distribution, give examples of its use in paramagnetism and atomic energy levels, study molecular and crystal heat capacities, and point the way to quantum mechanics. We describe a “problem” with applying Boltzmann to atomic energy levels, generalise the equipartition theorem to non-isolated systems, and examine the partition function. Finally, we extend the concept of counting states to non-isolated systems, and state the principles of data-transmission theory.

Up until now, we have studied only large systems that could be treated as isolated. But statistical methods cannot necessarily be applied to *small* isolated systems; we cannot, for example, give any meaning to the temperature of a single atom. And even if a system isn’t small, it might not be isolated. But envisage a small system of interest being in contact with an environment that is so large that its parameters don’t change significantly when it interacts with the small system. This environment is often called a *reservoir* or a *bath*. Since the bath *can* be treated statistically, we can calculate quantities of interest for the small system by studying how it interacts with the bath, without needing to know any details of the bath.

When calculating quantities relating to such a system–bath pair, it can be helpful to picture a large number of identically prepared systems that each interact with their own bath, with each system–bath pair in some different and random stage of its evolution. Shown in Figure 5.1, this imagined large number of system–bath pairs is known as an *ensemble*. We encountered ensembles for the first time while studying the random walk back in Section 1.3.2. At any chosen moment, the state of each pair in the ensemble is represented by some point in phase space. We assume that if we were to plot each of these points in phase space, all at the same moment—any moment will do—then the whole set of points would be identical to the path that a *single* system–bath pair will trace out in phase space as it evolves. This idea of being able to replace the time *evolution* of one system with a time *snapshot* of an ensemble of systems is called the *ergodic assumption*: it suggests that we can convert averages over time to averages over the ensemble. Although the ergodic assumption has never been completely validated, it’s a reasonable hypothesis that is used often in statistical mechanics.

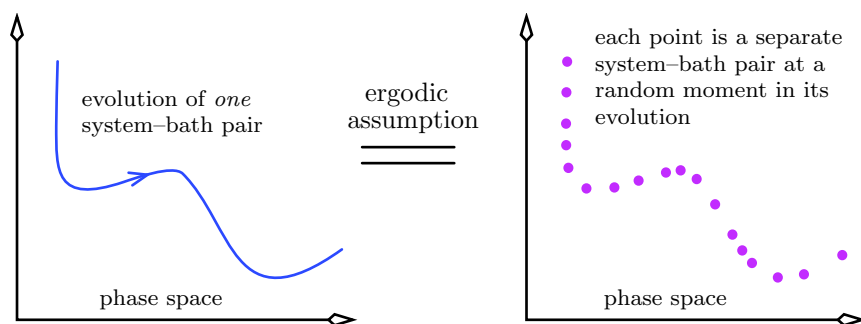


Fig. 5.1 The ergodic assumption states that the time evolution of *one* system-bath pair is representable by a snapshot of *many* system-bath pairs that have each effectively been captured at some random moment in their evolution

Ensembles are divided into three types, with names bestowed by historical usage.

Types of Ensemble

1. The **micro-canonical ensemble** has no interaction between system and bath; the system is isolated with a fixed energy. The bath is effectively not even present!
2. The **canonical ensemble** allows for thermal interactions only.
3. The **grand canonical ensemble** allows for all interactions: thermal, mechanical, and diffusive.

The word “canonical” implies a condition that is widely deemed to have especial importance. A *canon* is a body of work or set of concepts that is generally accepted as being of prime importance by the appropriate community of specialists. For example, the “canonical momentum” of lagrangian mechanics has more importance to that subject than the “kinematical momentum” $\mathbf{p} = m\mathbf{v}$. The ensemble that describes thermal-only interactions with a bath is called canonical because it describes a very common situation: for example, as you sit reading this book, you barely exchange volume or particles with your environment, but you certainly depend on this environment to help regulate your temperature. Many applications of statistical mechanics focus on this “canonical” interaction between system and bath, for which only thermal interactions are allowed. One important example is the study of the “Maxwell distribution” of the motions of gas particles in Chapter 6.

The other two ensembles that are generally considered describe systems that have either less or more interaction with their environments than the systems of the canonical ensemble, and so these other ensembles have taken on

the terms “micro-canonical” and “grand canonical”. The systems of a micro-canonical ensemble do not interact with their bath at all, and so the bath need not even be considered as being present. Rather, the micro-canonical ensemble is a description of systems that are truly isolated from the rest of the world. These systems thus each have a fixed energy. And in the case of the grand canonical ensemble, all interactions with the environment are allowed. This is the case for complex systems; a good example is the particle diffusion occurring in osmosis, where chemicals are separated using a membrane as a sieve with molecule-sized holes.

Note that the system need not be held in a separate “box” that is connected to the bath via a wall or membrane; it could be physically interspersed with the bath. Also, the system might well be a quantum state; and since particles can enter and leave a given quantum state, we must then take diffusive interactions into account, and so must use the grand canonical ensemble. We’ll treat such systems in Chapter 7.

The names of these three ensembles (particularly the canonical ensemble) are often used as an alternative to describing the assumed interaction between system and bath. Rather than say “We will assume the system interacts only thermally with the bath”, conventional phrasing might be “We will work with a canonical ensemble”. You will even sometimes encounter a somewhat bizarre phrasing such as “The canonical ensemble governs the behaviour of the system”. That phrase does not mean that a large number of imaginary systems governs the behaviour of the real system. It simply means “The system interacts only thermally with the bath”.

5.1 The Boltzmann Distribution

To begin to say something of how a system behaves when it’s in contact with a bath, we begin by calculating how likely some fluctuation in a system will be, when it contacts a bath. In particular, we require the probability $p(E_s, V_s, N_s)$ that, at some given moment, the system will have given values of energy E_s , volume V_s , and particle number N_s . The result, known as the Boltzmann distribution, is one of the central pillars of statistical mechanics.

This probability $p(E_s, V_s, N_s)$ is proportional to the number of states accessible to the system–bath pair when the system has the specified energy, volume, and particle number. And this number of states is the product of the individual numbers of states Ω_s and Ω_b that are occupiable by the system and bath individually, when the system has the specified energy, volume, and particle number:

$$p(E_s, V_s, N_s) \propto \Omega_s \Omega_b, \quad (5.1)$$

where subscripts “s” and “b” denote system and bath, respectively. To calculate Ω_s and Ω_b , we presume the system is simple enough that the number

of states Ω_s accessible to it is easily found by straightforward counting. The bath, on the other hand, is so huge that we simply cannot count the number of its accessible states Ω_b from first principles; but we *can* treat it statistically, and so can calculate Ω_b from a knowledge of its other parameters via (3.187).

When the system has energy, volume, and number of particles E_s, V_s, N_s , respectively, suppose the bath has equivalent parameters E_b, V_b, N_b . System and bath share common values of the intensive parameters T, P, μ . Then (3.187) yields

$$p(E_s, V_s, N_s) \propto \Omega_s \exp \frac{E_b + PV_b - \mu N_b}{kT}. \quad (5.2)$$

But we know only the system's parameters E_s, V_s, N_s , along with the fact that the unknown values of *total* energy, volume, and particle number (E, V, N) are all fixed. In that case, write the bath's parameters in terms of the system's parameters as

$$E_b = E - E_s, \quad V_b = V - V_s, \quad N_b = N - N_s. \quad (5.3)$$

Equation (5.2) now becomes

$$\begin{aligned} p(E_s, V_s, N_s) &\propto \Omega_s \exp \frac{E - E_s + P(V - V_s) - \mu(N - N_s)}{kT} \\ &= \Omega_s \exp \frac{E + PV - \mu N}{kT} \exp \frac{-E_s - PV_s + \mu N_s}{kT}. \end{aligned} \quad (5.4)$$

That is,

$$p(E_s, V_s, N_s) \propto \Omega_s \exp \frac{-E_s - PV_s + \mu N_s}{kT}. \quad (5.5)$$

This is the desired expression that involves only system parameters. It is the celebrated *Boltzmann distribution*, one of the central pillars of statistical mechanics. For the common case of the canonical ensemble, V_s and N_s are fixed, and can thus be absorbed into the constant of proportionality. For this case, (5.5) becomes

$$p(E_s) \propto \Omega_s \exp \frac{-E_s}{kT}. \quad (5.6)$$

5.1.1 The Exponential Atmosphere Again

In Section 4.3.1, we showed that the number density of atmospheric molecules (and equivalently, the mass density and pressure) decreases exponentially with height, assuming the temperature in the atmosphere is independent of height. (In Chapter 6, we'll relax this assumption.) An immediate application

of the Boltzmann distribution is to rederive this exponential drop in number density with height.

To do this, treat a single air molecule as a system that is in contact with the rest of the atmosphere, which thus forms the bath at temperature T . The system molecule does not exchange any appreciable volume with the bath, and nor does it exchange any particles. The energy of this molecule equals some base energy (set by the temperature) plus its potential energy mgz , where m is the molecule's mass, g is Earth's gravitational acceleration and z is the molecule's height above any reference point (say, sea level). The number density of the atmosphere, $\nu(z)$, is proportional to the probability that this system molecule will be found at height z . Thus, (5.6) produces

$$\nu(z) \propto \Omega_s \exp \frac{-mgz}{kT}. \quad (5.7)$$

If the atmosphere's molecules all have the same number Ω_s of internal states available, this number can be absorbed into the constant of proportionality, and (5.7) becomes

$$\nu(z) \propto \exp \frac{-mgz}{kT}, \quad (5.8)$$

which leads to

$$\nu(z) = \nu(0) \exp \frac{-mgz}{kT}. \quad (5.9)$$

This matches our previous result in both (4.149) [which used $\nu_0 \equiv \nu(0)$ for convenience] and (4.181). To recap, equation (4.149) was derived from the first-principles analysis of Figure 4.12, which focussed on the pressure differential in the atmosphere. Equation (4.181) was derived from knowledge of the chemical potential in (4.178), which itself came from the Sackur–Tetrode entropy expression (4.176), originally produced in Section 3.8.1 from our earlier phase-space arguments. We have now derived the number density's exponential drop-off with atmospheric height using three completely different methods. The first method used Newton's laws, and the other two used purely statistical mechanical ideas. The agreement of these wildly differing approaches supports the internal coherence of statistical mechanics.

5.2 Paramagnetism

Paramagnetism is a classic example of the Boltzmann distribution in action. The very strong magnetism of materials such as iron has been known from antiquity; but, in fact, materials can exhibit any of three different types of magnetic behaviour.

1. *Diamagnetism* is displayed by all materials. It was discovered by Michael Faraday, who found that a piece of bismuth was repelled by both the north

and south poles of a magnet. Diamagnetism is understood to be an example of *Lenz's law* in electromagnetism. Lenz's law states that when we interact magnetically with an electric circuit, the effect of all induced voltages and currents is to oppose the interaction. When we bring a magnet up to a material, we apply Lenz's law in the context of a classical model of electrons orbiting each atomic nucleus of that material. The magnetic field alters the electrons motion so as to produce a (usually very weak) magnetic field. This field opposes the field of our magnet, and the result is that the material is repelled. Diamagnetism is a fully electromagnetic phenomenon, and always produces a repulsive force.

2. *Paramagnetism* is displayed by only some materials, ones whose atoms have a configuration of electrons such that each atom has a permanent magnetic moment¹ that interacts, at most, weakly with its neighbours. Paramagnetism can be understood using statistical mechanics, as we'll show soon.
3. *Ferromagnetism* is displayed by only a few materials: principally iron, nickel, and cobalt, the rare metals gadolinium and dysprosium, and some chemical compounds. In such materials, the atoms' magnetic dipoles interact very strongly with their neighbours, with the result that the material can produce a strong permanent magnetic field.

To see how statistical mechanics describes paramagnetism, picture a set of atoms, each of which has one of two choices of spin, “up” or “down”, that arise from the configuration of electrons within the atoms. Each atom thus has some magnetic dipole moment $\boldsymbol{\mu}$. Suppose that a material composed of these atoms has a temperature T , and is placed in an external magnetic field $\mathbf{B}$. The magnetic field tries to align each magnetic dipole with itself.² The result would be a very strong magnetic field produced by the material—if it weren't for the competing action of the temperature, which acts to randomise the directions of the dipoles, and thus reduce the overall field. The question is, which effect dominates: the atomic dipoles lining up (when $\mathbf{B}$ dominates T), or those dipoles' directions being randomised (when T dominates $\mathbf{B}$)?

Suppose that the imposed magnetic field $\mathbf{B}$ acts in the z direction, whose unit basis vector is labelled $\mathbf{u}_z$. This field is then $\mathbf{B} = B\mathbf{u}_z$, where $B > 0$ is its strength. The two possible spins produce two possible magnetic dipole moments. Define:

$$\begin{aligned} \text{spin up: } \boldsymbol{\mu} &= \mu\mathbf{u}_z \equiv \boldsymbol{\mu}_\uparrow, \\ \text{spin down: } \boldsymbol{\mu} &= -\mu\mathbf{u}_z \equiv \boldsymbol{\mu}_\downarrow. \end{aligned} \tag{5.10}$$

¹ See Section 3.4.1 for an introduction to magnetic moments.

² Recall (3.60), which says the potential energy of the magnetic dipole $\boldsymbol{\mu}$ is $-\boldsymbol{\mu} \cdot \mathbf{B}$. This energy is lowest when $\boldsymbol{\mu}$ points in the same direction as $\mathbf{B}$.

The magnetic moment of N atoms is the sum of their individual magnetic moments. Hence, the expected value of the magnetic moment of N atoms is

$$\langle \boldsymbol{\mu}_1 + \cdots + \boldsymbol{\mu}_N \rangle = N \langle \boldsymbol{\mu} \rangle. \quad (5.11)$$

The expected value of the magnetic moment of a single atom is

$$\langle \boldsymbol{\mu} \rangle = p_{\uparrow} \boldsymbol{\mu}_{\uparrow} + p_{\downarrow} \boldsymbol{\mu}_{\downarrow}, \quad (5.12)$$

where $p_{\uparrow}$ is the probability of finding the atom with spin up (or, equivalently, the proportion of all the atoms that have spin up), and similarly for $p_{\downarrow}$. These probabilities are given by Boltzmann's distribution (5.6). The system that we are considering here is a single dipole, which has one state accessible to it when it has spin up and one state when it has spin down. Thus, $\Omega_s = 1$ in (5.6), and that equation becomes

$$p_{\uparrow} \propto \exp \frac{-E_{\uparrow}}{kT}, \quad p_{\downarrow} \propto \exp \frac{-E_{\downarrow}}{kT}, \quad (5.13)$$

where the energies of the up and down spins, $E_{\uparrow}, E_{\downarrow}$, are given by (3.60):

$$\begin{aligned} E_{\uparrow} &= -\boldsymbol{\mu}_{\uparrow} \cdot \mathbf{B} = -\mu \mathbf{u}_z \cdot B \mathbf{u}_z = -\mu B, \\ E_{\downarrow} &= -\boldsymbol{\mu}_{\downarrow} \cdot \mathbf{B} = \mu \mathbf{u}_z \cdot B \mathbf{u}_z = \mu B. \end{aligned} \quad (5.14)$$

In that case,

$$p_{\uparrow} \propto \exp \frac{\mu B}{kT}, \quad p_{\downarrow} \propto \exp \frac{-\mu B}{kT}. \quad (5.15)$$

For convenience, set

$$\alpha \equiv \frac{\mu B}{kT} > 0. \quad (5.16)$$

The normalised probabilities are then

$$p_{\uparrow} = \frac{e^{\alpha}}{e^{\alpha} + e^{-\alpha}}, \quad p_{\downarrow} = \frac{e^{-\alpha}}{e^{\alpha} + e^{-\alpha}}. \quad (5.17)$$

Let's check that the probabilities in (5.17) make sense. When the applied field is low or the temperature is high, $\alpha \approx 0$, and (5.17) becomes

$$p_{\uparrow} \simeq p_{\downarrow} \simeq 1/2. \quad (5.18)$$

That is, these conditions ensure that an even mix of dipoles is present. When the applied field is high or the temperature is low, $\alpha \rightarrow \infty$, and (5.17) becomes

$$p_{\uparrow} \simeq 1, \quad p_{\downarrow} \simeq 0. \quad (5.19)$$

Here, essentially all of the spins have become aligned with the field.

Equation (5.12) now becomes

$$\langle \mu \rangle = \frac{e^\alpha \mu \mathbf{u}_z + e^{-\alpha} \times -\mu \mathbf{u}_z}{e^\alpha + e^{-\alpha}} = \frac{e^\alpha - e^{-\alpha}}{e^\alpha + e^{-\alpha}} \mu \mathbf{u}_z = \text{th}(\alpha) \mu \mathbf{u}_z, \quad (5.20)$$

where “th” is the hyperbolic tangent function, also written as “tanh”. We’ll plot the tanh function shortly in Figure 5.2.

To discuss the extremes of low and high α , use the approximations

$$\text{th } x \simeq \begin{cases} x, & 0 \leq x \ll 1, \\ 1, & x \gg 1. \end{cases} \quad (5.21)$$

1. When $\alpha \ll 1$ (i.e., $\mu B \ll kT$: the applied magnetic field is relatively weak),

$$\langle \mu \rangle \simeq \alpha \mu \mathbf{u}_z = \frac{\mu B}{kT} \mu \mathbf{u}_z = \frac{\mu^2 \mathbf{B}}{kT}. \quad (5.22)$$

This expression is essentially Curie’s law, to be explained shortly.

2. When $\alpha \gg 1$ (so $\mu B \gg kT$: the applied magnetic field is relatively strong),

$$\langle \mu \rangle \simeq \mu \mathbf{u}_z = \mu_\uparrow. \quad (5.23)$$

This latter case is one of saturation: the magnetic field is so strong that essentially all of the spins have become aligned with the field.

If the material’s N atoms displace a volume V , its mean magnetic moment per unit volume is called the material’s *magnetisation* $\mathbf{M}$:

$$\mathbf{M} \equiv \frac{N \langle \mu \rangle}{V} = \nu \langle \mu \rangle, \quad (5.24)$$

where $\nu \equiv N/V$ is the particle-number density of the material. When the applied magnetic field is strong ($\mu B \gg kT$), the magnitude of the magnetisation saturates to a value of

$$M_s \equiv |\mathbf{M}|_{\alpha \rightarrow \infty} = \nu |\langle \mu \rangle|_{\alpha \rightarrow \infty} \stackrel{(5.23)}{=} \nu \mu. \quad (5.25)$$

For all values of α , we can express the magnitude M of the magnetisation in terms of this saturation value:

$$\frac{M}{M_s} = \frac{\nu |\langle \mu \rangle|}{\nu \mu} \stackrel{(5.20)}{=} \text{th } \alpha. \quad (5.26)$$

This relative magnetisation is plotted in Figure 5.2. For relatively weak magnetic fields ($\alpha \ll 1$), we have

$$\frac{M}{M_s} \simeq \alpha = \frac{\mu B}{kT}. \quad (5.27)$$

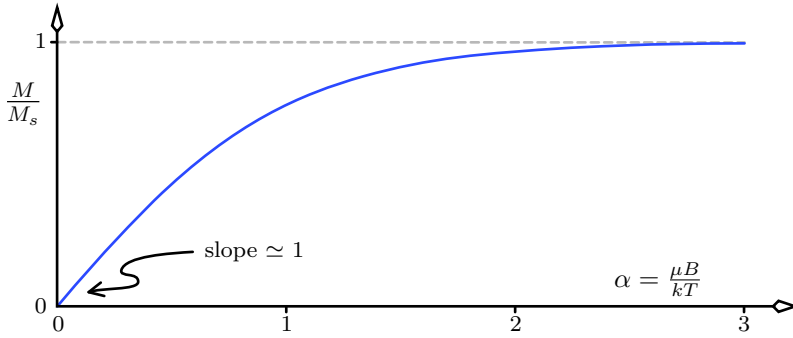


Fig. 5.2 The magnitude M of a material's magnetisation, expressed as a fraction of its saturation value M_s

This inverse dependence on temperature is known as *Curie's law*.

Atomic magnetic moments are usually expressed in terms of a useful basic amount known as a *Bohr magneton*:

$$\begin{aligned}
 1 \text{ Bohr magneton} &\equiv \frac{\text{proton charge} \times \hbar}{2 \times \text{electron mass}} \simeq \frac{1.602 \overset{-19}{\times} 6.626 \overset{-34}{\text{J}} / (2\pi)}{2 \times 9.11 \overset{-31}{\text{kg}}} \text{ J/T} \\
 &\simeq 9.27 \times 10^{-24} \text{ J/T}.
 \end{aligned} \tag{5.28}$$

In the laboratory, a typical magnetic moment is about one Bohr magneton and a strong magnetic field is around one tesla. Then, at room temperature,

$$\alpha = \frac{\mu_B}{kT} \simeq \frac{9.27 \overset{-24}{\times} 1}{1.381 \overset{-23}{\times} 298} \simeq 0.0023. \tag{5.29}$$

It's clear from this that most lab experiments in magnetisation are confined to the low- α regime.

5.3 Energy Levels, States, and Bands

The Boltzmann distribution (5.5) gives us the probability that a system will be found to have some specified value E_s of energy: that is, to be in some given *energy level*. As we saw with the case of paramagnetism, these energy levels are often quantised. You can read the word “level” here to denote “possibility”: if a system has three energy levels, then it can have any of three possible energies. These energy levels refer simply to amounts of energy, and are not states of the system. Each energy level may “contain” several states, meaning the system has available to it several states that each have that level's assigned value of energy. When energy levels are very close together so as to be well

approximated as a continuum, they are called an energy *band*. We discuss these bands more fully in Section 8.4.

Equation (5.5) is a proportionality; we normalise it by dividing by the sum over all energy levels of the system. As is conventional, we'll drop the subscript “ s ”, because it's understood that we are focussing on the system and not the bath; and instead, we'll index the system's energy, volume, and particle number by n —but will explicitly denote that n refers to a level. The normalised version of (5.5) is then

$$p_{\text{level } n} \equiv p(E_n, V_n, N_n) = \frac{\Omega_n}{Z} \exp \frac{-E_n - PV_n + \mu N_n}{kT}, \quad (5.30)$$

where (with n now a summation variable over all energy levels)

$$Z \equiv \sum_{\text{all levels } n} \Omega_n \exp \frac{-E_n - PV_n + \mu N_n}{kT}. \quad (5.31)$$

The normalisation Z is called the system's *partition function*, and is discussed in detail in Section 5.9.

The canonical ensemble gives the simplest example of the Boltzmann distribution: we wrote it in (5.6). For that energy-only version of the exponential, (5.30) and (5.31) become

$$p_{\text{level } n} = \frac{\Omega_n}{Z} \exp \frac{-E_n}{kT}, \quad \text{where } Z \equiv \sum_{\text{all levels } n} \Omega_n \exp \frac{-E_n}{kT}. \quad (5.32)$$

Note that the sum is over energy levels: level n has energy E_n and contains Ω_n states, and $p_{\text{level } n}$ is the probability that the system is found to have energy E_n : meaning, is found to be in *any one* of those Ω_n states. For example, when Ω_n is large, there are many such states, and then $p_{\text{level } n}$ can be large.

Here, we must be mindful of the terminology, and so we appeal to the quantum mechanics of the hydrogen atom as a classic example to sort out the language and notation. We will simplify the situation by considering only the energy of the atom, and not its volume. Quantum theory labels the hydrogen atomic states with a set of quantum numbers conventionally denoted (n, ℓ, m, s_z) . Solving Schrödinger's equation in detail (you'll find that calculation in introductory books on quantum mechanics) shows that the hydrogen atom's quantised internal energy is determined only by n , and is $E_n = -13.6 \text{ eV}/n^2$. This number essentially describes how far the bound electron is likely to be found from the proton. The quantum number ℓ relates to the electron's orbital angular momentum, and m relates to the z component of this angular momentum:

- n can take on values $1, 2, 3, \dots$;
- for each n : ℓ can take on values $0, 1, \dots, n-1$;
- for each ℓ : m can take on values $-\ell, \dots, \ell$.

energy level 3	<u>$300\ 1/2\rangle$ $300\ -1/2\rangle$ $\dots$ $322\ -1/2\rangle$</u>	$\Omega_3 = 18$ states, each with energy $E_3 = -13.6/3^2$ eV
energy level 2	<u>$200\ 1/2\rangle$ $200\ -1/2\rangle$ $\dots$ $211\ -1/2\rangle$</u>	$\Omega_2 = 8$ states, each with energy $E_2 = -13.6/2^2$ eV
energy level 1	<u>$100\ 1/2\rangle$ $100\ -1/2\rangle$</u>	$\Omega_1 = 2$ states, each with energy $E_1 = -13.6$ eV

Fig. 5.3 The energy states and levels of the hydrogen atom. Each level has $\Omega_n = 2n^2$ states

The last quantum number, s_z , relates to the electron's spin angular momentum. Experiments indicate that for each of these sets of (n, l, m) , the number s_z can take on either of the values $\pm 1/2$.

These rules on the allowed quantum numbers yield $2n^2$ sets of (n, ℓ, m, s_z) values that describe solutions to Schrödinger's equation. Each of these sets of values describes a unique (quantum) state, conventionally denoted " $|n, \ell, m, s_z\rangle$ ". The number of quantum states with energy E_n is then

$$\Omega_n \equiv \Omega(E_n) = 2n^2. \quad (5.33)$$

Put another way, energy level n consists of $\Omega_n = 2n^2$ states.

For example, the lowest energy level or "ground level" ($n = 1$) of a hydrogen atom can be populated by atoms in either of the two states described by the set of quantum numbers

$$(n, \ell, m, s_z) = (1, 0, 0, 1/2) \text{ or } (1, 0, 0, -1/2); \quad (5.34)$$

hence, this level has two states available, or $\Omega_1 = 2$. The first excited level ($n = 2$) has atoms with

$$(n, \ell, m) = (2, 0, 0) \text{ or } (2, 1, -1) \text{ or } (2, 1, 0) \text{ or } (2, 1, 1). \quad (5.35)$$

Each of these four sets of quantum numbers is then supplemented with either of two electron spins $s_z = \pm 1/2$, making a total of $\Omega_2 = 8$ states. The first three levels with their states are shown in Figure 5.3.

What is the partition function Z here? Remember that we are treating a simplified situation where the volume of the atom is assumed fixed. (In

Section 5.7, we'll include the volume in detail.) Apply (5.32) by performing a sum over energy levels, where each level has $\Omega_n = 2n^2$ states, and $\beta \equiv 1/(kT)$ is a widely used shorthand:

$$\begin{aligned} Z &\equiv \sum_{\text{all levels } n} \Omega_n e^{-\beta E_n} = \Omega_1 e^{-\beta E_1} + \Omega_2 e^{-\beta E_2} + \Omega_3 e^{-\beta E_3} + \dots \\ &= 2e^{-\beta E_1} + 8e^{-\beta E_2} + 18e^{-\beta E_3} + \dots \end{aligned} \quad (5.36)$$

Some care with the conventional notation is needed. Aside from the probability of finding a system to be in some energy *level*, we often require the probability of finding the system to be in some *given state* at that energy level. Here, instead of calculating $p_{\text{level } n}$ (the chance that the system is found in *any* state at level n), we calculate $p_{\text{state } n}$, the chance that the system is found to be in a *specific state* n with energy E_n :

$$p_{\text{state } n} = \frac{1}{Z} e^{-\beta E_n}. \quad (5.37)$$

Compare this with $p_{\text{level } n}$ in (5.32). For the hydrogen atom, $p_{\text{state } 1}$ and $p_{\text{state } 2}$ denote the probabilities of finding the atom in either of the two states that comprise the ground level (see Figure 5.3). These states' equal energies are now denoted E_1 and E_2 , *both* of which equal -13.6 eV . The next energy level's eight states are found with equal probabilities $p_{\text{state } 3}, \dots, p_{\text{state } 10}$ and energies $E_3, \dots, E_{10}$, where

$$E_3 = E_4 = \dots = E_{10} = -13.6/2^2 \text{ eV}, \quad (5.38)$$

and so on for higher energy levels. We are now treating each of the terms in (5.32) as a set of $\Omega_n = 2n^2$ separate terms to be summed. In that case, the normalisation for (5.37) will be

$$Z \equiv \sum_{\text{all states } n} e^{-\beta E_n}. \quad (5.39)$$

Refer to Figure 5.4 for this alternative language of enumerating the states. The partition function is

$$Z \equiv \sum_{\text{all states } n} e^{-\beta E_n} = e^{-\beta E_1} + e^{-\beta E_2} + e^{-\beta E_3} + \dots \quad (5.40)$$

Note that this is the *same* Z that appeared in (5.32) and (5.36)—only the meaning of “ E_n ” differs across the two choices of notation. The simpler look of (5.37) and (5.39)—no factor of Ω_n present—leads to their being used extremely widely throughout statistical mechanics. We must only be aware that these equations refer to *states*, not *levels*.

This language of states and levels can be confusing. It is common to find energy level n referred to as a *state* with *degeneracy* $\Omega_n = 2n^2$. Hence, the

energy level 3	<u> 3 0 0 ¹/₂⟩ 3 0 0 ⁻¹/₂⟩ ... 3 2 2 ⁻¹/₂⟩</u>	$\Omega_3 = 18$ states, with energies $E_{11}, \dots, E_{28}$, all = $-13.6/3^2$ eV
energy level 2	<u> 2 0 0 ¹/₂⟩ 2 0 0 ⁻¹/₂⟩ ... 2 1 1 ⁻¹/₂⟩</u>	$\Omega_2 = 8$ states, with energies $E_3, \dots, E_{10}$, all = $-13.6/2^2$ eV
energy level 1	<u> 1 0 0 ¹/₂⟩ 1 0 0 ⁻¹/₂⟩</u>	$\Omega_1 = 2$ states, with energies E_1, E_2 , both = -13.6 eV

Fig. 5.4 An alternative way of enumerating the states of the hydrogen atom that focusses attention on individual states in the partition-function sum. Compare this with Figure 5.3

ground energy level ($n = 1$) is also called “the ground state with its 2-fold degeneracy”. The first excited energy level ($n = 2$) is also called “the first excited state with its 8-fold degeneracy”. Energy level n with its $\Omega_n = 2n^2$ states becomes “the $n - 1^{\text{th}}$ excited state, with $2n^2$ -fold degeneracy”. This language is summarised in Table 5.1.

Table 5.1 A comparison of alternative phrasing used to describe the occupation of energy levels

Language used here	Alternative language
Ground level with 2 states	Ground state with 2-fold degeneracy
1 st excited level with 8 states	1 st excited state with 8-fold degeneracy
2 nd excited level with 18 states	2 nd excited state with 18-fold degeneracy
⋮	⋮
$n - 1^{\text{th}}$ excited level with $2n^2$ states	$n - 1^{\text{th}}$ excited state with $2n^2$ -fold degeneracy

5.4 Hydrogen Energy Levels

Stepping up in complexity from paramagnetism, the hydrogen atom is our next example of the practical details of applying the Boltzmann distribution. We streamline the discussion by examining only monatomic hydrogen, rather than a molecule of H₂. The monatomic case is simpler, but how realistic

is it? Recall the basic quantum mechanical result that energy level n of the hydrogen atom has energy

$$E_n = -13.6 \text{ eV}/n^2. \quad (5.41)$$

The energy required to excite a hydrogen atom from the ground level ($n = 1$) to the first excited level ($n = 2$) is thus

$$E_2 - E_1 = (-13.6/2^2 - -13.6) \text{ eV} = 10.2 \text{ eV}. \quad (5.42)$$

This is rather more than the 4.5 eV of energy required to split a hydrogen molecule into its two atoms. It follows that in the interesting case of a gas of hydrogen having a spread of energy levels, they will already all be dissociated into atoms.

Thus, picture a volume of monatomic hydrogen gas. As two atoms collide, the kinetic energy of one can excite the other into a higher energy level. An excited hydrogen atom will de-excite very quickly to drop back down to the ground level; but in equilibrium with collisions continuously occurring, we suppose that the numbers of hydrogen atoms occupying each energy level remain constant with time.

What are the various fractions of atoms in each energy level? We can answer this question by setting our system to be a *single* “system hydrogen atom” that we can separate from the rest—so it’s *distinguishable*—while the rest of the gas is the bath with which that atom continually interacts via collisions, which can bump the system atom from one energy level to another. If, for example, the system atom has a 10% chance of being found in its ground level ($n = 1$), then we infer that 10% of all the atoms will be found with energy E_1 . We are not interested in the system atom’s kinetic energy, and so will work the problem in that atom’s (non-inertial) frame.

We then apply (5.30) and (5.31) to our system atom. The system contains just one particle, so $N_n = 1$ for all n in the normalisation sum (5.31). This allows for the μ part of the exponential to be factored outside the sum in (5.31). The result is

$$p_n \equiv p(E_n) = e^{\mu/(kT)} \frac{\Omega_n}{Z} \exp \frac{-E_n - PV_n}{kT}, \quad (5.43)$$

with normalisation

$$Z = e^{\mu/(kT)} \sum_n \Omega_n \exp \frac{-E_n - PV_n}{kT}. \quad (5.44)$$

We can cancel out the $e^{\mu/(kT)}$ with the same term in Z in (5.43) to obtain simpler expressions:

$$p_n = \frac{\Omega_n}{Z'} \exp \frac{-E_n - PV_n}{kT}, \text{ with } Z' = \sum_n \Omega_n \exp \frac{-E_n - PV_n}{kT}. \quad (5.45)$$

We will use (5.45), and drop the prime from Z' when referring to it later. (Remember, Z' is just a normalisation that we are generically calling Z .)

Here is a traditional question that aims to give some idea of applying the Boltzmann distribution:

A star's surface is modelled as being fully composed of hydrogen atoms.³ Its temperature is such that the number of atoms in the ground level is one million times the number of atoms in the first excited level. What is this temperature?

We can answer this question by using (5.45) to compute the ratio of probabilities p_2/p_1 , setting this to equal 10^{-6} , and then solving for T . We refer to (5.33) for the numbers of states Ω_n at energy levels $n = 1$ and 2. The normalisation Z' cancels out in the ratio of probabilities. Also, on a first attempt, we will assume that $V_1 = V_2$, and so the volume factors will also cancel out. So, we need deal only with *relative* probabilities p_n^{rel} that don't include the normalisation. The ratio p_2/p_1 then equals $p_2^{\text{rel}}/p_1^{\text{rel}}$, where

$$p_n^{\text{rel}} = \Omega_n \exp \frac{-E_n}{kT}. \quad (5.46)$$

We proceed to write

$$\frac{1}{10^6} = \frac{p_2}{p_1} = \frac{p_2^{\text{rel}}}{p_1^{\text{rel}}} = \frac{2 \times 2^2 \exp \frac{-E_2}{kT}}{2 \times 1^2 \exp \frac{-E_1}{kT}} = 4 \exp \frac{-E_2 + E_1}{kT}. \quad (5.47)$$

Solve this for T , obtaining

$$T = \frac{-E_2 + E_1}{-k \ln(4 \frac{6}{\equiv})} = \frac{13.6 \text{ eV} \times \left(\frac{1}{2^2} - \frac{1}{1^2} \right)}{-k \ln(4 \frac{6}{\equiv})} = \frac{13.6 \times 1.602 \frac{-19}{\equiv} \times -3/4}{-1.381 \frac{-23}{\equiv} \ln(4 \frac{6}{\equiv})} \text{ K} \\ \simeq 7800 \text{ K}. \quad (5.48)$$

Given that we set the volume terms in (5.45) to be equal, can we have confidence in this value of temperature? It turns out that we cannot. We'll explore the idea of including the volume terms later in Section 5.7.

³ Stars are composed of a plasma of atoms whose electrons have been stripped away, which makes this problem artificial. But it serves to illustrate the way to apply the Boltzmann distribution.

5.5 Excitation Temperature

It's clear from (5.5) that kT plays a central role in the Boltzmann distribution: it determines how sharply the probability of atoms occupying higher energy levels decreases as the energies of those levels increase. Its value at room temperature, $kT \simeq 1/40$ eV, is useful to memorise. This value forms a good rule of thumb for whether an appreciable number of a system's entities will be excited into higher energy levels. To see this in more detail, use the fact that for systems that are not stars with their ultra-high pressure, the pressure/volume contribution to the Boltzmann exponential is negligible for the first few energy levels. In particular, with the standard shorthand $\beta \equiv 1/(kT)$,

$$\frac{p_2}{p_1} \simeq \frac{\Omega_2 e^{-\beta E_2}}{\Omega_1 e^{-\beta E_1}} = \frac{\Omega_2}{\Omega_1} e^{-\beta(E_2 - E_1)}. \quad (5.49)$$

For hydrogen atoms, the energy “distance” from the ground to the first excited level is $E_2 - E_1 = 10.2$ eV, in which case (with $\Omega_n = 2n^2$)

$$\frac{p_2}{p_1} \simeq \frac{2 \times 2^2}{2 \times 1^2} e^{-10.2 \text{ eV}/(kT)}. \quad (5.50)$$

This 10.2 eV is so much larger than room temperature's $kT \simeq 1/40$ eV that we can see immediately that the amount of excitation will be vanishingly small. The equipartition theorem says that a typical energy exchanged in any interaction between the hydrogen atoms is kT . A much larger energy of around 10 eV might well be exchanged, but that will happen only *exceedingly* rarely. Hence, almost all of the atoms will be in the ground state.

This example suggests defining the *excitation temperature* T_e of a general system, such that

$$kT_e \equiv E_2 - E_1, \quad (5.51)$$

where E_1 and E_2 are the energies of its ground and first excited levels. T_e is the temperature at which an appreciable number of particles are beginning to occupy the first excited level, with a very few also beginning to climb into higher energy levels. For our gas of monatomic hydrogen, the excitation temperature is

$$T_e = \frac{E_2 - E_1}{k} = \frac{10.2 \times 1.602 \overset{-19}{\text{eV}}}{1.381 \overset{-23}{\text{J/K}}} \text{ K} \simeq 118,300 \text{ K}. \quad (5.52)$$

Only at stellar temperatures must we allow for the possibility that some hydrogen atoms populate excited energy levels.

5.6 Diatomic Gases and Heat Capacity

A molecule can *always* translate sideways, and thus store energy in this motion. It can also rotate and vibrate; but these two types of motion turn out to be quantised, and so, like an atom's energy levels, give rise to their own excitation temperatures. In this section, we'll derive these excitation temperatures for molecular rotation and vibration. We'll see that after translation, rotation is a molecule's preferred method of storing energy, followed by vibration.

Recall that we defined the ratio of heat capacities $\gamma \equiv C_P/C_V = 1 + 2/\nu$ for an ideal gas in (4.24). Some representative values of γ are given for room temperature in Table 5.2. These values begin at about $\gamma = 1 + 2/5$ for the lighter diatomic gases such as NO and HCl. Such gas molecules thus have $\nu = 5$ quadratic energy terms. Because gas molecules can always translate in space, three of these modes must be translational. We'll show soon that the remaining two modes are rotational rather than vibrational. The lighter diatomic molecules are thus rigid rotors, or “dumb-bells”.

In the case of a non-rigid rotor, we can expect vibration to contribute two more quadratic energy terms to make $\nu = 7$ in total, giving $\gamma = 1 + 2/7 \simeq 1.29$. The values of γ in Table 5.2 suggest that the molecules appearing there can rotate at room temperature; additionally, the heavier molecules chlorine and bromine can vibrate by some small amount at room temperature. The even-heavier iodine molecule is almost fully non-rigid, and vibrates freely.

We'll soon see that the heat capacity of a given gas increases through a series of levels with increasing temperature, as the quadratic energy terms change from being purely translational at low temperatures, to translational + rotational at medium temperatures, and finally to translational + rotational + vibrational at high temperatures. Let's investigate this using the Boltzmann distribution and quantum mechanics.

Table 5.2 Values of $\gamma \equiv C_P/C_V$ measured at room temperature for some common molecules with a range of masses from light to heavy, followed by the inferred values of their number of quadratic energy terms ν

Molecule:	NO	HCl	Cl ₂	Br ₂	I ₂
Molar mass (g):	30	36	71	160	254
Measured γ :	1.40	1.41	1.36	1.32	1.30
$\gamma \approx$	$12/5$	$12/5$	$(12/5 \text{ to } 12/7)$		$12/7$
$\nu \approx$	5	5	(5 to 7)		7

5.6.1 Quantised Rotation

Before we investigate the quantisation of rotation, one observation should be made. Spinning objects are described by their *moment of inertia*. The moment of inertia is often described as relating to rotation about a given axis. In fact, this is not quite right; the moment of inertia is always defined relative to a *point*, not an axis. To see what is happening here, realise that the angular *momentum* vector $\mathbf{L}$ of a spinning object is generally not parallel to the object's angular *velocity* vector $\boldsymbol{\omega}$. These two vectors are related by the moment-of-inertia tensor I (or “inertia tensor” for short), using the following formalism:

$$\mathbf{L} = I\boldsymbol{\omega}. \quad (5.53)$$

This expression makes no reference to coordinates; it merely says that I operates on the vector $\boldsymbol{\omega}$ to give the vector $\mathbf{L}$. What does “operates on” mean? In analogy, picture rotating a vector $\mathbf{a}$ to generate a new vector $\mathbf{b}$, writing this operation as “ $\mathbf{b} = R\mathbf{a}$ ”. This is not a multiplication; the “ $R\mathbf{a}$ ” simply denotes the operation of “rotation” on $\mathbf{a}$. But recall the discussion in Section 1.9.3, where we emphasised the importance of distinguishing an object from its *representation* in some system of coordinates. If we *coordinatise* the vectors $\mathbf{a}, \mathbf{b}$ as 3×1 matrices $[\mathbf{a}], [\mathbf{b}]$ (usually called “column vectors”), then *it turns out* that the rotation operator R can be written as a matrix $[R]$, and the rotation operation can be written as a matrix multiplication

$$[\mathbf{b}] = [R][\mathbf{a}]. \quad (5.54)$$

The rotation R is a *tensor*, and its coordinate representation is a *matrix*. It turns out, from a classical-mechanics analysis, that the same idea holds for I : this too is a tensor, and can be coordinatised as a matrix. Hence, when we coordinatise $\mathbf{L}$ and $\boldsymbol{\omega}$ to make 3×1 matrices $[\mathbf{L}]$ and $[\boldsymbol{\omega}]$, (5.53) becomes a matrix multiplication

$$[\mathbf{L}] = [I][\boldsymbol{\omega}]. \quad (5.55)$$

Classical mechanics tells us that when an object is set spinning and left to spin with no outside intervention, it will always quickly settle into spinning evenly about any one of three orthogonal axes known as *principal axes*.⁴ Principal axes always coincide with any symmetry axes present; but even the most asymmetric objects will always have three principal axes that are mutually orthogonal.⁵ For this case of spin about a principal axis, $\mathbf{L}$ will be *parallel*

⁴ By “spinning evenly”, we mean that the object's angular momentum does not change with time; hence, no torques are involved, which means the object does not stress its bearings. Most examples of everyday spin fall into this category; an extreme example of an engineering requirement to spin about a principal axis is the turbine in a jet engine.

⁵ A sphere can spin evenly about any of its diameters, and so might be said to have an infinite number of principal axes. But that is a special case of high symmetry.

to ω . In that case, $\mathbf{L}$ equals a *number* times ω ; and this number is then usually called the “moment of inertia about the spin axis”. It is the *eigenvalue* of the inertia tensor corresponding to the *eigenvector* ω . (Equivalently, we can say it is the eigenvalue of the inertia *matrix* corresponding to the eigenvector $[\omega]$.) But note that the inertia matrix is *not* equal to this number times the identity matrix.

For the case of spin about a principal axis “ α ”, we’ll write this angular-momentum eigenvalue as I_α , and use the result of classical mechanics:

$$I_\alpha = \sum_i m_i r_i^2, \quad (5.56)$$

where m_i is the mass of the i^{th} particle and r_i is the perpendicular distance of that particle from the α axis.

With this observation in place, we return to our gas molecules. Quantum mechanics says that the rotation of a rigid body with a “moment of inertia I about some axis” (we really mean the eigenvalue here, as pointed out above) is quantised about that axis, with quantum number $\ell = 0, 1, 2, \dots$. The corresponding energy levels are

$$E_\ell = \frac{\ell(\ell+1)\hbar^2}{2I}. \quad (5.57)$$

At temperature T , some gas particles are excited into higher levels. If we ignore the pressure/volume term in the Boltzmann probability—at least for the first few levels—the population of level ℓ compared to the ground level ($\ell = 0$) is [with $E_0 = 0$ from (5.57)]

$$\frac{N_\ell}{N_0} = \frac{\Omega_\ell e^{-\beta E_\ell}}{\Omega_0 e^{-\beta E_0}} = \frac{\Omega_\ell}{\Omega_0} e^{-\beta E_\ell}. \quad (5.58)$$

Each level is described by the quantum numbers (ℓ, m) where $\ell = 0, 1, 2, \dots$ quantifies the energy and m can take any integer value in the range $-\ell, \dots, \ell$. That is, each energy level is associated with $\Omega_\ell = 2\ell + 1$ states; again, this number is usually called the degeneracy of “state” ℓ . In that case,

$$\frac{N_\ell}{N_0} = (2\ell + 1) \exp \frac{-\ell(\ell+1)\frac{\hbar^2}{2I}}{kT}. \quad (5.59)$$

Almost analogously to (5.51), define “the characteristic temperature of (the onset of) rotation” as T_R , where

$$kT_R \equiv \frac{\hbar^2}{2I}. \quad (5.60)$$

[I say “almost analogously” because (5.60) is defined for convenience. Whereas (5.51) defines kT_e to be the energy difference between the first excited atomic

level and ground level, (5.60) essentially defines kT_R to be *half* the energy difference between the first excited rotational level and ground level. The factor of one half is not important, as these characteristic temperatures are defined arbitrarily in order to be convenient.] It follows that

$$\frac{N_\ell}{N_0} = (2\ell + 1) \exp \frac{-\ell(\ell + 1)T_R}{T}. \quad (5.61)$$

At relatively low temperatures T , we will have $T \ll T_R$, and so $N_\ell/N_0 \approx 0$. The rotational levels are thus “frozen out”: the bath simply has too little energy (characterised by kT) to excite rotation. This is because the rotational energy level spacing (characterised by kT_R) is comparatively large.

But at relatively high temperatures, $T \gg T_R$, and the ratio N_ℓ/N_0 is then nonzero for many values of ℓ : rotational energy levels are now well populated. The rotational energy level spacing (characterised by kT_R) is now small compared to kT . With so many rotational levels able to be accessed, a gas of such molecules well and truly has rotational quadratic energy terms, and they will each appear in the equipartition theorem with the usual value of $kT/2$ per particle.

The rotational energy levels in (5.57) depend on the axis about which the object spins. But a study of the measurements of $\gamma \equiv C_P/C_V$ back in Table 5.2 suggests that diatomic molecules have $\nu = 5$ quadratic energy terms. Three of these are translational, which makes for just *two* rotational terms, not three (we’ll show soon that these terms are not vibrational for the light gases in the table). Why should that be? Classically, the answer could be said to lie in the moment of inertia. We must focus on principal axes here, because gas molecules—being free—will rotate about those axes. These axes consist of the interatomic axis that joins the two atoms in the molecule, and any two axes orthogonal to this.

Rotation About a Non-Interatomic Axis

Figure 5.5 shows a classical rigid rotor formed from two masses that are momentarily lying along the x axis. They can spin in the xy plane around the z axis through their centre of mass, which lies at the origin. To analyse rotation about this non-interatomic axis, an acceptable approximation assumes that the masses forming the rotor are localised as points. These masses are m_1 , distance r_1 from the origin, and m_2 , distance r_2 from the origin, and so are a distance $D = r_1 + r_2$ apart. If the rotation of this system is quantised as per (5.57), what is its characteristic temperature of rotation T_R ?

Equation (5.60) writes T_R in terms of the moment of inertia I_z for rotation about the z axis, where the centre of mass lies on the z axis (at the origin). The moment of inertia is then

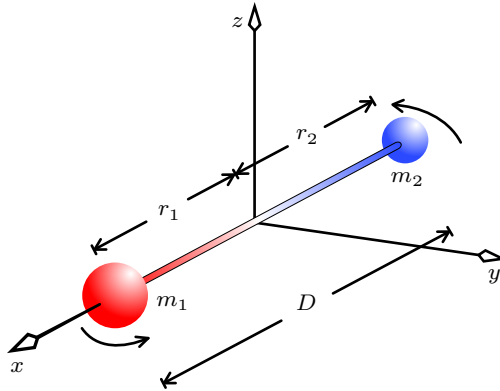


Fig. 5.5 Modelling a diatomic molecule as a rigid rotor. The “connecting rod” has no mass

$$I_z \stackrel{(5.56)}{=} m_1 r_1^2 + m_2 r_2^2. \quad (5.62)$$

What are r_1 and r_2 ? The centre of mass $\mathbf{r}_{\text{CM}}$ of a set of masses m_i with positions $\mathbf{r}_i$ is given by

$$\mathbf{r}_{\text{CM}} \equiv \sum_i m_i \mathbf{r}_i / \sum_i m_i. \quad (5.63)$$

With the centre of mass at the origin, this implies that

$$m_1 r_1 - m_2 r_2 = 0. \quad (5.64)$$

Also, we know that

$$r_1 + r_2 = D. \quad (5.65)$$

These two linear simultaneous equations in r_1, r_2 are easily solved, yielding

$$\begin{Bmatrix} r_1 \\ r_2 \end{Bmatrix} = \frac{D}{m_1 + m_2} \begin{Bmatrix} m_2 \\ m_1 \end{Bmatrix}. \quad (5.66)$$

Substituting these into (5.62) produces

$$I_z = \mu D^2, \quad \text{where} \quad \frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2}. \quad (5.67)$$

μ is called the *reduced mass* of the system, because it is less than each of m_1 and m_2 . The reduced mass is a mathematically useful quantity that simplifies expressions in various areas of classical mechanics.⁶

⁶ I find it remarkable that whereas the reduced mass is universally accepted as useful despite not existing in its own right as separate matter, the idea of *relativistic mass*

To demonstrate, we determine the characteristic temperature of rotation for CO about an axis orthogonal to its interatomic axis, given that the C and O atoms are a distance of $D = 0.112\text{ nm}$ apart. The masses of an atom of carbon and an atom of oxygen are $m_C = 12\text{ g}/N_A$ and $m_O = 16\text{ g}/N_A$, respectively. Equation (5.60) produces

$$\begin{aligned} T_R &= \frac{\hbar^2}{2I_z k} = \frac{\hbar^2}{2\mu D^2 k} = \frac{\hbar^2}{2D^2 k} \left(\frac{1}{m_C} + \frac{1}{m_O} \right) \\ &= \frac{(1.0546 \times 10^{-34})^2 \times 6.022 \times 10^{23}}{2 \times (0.112 \times 10^{-9})^2 \times 1.381 \times 10^{-23}} \left(\frac{1}{0.012} + \frac{1}{0.016} \right) \text{ K} = 2.8 \text{ K}. \end{aligned} \quad (5.68)$$

This low value of temperature implies that rotation about the non-interatomic axis occurs very easily. At room temperature, a gas of CO molecules has many rotational levels occupied, because $k \times \text{room temperature} \gg kT_R$. This gas thus obeys the equipartition theorem very well.

Rotation About the Interatomic Axis

For the masses in Figure 5.5, we can try a classical calculation of the moment of inertia I_x for rotation about the x axis, meaning about the line joining the masses. This moment of inertia will be the sum of the moments of four “objects”: the carbon and oxygen nuclei, and the carbon and oxygen electron clouds, all modelled as spherical.

Call on the result that the moment of inertia of a sphere of mass M and radius R relative to its centre, for rotation about a diameter, is $\frac{2}{5}MR^2$. [We re-iterate the footnote just before (5.62): the moment of inertia is a matrix, but its effect mathematically for this symmetrical situation is as if it were the number $\frac{2}{5}MR^2$.] The carbon electron cloud has a radius of about 0.08 nm , and that of oxygen is 0.07 nm .⁷ These clouds have masses of 6 and 8 electron masses, respectively. The (equivalent scalar value of the) moment

in special relativity is sometimes criticised due to supposedly not being “really there”. Yet both of these mass types are—at least in principle—physically measurable, and each of them gives a little simplification to the relevant mathematics, even though neither is absolutely necessary, physically or mathematically. Many things that are not *necessary* are still *useful*.

⁷ In fact, most atoms are of a similar size; the average radius over the whole periodic table is around 0.12 nm , not much more than that of “smaller” atoms such as carbon and oxygen. If it surprises you that a hydrogen atom has almost the same size as a uranium atom, realise that an atom with a large atomic number has many electrons in inner shells, and these inner shells then have very small radii due to the pull of the large nuclear charge. These inner shells shield the outer electrons from feeling a strong nuclear attraction. Thus, the outer electrons effectively “see” a nucleus with a similar charge to that seen by the outer electrons of an element with a low atomic number.

of inertia of carbon's electron cloud is then (with an electron's mass being about 9.11×10^{-31} kg)

$$\begin{aligned} I_x &= \frac{2}{5} MR^2 = \frac{2}{5} \times 6 \times 9.11 \overset{-31}{=} \times (0.08 \overset{-9}{=})^2 \text{ kg m}^2 \\ &\simeq 1.4 \times 10^{-50} \text{ kg m}^2. \end{aligned} \quad (5.69)$$

Oxygen's electron cloud has approximately the same moment of inertia. What about the nuclear contributions? Each nucleon can be treated as though it were a ball of radius 1.5 fm, with the nucleus well modelled by these balls being packed together to form a sphere. The number of nucleons N of these packed balls makes a sphere of radius R (where R now relates to the nucleus, not the electron cloud), with volume

$$\begin{aligned} \frac{4}{3} \pi R^3 &= N \times \text{volume of one nucleon} \\ &= N \times \frac{4}{3} \pi \times (1.5 \text{ fm})^3. \end{aligned} \quad (5.70)$$

It follows that the nuclear radius is $R = N^{1/3} \times 1.5 \text{ fm}$, and so

$$\begin{aligned} \text{radius of carbon nucleus} &= 12^{1/3} \times 1.5 \text{ fm}, \\ \text{radius of oxygen nucleus} &= 16^{1/3} \times 1.5 \text{ fm}. \end{aligned} \quad (5.71)$$

The moment of inertia of the carbon nucleus is thus (with a nucleon's mass being about 1.67×10^{-27} kg)

$$\begin{aligned} I_x &= \frac{2}{5} MR^2 = \frac{2}{5} \times 12 \times 1.67 \overset{-27}{=} \times \left(12^{1/3} \times 1.5 \overset{-15}{=} \right)^2 \text{ kg m}^2 \\ &\simeq 9.5 \times 10^{-56} \text{ kg m}^2. \end{aligned} \quad (5.72)$$

The value for oxygen's nucleus is about $1.5 \times 10^{-55} \text{ kg m}^2$. In summary:

constituent of CO	I_x (units of kg m^2)
C electron cloud	1.4×10^{-50}
O electron cloud	1.4×10^{-50}
C nucleus	9.5×10^{-56}
O nucleus	1.5×10^{-55}

Despite their far larger masses, the nuclei are vastly smaller than the electron clouds; and radius appears as a *square* in the moment of inertia. So, the nuclear moments of inertia end up being 100,000 times smaller than the electron-cloud moments. The total moment of inertia I_x of CO is the sum of the above tabulated values, or about $2.8 \times 10^{-50} \text{ kg m}^2$. The characteristic temperature of rotation about the interatomic axis is then

$$T_R = \frac{\hbar^2}{2I_x k} = \frac{(1.0546 \overset{-34}{\text{}})^2}{2 \times 2.8 \overset{-50}{\text{}} \times 1.381 \overset{-23}{\text{}}} \text{ K} \simeq 14,000 \text{ K}. \quad (5.73)$$

A very hot room is needed to impart energy kT_R to the dumb-bell, and so it's clear that rotation about the interatomic axis is well and truly frozen out at room temperature. In fact, quantum mechanics says that no rotation *at all* can occur about the interatomic axis, because the spherical symmetry of the nuclei and electron clouds implies that they have no features or “handles”, so to speak, that could be used to impart any rotation to them. Thus, rotation about the interatomic axis is actually frozen out at *all* temperatures.

5.6.2 Quantised Vibration

All molecules have three translational quadratic energy terms. Rotational modes are a little more difficult to access. But the energy levels corresponding to the internal vibration of a molecule are the most difficult to populate. To see why, we apply some basic quantum mechanics to a diatomic molecule, treating it as a one-dimensional harmonic oscillator of frequency f . The solution to Schrödinger's equation for a one-dimensional harmonic oscillator is found in all texts on introductory quantum mechanics. The oscillator's vibrations turn out to be quantised into energy levels. The n^{th} level has energy E_n :

$$E_n = (n + 1/2) hf, \quad \text{where } n = 0, 1, 2, \dots \quad (5.74)$$

Each energy level has just one state: $\Omega_n = 1$ for all n , because there is only one way that a harmonic oscillator can oscillate. The ground level of vibration corresponds to $n = 0$. (n is standard notation; don't confuse it with the atomic quantum number n that has a ground-level value of 1 for, say, the hydrogen atom that we met in Section 5.3.) The pressure/volume term PV_n is presumably negligible—at least for simple molecules and low energy levels—so we ignore it, and write the relative populations of vibrational level n and the ground level as

$$\frac{N_n}{N_0} = \frac{\Omega_n e^{-\beta(n+1/2)hf}}{\Omega_0 e^{-\beta 1/2 hf}} = \exp \frac{-nhf}{kT}. \quad (5.75)$$

Define the characteristic temperature T_V of (the onset of) vibration by

$$kT_V \equiv hf. \quad (5.76)$$

Then,

$$\frac{N_n}{N_0} = e^{-nT_V/T}. \quad (5.77)$$

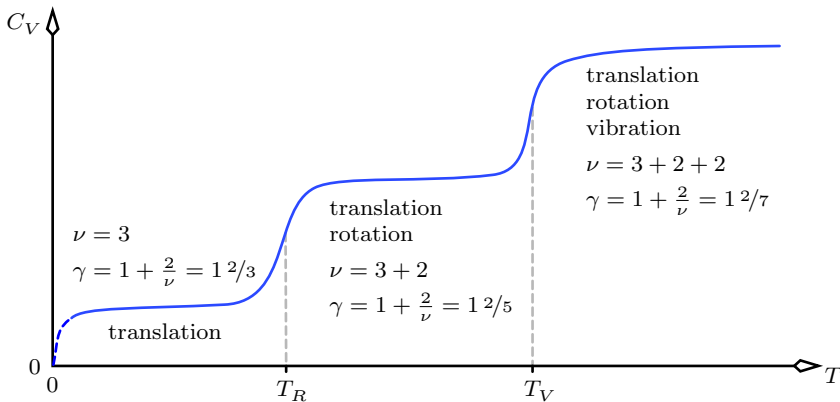


Fig. 5.6 A schematic showing the contributions to the heat capacity of a gas of diatomic molecules as a function of temperature. Initially, only translational quadratic energy terms contribute. At modestly higher temperatures, rotation comes to life, and at much higher temperatures, vibration is allowed

At low temperatures, $T \ll T_V$, and so $N_n/N_0 \approx 0$: the vibrational levels are frozen out. Another way of seeing this is to note that in the regime $T \ll T_V$, it's clear that $kT \ll kT_V = hf$. This means that the thermal energy kT available to impart vibration to the oscillators is much less than the vibrational energy-level spacing hf , and so vibrational levels are not populated.

At relatively high temperatures, $T \gg T_V$, and then N_n/N_0 is nonzero for many values of n , meaning the vibrational energy levels are well populated. Now $kT \gg kT_V = hf$, meaning the thermal energy kT is much greater than the vibrational energy-level spacing hf . Vibrational levels are now well occupied. A gas of such molecules has vibrational quadratic energy terms, and can be treated using the equipartition theorem.

Simple molecules can oscillate at typical frequencies of hundreds of terahertz; for example, HCl has a frequency of $f \simeq 89$ THz. For this, (5.76) yields

$$T_V = \frac{hf}{k} = \frac{6.626 \overset{-34}{\text{J}\cdot\text{s}} \times 89 \overset{12}{\text{THz}}}{1.381 \overset{-23}{\text{J/K}}} \text{ K} \simeq 4300 \text{ K}. \quad (5.78)$$

We see that at room temperature and well beyond, simple molecules don't vibrate. It's as if their atoms are connected by stiff springs that need a very high energy to be set into vibration.

Figure 5.6 is a schematic showing how translation, rotation, and vibration contribute quadratic energy terms to a diatomic molecule's heat capacity, as a function of temperature. For the gaseous phase, translation is always allowed, even at the lowest temperatures. Rotation enters at slightly higher temperatures, and vibration comes to life last of all.

5.7 Another Look at the Hydrogen Atom

In Section 5.4, we attempted to calculate the relative populations of the ground and first excited levels of a hydrogen atom by assuming that the atom's volume doesn't change when it jumps from level 1 to level 2. The equivalent assumption (at least, in this case), that the volume term in Boltzmann's distribution can be ignored, is actually very common. But is it valid?

To check, let's investigate the higher energy-level populations that result from ignoring the volume term. We write out the first few relative probabilities p_n^{rel} from (5.46), using

$$p_n^{\text{rel}} = 2n^2 \exp \frac{13.6 \text{ eV}}{n^2 kT} = 2n^2 \exp \frac{13.6 \times 1.602 \overset{-19}{\text{eV}}}{n^2 \times 1.381 \overset{-23}{\text{J/K}} \times 7800} \simeq 2n^2 \exp \frac{20.27}{n^2}. \quad (5.79)$$

The first two relative probabilities are in the ratio of $10^6:1$, as required:

$$p_1^{\text{rel}} = 1270 \times 10^6, \quad p_2^{\text{rel}} = 1270. \quad (5.80)$$

The next few are smaller, but they start to increase at $n = 6$:

$$p_3^{\text{rel}} = 171, \quad p_4^{\text{rel}} = 114, \quad p_5^{\text{rel}} = 112, \quad p_6^{\text{rel}} = 126. \quad (5.81)$$

Populations of higher energy levels are increasing more drastically:

$$p_{10}^{\text{rel}} = 245, \quad p_{100}^{\text{rel}} = 20,041. \quad (5.82)$$

What has happened? It's clear from (5.79) that for large n , the relative probability p_n^{rel} tends toward $2n^2$. This number blows up as n increases—meaning the normalisation $\sum_n p_n^{\text{rel}}$ is not defined. Maybe levels with large values of n are too difficult to access in practice? In practice, possibly; but not in principle, when we invoke the fundamental postulate of statistical mechanics. This says that in equilibrium, every state is just as accessible as every other state; and as a result, we expect Boltzmann's theory to be a self-consistent, complete description of state populations. Also, *all* of a hydrogen atom's energy levels lie within 13.6 eV of its ground state, and since the number of states at level n is $2n^2$, higher levels have significantly more states available for the atom to occupy.

Perhaps the fact that we have essentially omitted the volume term from (5.45) has caused the problem. After all, quantum mechanics shows that the size of an atom depends on its energy level. The situation is rather like having a room full of light balloons, with each containing a chemical mixture that reacts to jostling of the balloon by releasing gas that inflates the balloon somewhat; this mixture then re-absorbs some amount of gas after a short period, deflating the balloon somewhat. As the balloons fly into each other, work is continually being performed to increase their volumes for a short

period. At equilibrium, the room contains fixed numbers of balloons of various sizes.

Excited hydrogen atoms are larger than ground-state atoms, and so the PV_n in (5.45) might well be important to retain. The atom's volume V_n at energy level n can have any of a spread of values, because the electron's position is spread out quantum mechanically; but the volume has an *expected value* of $\langle V_n \rangle = 4/3 \pi \langle r_n^3 \rangle$, where r_n is the atom's radius at level n . A complication is that the expected value of any power of r_n in the quantum mechanics of the hydrogen atom depends on the atom's orbital angular-momentum quantum number ℓ . This means we must work with this ℓ -dependent radius, $r_{n\ell}$. For example,

$$\langle r_{n\ell} \rangle = \frac{a_0}{2} [3n^2 - \ell(\ell + 1)] , \quad (5.83)$$

where a_0 is the *Bohr radius*, about 0.0529 nm. To calculate $\langle r_n \rangle$, or $\langle r_n^3 \rangle$ in our case, we then must average over all values of ℓ —an exercise that is beginning to look excessive. On the other hand, the expected value of the *reciprocal* of r does *not* depend⁸ on ℓ :

$$\langle 1/r_n \rangle = \frac{1}{a_0 n^2} . \quad (5.84)$$

There is some latitude in these definitions that lets us simplify the problem, and so we'll use the simpler expression, opting for $\langle 1/r_n \rangle$ in place of $\langle r_n \rangle$. We then *define* the volume of the atom to be

$$V_n \equiv 4/3 \pi \langle 1/r_n \rangle^{-3} \simeq 4a_0^3 n^6 . \quad (5.85)$$

The expression (5.46) for the relative probability to occupy level n must now be modified to include the volume term:

$$p_n^{\text{rel}} = 2n^2 \exp \frac{-E_n - PV_n}{kT} . \quad (5.86)$$

⁸ It is strange and interesting that the expected value of the reciprocal of radial distance should be a simpler expression than the expected value of the radial distance itself. A similar phenomenon appears in classical orbital mechanics: when solving Newton's equation for the motion of two bodies orbiting their centre of mass, one of the early steps in the mathematics switches from the radial variable r to $1/r$, which renders the equation of motion tractable—and the resulting expressions are simpler when written in terms of $1/r$. For example, using polar coordinates r, θ in the orbital plane and centred on the Sun, the inverse distance of a planet is a displaced sinusoid: $1/r = \alpha + \beta \cos(\theta - \theta_0)$ for constants α, β, θ_0 . Another example is the *vis-viva equation* that relates the planet's speed v (relative to the Sun) to its distance r from the Sun: $v^2 = \mu(2/r - 1/a)$ for a constant μ , where a the orbit's semi-major axis length. Note that $1/r$, both for the hydrogen atom and in orbital mechanics, is the form of the relevant potential (Coulomb for hydrogen, gravity for orbits). So, perhaps the form of the potential is what determines the useful quantity. In the case of a spring, its stretch x is more useful than $1/x$. Nonetheless, the potential energy of a spring depends on x^2 rather than x .

Let's use SI units throughout, to write

$$\begin{aligned} \frac{-E_n - PV_n}{kT} &= \frac{\frac{13.6 \times 1.602 \times 10^{-19}}{n^2} - P \times 4 \times (5.29 \times 10^{-11})^3 n^6}{1.381 \times 10^{-23} T} \\ &= \frac{1}{T} \left(\frac{1.58 \times 10^{-5}}{n^2} - P n^6 \times 4.29 \times 10^{-8} \right). \end{aligned} \quad (5.87)$$

The first term in the parentheses of (5.87) is the usual $1/n^2$ that we saw in (5.79). But a second term is now present (as a result of including the volume), proportional to n^6 . For large n , this second term will dominate the first term in the parentheses, and the whole expression will be a negative number with large absolute value. This will have the effect of reducing p_n^{rel} in (5.86) to zero for large n : and *that* allows the normalisation series $\sum_n p_n^{\text{rel}}$ to be well defined, which thus solves our problem. Equation (5.47) is then modified to be (again with P and T in SI units)

$$\begin{aligned} \frac{1}{10^6} = \frac{p_2^{\text{rel}}}{p_1^{\text{rel}}} &= \frac{2 \times 2^2 \exp \left(\frac{1.58 \times 10^{-5}}{2^2 T} - \frac{P \times 2^6 \times 4.29 \times 10^{-8}}{T} \right)}{2 \times 1^2 \exp \left(\frac{1.58 \times 10^{-5}}{1^2 T} - \frac{P \times 1^6 \times 4.29 \times 10^{-8}}{T} \right)} \\ &\simeq 4 \exp \frac{-118,300 - P \times 2.70 \times 10^{-6}}{T}. \end{aligned} \quad (5.88)$$

[This is the same 118,300 K that appeared in (5.52).] Initially, when we ignored the volume term, our treatment was equivalent to writing down (5.88) without the pressure term. And certainly,

$$4 \exp \frac{-118,300}{7800} \simeq 1.0 \times 10^{-6}, \quad (5.89)$$

which is (5.48) again. But it seems that we must now solve (5.88) for *pressure* as well as for temperature. A star's pressure and temperature are interrelated, and (5.88) should incorporate this fact by replacing P with the appropriate function of T . But that is a path for astrophysicists to pursue, and so we will finish the discussion with two representative solutions:⁹

$$\begin{aligned} P &= 10^{11} \text{ Pa} \simeq 10^6 \text{ Earth atmospheres,} & T &= 25,600 \text{ K;} \\ P &= 10^{12} \text{ Pa} \simeq 10^7 \text{ Earth atmospheres,} & T &= 185,500 \text{ K.} \end{aligned} \quad (5.90)$$

⁹ In calculating these, I've included a few more significant figures than are present in the text.

An appropriate model must give meaningful mathematics, but should we have really included a PV_n mechanical work term above, as opposed to other mechanical work terms involving electric or magnetic fields? (Recall the discussion in Section 3.4.1.) The answer to this can only depend on the applicability of the model being used. Historically, the PV_n term has *not* been included in discussions of the hydrogen atom, with the resulting awkward infinite normalisation being avoided through the use of *relative* probabilities and by avoiding scenarios that involve high energy levels. Certainly, the system involved in the derivation of the Boltzmann distribution is assumed to be much smaller than the bath; but this assumption breaks down when an atom gets so large that it is no longer able to fit inside the star. This suggests that the normalisation series be truncated at some value of n . These ideas show that whether or not we include PV_n , we must modify the original model of the atom to interact with its environment in some other way. The mathematics must continue to make sense, and infinite series fail that criterion. But the physics should also make sense, and star-sized atoms don't fit into the assumptions of the Boltzmann distribution.

5.8 Equipartition for a System Contacting a Thermal Bath

We introduced the equipartition theorem in Section 3.5 for an *isolated system* that we had split into two subsystems, and showed that each quadratic energy term contributes an energy of $1/2 kT$ to the total. Now, we ask: is there an equivalent quantity or theorem for a system in contact with a bath? The bath gives rise to continuous fluctuations in the system, and, as a result, neither system nor bath have a fixed energy. Instead, we'll investigate the *mean* energy that each of its quadratic energy terms contributes to the system.

The mean internal energy associated with any particular coordinate u is

$$\langle E_u \rangle \stackrel{(1.157)}{=} \int_0^\infty E_u p(E_u) dE_u, \quad (5.91)$$

where

$$\begin{aligned} p(E_u) dE_u &= \text{probability that the system has energy in } E_u \text{ to } E_u + dE_u \\ &= \left[\begin{array}{c} \text{probability that the system} \\ \text{is in a state with energy } E_u \end{array} \right] \times \left[\begin{array}{c} \text{number of states} \\ \text{in } E_u \text{ to } E_u + dE_u \end{array} \right]. \end{aligned} \quad (5.92)$$

Restrict attention to the canonical ensemble—meaning we assume the system interacts only thermally with the bath. The probability that the system occupies a state with energy E_u is then proportional to $e^{-\beta E_u}$, where, as usual, $\beta \equiv 1/(kT)$. As in Sections 3.5 to 3.5.2, where we first studied the

equipartition theorem, we restrict consideration to quadratic energy dependence: $E_u = bu^2$ for some positive constant b . Given that the internal energy depends on the *square* of u , we need only consider $u \geq 0$. (In fact, if we do insist on treating negative values of u separately from positive values, we will arrive at the same result in what follows.)

Note: despite the resemblance of “ bu^2 ” to kinetic energy $\frac{1}{2}mv^2$, don’t confuse u with speed v when more than one dimension is being used. u is a coordinate, so when dealing with velocity, we must set u to be a *component* of the velocity vector: v_x, v_y , or v_z .

We define the number of states in the energy interval $E_u \rightarrow E_u + dE_u$ to be the number of states in the coordinate interval $u \rightarrow u + du$, and so that number of states is proportional to du . Hence, (5.92) becomes

$$p(E_u) dE_u = A e^{-\beta E_u} du \quad (5.93)$$

for some normalisation constant A . Equation (5.91) then becomes

$$\langle E_u \rangle = \int_0^\infty bu^2 A e^{-\beta bu^2} du. \quad (5.94)$$

We might consider calculating this integral using (1.98), but the result will still contain the normalisation A , whose value must then be found by evaluating and setting the integral of (5.93) equal to one (because the integral of the probability equals one). Alternatively, we can avoid calculating A by integrating (5.94) by parts, where the following under-brackets denote each part [just as we did in (1.96)]:

$$\langle E_u \rangle = \int_0^\infty \underbrace{u}_{\frac{-u}{2\beta}} \underbrace{bu A e^{-\beta bu^2}}_{\left[\frac{-u}{2\beta} A e^{-\beta bu^2} \right]_0^\infty} du = \left[\frac{-u}{2\beta} A e^{-\beta bu^2} \right]_0^\infty + \frac{1}{2\beta} \int_0^\infty A e^{-\beta bu^2} du. \quad (5.95)$$

The brackets term $[\dots]_0^\infty$ in (5.95) equals zero. The integral in the last term is just $\int_0^\infty p(E_u) dE_u$ [recall (5.93)], which we know equals one. The mean energy associated with coordinate u for one particle then becomes

$$\boxed{\langle E_u \rangle = kT/2.} \quad (5.96)$$

This is the generalisation of the equipartition theorem to a non-isolated system. Each quadratic energy term now contributes an *average* value of $\frac{1}{2}kT$ to the internal energy of a particle.

To demonstrate, suppose we have a box of ideal-gas point particles in thermal equilibrium with a bath at temperature T . How fast are the particles moving? They will have a spread of speeds, but we’ll be content to calculate the root-mean-square or “rms” average, by using the equipartition theorem as applied to the system–bath pair. With three quadratic energy terms (all translational), the average energy of a gas particle will be $\langle E \rangle = \frac{3}{2}kT$. The

energy of any one particle of mass m is all kinetic, and so is $E = 1/2 mv^2$. Hence,

$$\langle 1/2 mv^2 \rangle = 3/2 kT, \quad (5.97)$$

in which case

$$\langle v^2 \rangle = 3kT/m = 3RT/M_{\text{mol}}, \quad (5.98)$$

where R is the gas constant and M_{mol} is the molar mass. The rms speed of the particles is then

$$v_{\text{rms}} \equiv \sqrt{\langle v^2 \rangle} = \sqrt{3kT/m} = \sqrt{3RT/M_{\text{mol}}}. \quad (5.99)$$

The “average” molecular speed can be defined in other ways, and we’ll study these further in Section 6.3 ahead. In particular, we’ll calculate v_{rms} in a different way in that section.

Equipartition for Non-Quadratic Energy Terms

Our earliest discussions of counting microstates (in Chapter 2) were confined to variables that contributed quadratically to a system’s energy: momentum p gave kinetic energy $p^2/(2m)$ [alternatively, velocity v gave kinetic energy $1/2 mv^2$], a spring’s stretch x gave potential energy $1/2 kx^2$, and so on. These terms led to calculations of the volumes of ellipsoids. But what can be said about *non-quadratic* energy terms? The above calculation of $\langle E_u \rangle$ from (5.91)–(5.96) extends easily to the case when the energy is proportional to u^α , where both u and α are positive, and α may or may not equal 2. To see how, note that (5.91)–(5.93) are unchanged in the general case of u^α , but (5.94) becomes

$$\langle E_u \rangle = \int_0^\infty bu^\alpha Ae^{-\beta bu^\alpha} du. \quad (5.100)$$

The rest of the calculation proceeds as in (5.95), but with the exponent 2 replaced by α . That is, we evaluate (5.100) by parts, where the underbrackets in what follows denote those parts:

$$\begin{aligned} \langle E_u \rangle &= \int_0^\infty \underbrace{u}_{\text{u}} \underbrace{bu^{\alpha-1} Ae^{-\beta bu^\alpha}}_{\text{bu}^{\alpha-1} Ae^{-\beta bu^\alpha}} du \\ &= \left[\frac{-u}{\alpha\beta} Ae^{-\beta bu^\alpha} \right]_0^\infty + \frac{1}{\alpha\beta} \int_0^\infty Ae^{-\beta bu^\alpha} du = \frac{1}{\alpha\beta}. \end{aligned} \quad (5.101)$$

That is,

$$\langle E_u \rangle = kT/\alpha. \quad (5.102)$$

This is the extension of (5.96) to non-quadratic energy terms. Of course, it reduces to (5.96) when $\alpha = 2$.

5.8.1 Fluctuation of the System's Energy

In Section 1.3.1, we found the relative fluctuation $\sigma_n/\bar{n}$ to be expected of the number of molecules n in some specified part of a room, and showed that it is typically *extremely* small. This suggests that the energy of those particles will also fluctuate by similarly minuscule amounts in the canonical ensemble, meaning a system whose only interaction with the environment is thermal. We know this to be true from our everyday experience: after all, the temperature of the air in a room tends to be stable when conditions outside the room are stable.

Let's show this by calculating $\sigma_E/\bar{E}$ for a canonical ensemble. Begin with the basic expression for the variance of any quantity, (1.47):

$$\sigma_E^2 = \langle E^2 \rangle - \langle E \rangle^2, \quad (5.103)$$

where $\langle E \rangle$ means the same as $\bar{E}$, but notationally, the brackets " $\langle \cdot \rangle$ " are a little more flexible here than an overbar. Write the following for *states* n of the canonical ensemble:

$$\begin{aligned} \langle E \rangle &= \sum_n p_n E_n \stackrel{(5.37)}{=} \frac{1}{Z} \sum_n e^{-\beta E_n} E_n, \\ \langle E^2 \rangle &= \sum_n p_n E_n^2 = \frac{1}{Z} \sum_n e^{-\beta E_n} E_n^2, \end{aligned} \quad (5.104)$$

where Z is given by (5.39):

$$Z \equiv \sum_{\text{all states } n} e^{-\beta E_n}. \quad (5.105)$$

Now notice that

$$\begin{aligned} \frac{\partial Z}{\partial \beta} &= \frac{\partial}{\partial \beta} \sum_n e^{-\beta E_n} = - \sum_n e^{-\beta E_n} E_n = -Z \langle E \rangle, \\ \frac{\partial^2 Z}{\partial \beta^2} &= \frac{\partial}{\partial \beta} \left(- \sum_n e^{-\beta E_n} E_n \right) = \sum_n e^{-\beta E_n} E_n^2 = Z \langle E^2 \rangle. \end{aligned} \quad (5.106)$$

These last two expressions enable (5.103) to be written as

$$\sigma_E^2 = \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2} - \frac{1}{Z^2} \left(\frac{\partial Z}{\partial \beta} \right)^2 = \frac{\partial}{\partial \beta} \left(\frac{1}{Z} \frac{\partial Z}{\partial \beta} \right) \stackrel{(5.106)}{=} \frac{-\partial \langle E \rangle}{\partial \beta}. \quad (5.107)$$

But

$$-\frac{\partial}{\partial \beta} = \frac{-dT}{dT} \frac{\partial}{\partial T} = kT^2 \frac{\partial}{\partial T}. \quad (5.108)$$

This lets us rewrite (5.107) as

$$\sigma_E^2 = kT^2 \frac{\partial \langle E \rangle}{\partial T}. \quad (5.109)$$

Now, recall that (5.96) tells us that an ideal gas with ν quadratic energy terms has a mean energy of $\langle E \rangle = \nu NkT/2$. Hence,

$$\sigma_E^2 = kT^2 \times \nu Nk/2 = k^2 T^2 \nu N/2. \quad (5.110)$$

It follows that

$$\frac{\sigma_E}{\langle E \rangle} = \frac{\sqrt{k^2 T^2 \nu N/2}}{\nu NkT/2} = \sqrt{\frac{2}{\nu N}}. \quad (5.111)$$

Here, we see the characteristic $1/\sqrt{N}$ for relative fluctuations that we first encountered in (1.55). For one mole of an ideal diatomic gas at any temperature (thus $N = 6.022 \times 10^{23}$ particles and $\nu = 5$ quadratic energy terms), the relative fluctuation in energy is

$$\frac{\sigma_E}{\langle E \rangle} = \sqrt{\frac{2}{5 \times 6.022^{23}}} \simeq 10^{-12}. \quad (5.112)$$

The fact that the system's energy typically fluctuates by only one part in 10^{12} demonstrates the extreme energy stability of a large system in contact with a bath.

5.9 The Partition Function in Detail

In equation (5.106), expected values of energy and energy squared were written using derivatives of the partition function Z . Indeed, Z can be used to calculate other system parameters. We encountered this idea previously in (3.175), which expressed the *intensive* variable I_n in terms of the partial derivative of entropy S with respect to the conjugate *extensive* variable X_n . We will demonstrate shortly that the partition function Z is closely related to entropy S . The key idea here is that Z tends to be easier to calculate than S , and so we may wish to replace S with Z where possible.

Recall the start of Section 5.1, where we considered the numbers of states available to system and bath, Ω_s and Ω_b , respectively, when their energies, volumes, and particle numbers were E_s, V_s, N_s and E_b, V_b, N_b . The number of states accessible to the system–bath pair was

$$\Omega_{sb}(E_s, V_s, \dots, E_b, V_b, \dots) = \Omega_s \Omega_b. \quad (5.113)$$

The system–bath combination has fixed energy, volume, and number of particles E, V, N . The system might not have its own well-defined temperature (it might be just one atom), but the bath is effectively always at some temperature, pressure, and chemical potential T, P, μ . For the bath, then, we can write the integrated First Law:

$$E_b = TS_b - PV_b + \mu N_b. \quad (5.114)$$

The bath's number of available states is then

$$\begin{aligned} \Omega_b &= \exp \frac{S_b}{k} = \exp \frac{E_b + PV_b - \mu N_b}{kT} \\ &= \exp \frac{E - E_s + P(V - V_s) - \mu(N - N_s)}{kT}. \end{aligned} \quad (5.115)$$

Equation (5.113) becomes

$$\begin{aligned} \Omega_{sb} &= \Omega_s \exp \frac{E - E_s + P(V - V_s) - \mu(N - N_s)}{kT} \\ &= \exp \left[\frac{E + PV - \mu N}{kT} \right] \Omega_s \exp \frac{-E_s - PV_s + \mu N_s}{kT}. \end{aligned} \quad (5.116)$$

The total number of states available to the system–bath pair is the sum of Ω_{sb} over all values of E_s :

$$\Omega_{sb, \text{tot}} = \sum_s \Omega_{sb} \stackrel{(5.116)}{=} \exp \left[\frac{E + PV - \mu N}{kT} \right] Z, \quad (5.117)$$

where we have summed over all energy levels E_s of the system to produce the partition function Z , following (5.31). The entropy of the system–bath pair is then

$$S_{sb} \simeq k \ln \Omega_{sb, \text{tot}} = \frac{E + PV - \mu N}{T} + k \ln Z. \quad (5.118)$$

The system and bath have mean parameters $\bar{E}_s, \bar{E}_b, \bar{V}_s$, and so on, where $E = \bar{E}_s + \bar{E}_b$, and similarly for volume and particular number. These enable the system–bath entropy in (5.118) to be written as

$$S_{sb} = \frac{\bar{E}_s + \bar{E}_b + P(\bar{V}_s + \bar{V}_b) - \mu(\bar{N}_s + \bar{N}_b)}{T} + k \ln Z. \quad (5.119)$$

But remember that the entropy S_{sb} of the system–bath pair is the sum of the system and bath entropies S_s and S_b :

$$S_{sb} = S_s(\bar{E}_s, \dots) + S_b(\bar{E}_b, \dots)$$

$$= S_s(\bar{E}_s, \dots) + \frac{\bar{E}_b + P\bar{V}_b - \mu\bar{N}_b}{kT}. \quad (5.120)$$

Equations (5.119) and (5.120) combine to give us the entropy of the system:

$$S_s(\bar{E}_s, \dots) = \frac{\bar{E}_s + P\bar{V}_s - \mu\bar{N}_s}{T} + k \ln Z. \quad (5.121)$$

This expression doesn't refer to the bath parameters; so we will drop the "system" subscript, to write our final expression for the entropy of the system:

$$S = \frac{\bar{E} + P\bar{V} - \mu\bar{N}}{T} + k \ln Z. \quad (5.122)$$

A quick check on (5.122) can be made in the zero-temperature limit, when the system parameters become E_0, V_0, N_0 and its number of states is Ω_0 . Its partition function has a single term in the sum (5.31), so that (5.122) becomes

$$\begin{aligned} S_s(E_0, V_0, N_0) &= \lim_{T \rightarrow 0} \frac{E_0 + PV_0 - \mu N_0}{T} + k \ln \left[\Omega_0 \exp \frac{-E_0 - PV_0 + \mu N_0}{kT} \right] \\ &= \lim_{T \rightarrow 0} \frac{E_0 + PV_0 - \mu N_0}{T} + k \ln \Omega_0 + \frac{-E_0 - PV_0 + \mu N_0}{T} \\ &= k \ln \Omega_0, \end{aligned} \quad (5.123)$$

as expected.

Equation (5.122) was derived assuming that energy, volume, and particle number were being exchanged with the bath. Often, this is not so: the system might be interacting only thermally with the bath. In that case, the volume and particle-number terms vanish from the previous discussion, and (5.122) becomes

$$S(\bar{E}) = \bar{E}/T + k \ln Z, \quad (5.124)$$

again with all parameters relating to the system only, and where mean values are understood. We'll encounter this expression for entropy again at the start of Section 5.10.

Equation (5.122) [or perhaps (5.124)] gives an easier way to calculate the entropy of a system, when compared with the somewhat laborious approach we followed in Chapter 2. To demonstrate this, let's use (5.124) to calculate the entropy of an ideal gas of point particles, for both the distinguishable and identical-classical cases. Before doing so, recall the approach from Section 2.4, where we calculated the volume of the $3N$ -dimensional hypersphere (2.42) to arrive at the number of states for distinguishable particles, (2.50), and its identical-classical counterpart, (2.84). These two numbers of states were rewritten for convenience in (3.144), and led to the entropies in (3.145) and (3.146). We wish to recalculate these entropies using the partition-function

approach of (5.124). Our gas will have a fixed volume V and a fixed number of particles N .

Begin by rearranging (5.124), to produce

$$Z = \Omega \exp \frac{-\bar{E}}{kT}. \quad (5.125)$$

Here, we see that Z is a kind of weighted number of states for the system. That allows us to refer back to (2.25) to write

$$Z = \int_{-\infty}^{\infty} e^{-\bar{E}/(kT)} \frac{dx_1 \dots dz_N dp_{x_1} \dots dp_{z_N}}{h^{3N}}. \quad (5.126)$$

The gas's mean energy $\bar{E}$ is all kinetic, and given by (2.42):

$$\bar{E} = \frac{p_{x_1}^2}{2m} + \frac{p_{y_1}^2}{2m} + \frac{p_{z_1}^2}{2m} + \dots + \frac{p_{x_N}^2}{2m} + \frac{p_{y_N}^2}{2m} + \frac{p_{z_N}^2}{2m}. \quad (5.127)$$

The partition function is then

$$\begin{aligned} Z &= \frac{1}{h^{3N}} \int_{-\infty}^{\infty} \exp \frac{-p_{x_1}^2 - \dots - p_{z_N}^2}{2mkT} dx_1 \dots dz_N dp_{x_1} \dots dp_{z_N} \\ &= \frac{V^N}{h^{3N}} \left[\int_{-\infty}^{\infty} \exp \frac{-u^2}{2mkT} du \right]^{3N} \stackrel{(1.91)}{=} \frac{V^N (2\pi mkT)^{3N/2}}{h^{3N}}. \end{aligned} \quad (5.128)$$

Now apply (5.124), setting $\bar{E} = 3/2 NkT$ from the equipartition theorem. The entropy for distinguishable particles is then

$$\begin{aligned} S^{\text{dist}} &= \bar{E}/T + k \ln Z = 3/2 Nk + k \ln \frac{V^N (2\pi mkT)^{3N/2}}{h^{3N}} \\ &= Nk \left[\frac{3}{2} + \ln \frac{V(2\pi mkT)^{3/2}}{h^3} \right] \\ &= Nk \left[\frac{3}{2} + \ln V + \frac{3}{2} \ln \frac{2\pi mkT}{h^2} \right]. \end{aligned} \quad (5.129)$$

This agrees with (3.145). If the particles are identical classical, divide the number of states by $N!$: that is, since Z is a kind of weighted number of states, divide it by $N! \approx (N/e)^N$ to obtain the identical-classical partition function:

$$Z^{\text{ic}} = Z(e/N)^N. \quad (5.130)$$

Then,

$$S^{\text{ic}} = \bar{E}/T + k \ln Z^{\text{ic}} = \bar{E}/T + k \ln \left[Z \left(\frac{e}{N} \right)^N \right]$$

$$\begin{aligned}
&= S^{\text{dist}} + Nk(1 - \ln N) \\
&= Nk \left[\frac{5}{2} + \ln \frac{V}{N} + \frac{3}{2} \ln \frac{2\pi mkT}{h^2} \right], \tag{5.131}
\end{aligned}$$

which agrees with (3.146). This calculation of the ideal-gas entropy via the partition function with (5.124) involved only a gaussian integral over all space, as opposed to the discussion of the hypersphere volume in Section 2.4. This simplicity of use is why the partition function tends to replace Ω and Ω_{tot} in statistical mechanics.

Using Z to Calculate Intensive Variables of the System

We began this section by stating that Z can be used to calculate more than just the system's mean energy and mean-square energy. We show here how it's used to calculate the value of any intensive variable. Recall the First Law as written in (3.173), from which (3.175) followed easily:

$$I_n = -T \left(\frac{\partial S}{\partial X_n} \right)_{E \text{ and all other variables}}. \tag{5.132}$$

Suppose we replace S here with the expression in (5.124):

$$I_n = -T \frac{\partial}{\partial X_n} (\bar{E}/T + k \ln Z). \tag{5.133}$$

For classical systems, the equipartition theorem says $\bar{E}/T = \nu Nk/2$ —which is a constant, and thus has no dependence on X_n . Hence, (5.133) becomes

$$I_n = -kT \frac{\partial \ln Z}{\partial X_n}. \tag{5.134}$$

This partial derivative is taken with energy and all extensive variables other than X_n held fixed. For an example when $I_n = -P$ and $X_n = V$, consider an ideal gas. For both the distinguishable and the identical-classical cases (the latter dividing Z by $N!$), (5.128) yields

$$\ln Z = N \ln V + \text{terms involving } T \text{ and } N. \tag{5.135}$$

Equation (5.134) then gives us

$$-P = -kT \frac{\partial}{\partial V} (N \ln V + \text{terms in } T \text{ and } N) = \frac{-kTN}{V}, \tag{5.136}$$

which reduces to $PV = NkT$, as expected.

The Helmholtz energy $F \equiv E - TS$, from Section 3.14.1, is sometimes used in this context. Begin with the generalised version of (3.190):

$$dF = -S dT + \sum_n I_n dX_n. \quad (5.137)$$

It follows immediately that

$$I_n = \left(\frac{\partial F}{\partial X_n} \right)_{T, \text{ all other } X_i}. \quad (5.138)$$

But recall (5.124): $S(\bar{E}) = \bar{E}/T + k \ln Z$, which rearranges to

$$-kT \ln Z = \bar{E} - TS(\bar{E}) = F(\bar{E}). \quad (5.139)$$

Equation (5.138) then becomes

$$I_n = -kT \frac{\partial \ln Z}{\partial X_n}, \quad (5.140)$$

which is just (5.134) again. We see here a convergence of approaches that differentiate entropy, the partition function, and the Helmholtz energy to give the values of a system's intensive variables.

5.10 Entropy of a System Contacting a Thermal Bath

In the next few pages, we'll apply the Boltzmann distribution to construct a view of entropy that extends the standard counting of states to systems that are in contact with a heat bath.

Recall the central postulate of statistical mechanics: an isolated system is equally likely to occupy any one of its Ω accessible states. Early on, we defined the entropy of that isolated system as $S \equiv k \ln \Omega$, where k is Boltzmann's constant. But a system in contact with a bath is no longer isolated. In this case, although all states of the system-plus-bath *are* postulated to be equally likely, the states of the system itself might not be equally likely. Can we still count the system states in a meaningful way while ignoring the bath—and, if so, how does their number relate to the system's entropy in (5.124)?

Each system state can be accessed with some probability given by the Boltzmann distribution; but if the probability of a given system state is negligible, then perhaps it should not be counted at all. Defining Ω for the system has thus become problematic. If we do want to count the states of a system, which should we include and which should we leave out as being negligible? If we can find a way to answer this question in a useful manner, then it should

be possible to extend the above definition of entropy as counting states to incorporate systems that interact with a bath.

We can get a taste of what is to come by studying the canonical ensemble via (5.124), meaning a system that interacts only thermally with the bath. Begin with the Boltzmann-distributed probability for the system to occupy a state with energy E_n :

$$p_n = \frac{e^{-\beta E_n}}{Z}. \quad (5.141)$$

Rearrange this to give

$$\ln(p_n Z) = -\beta E_n. \quad (5.142)$$

Now bring in (5.124), writing it as

$$\begin{aligned} S/k &= \beta \bar{E} + \ln Z = \ln Z + \sum_n p_n \beta E_n \\ &\stackrel{(5.142)}{=} \ln Z + \sum_n p_n \times -\ln(p_n Z) \\ &= \ln Z - \sum_n (p_n \ln p_n + p_n \ln Z) \\ &= \ln Z - \sum_n p_n \ln p_n - \ln Z. \end{aligned} \quad (5.143)$$

We arrive at

$$\boxed{S/k = - \sum_n p_n \ln p_n.} \quad (5.144)$$

This expression for the entropy S of a system contacting a bath thermally is known as the system's *Gibbs entropy*. The name is not meant to imply that it is a different type of entropy from what we have used up until now; it is simply an expression that is tailored to a system contacting a bath thermally. The remainder of this chapter revolves around this formula.

Let's show that the entropy calculated from (5.144) is additive, as it should be. Consider two systems with, respectively, probabilities p_n and P_n of occupying state n . Place them in contact and calculate the entropy of the combined system before any interaction occurs. This is, with entropy again divided by Boltzmann's k for simplicity,

$$\begin{aligned} S/k &= - \sum_{mn} p_m P_n \ln(p_m P_n) = - \sum_{mn} p_m P_n (\ln p_m + \ln P_n) \\ &= - \sum_{mn} p_m P_n \ln p_m - \sum_{mn} p_m P_n \ln P_n \\ &= - \sum_m p_m \ln p_m - \sum_n P_n \ln P_n \\ &= S/k \text{ (system 1)} + S/k \text{ (system 2)}. \end{aligned} \quad (\text{QED}) \quad (5.145)$$

For the rest of this section, we derive some insight into the precise form of (5.144). Begin with the idea that if we *define* a modified number Ω of states accessible to the non-isolated system such that its entropy can still be written as $S = k \ln \Omega$, then clearly,

$$\ln \Omega = S/k = - \sum_n p_n \ln p_n. \quad (5.146)$$

It follows that

$$\Omega = \prod_n p_n^{-p_n}. \quad (5.147)$$

What is this strange-looking product? In some sense, it's the result of a counting procedure that does not give equal weighting to all the objects being counted. We can gain insight into this Ω by employing a numerical argument that makes no reference to the Boltzmann distribution.

Begin with the accessible states $1, 2, \dots, M$ of the system, where M might be infinite. To restate the problem: just as for the isolated system, if the probabilities $p_1, \dots, p_M$ of the system being found in one of these states $1, 2, \dots, M$, respectively, are all equal, then we count the states in the usual way: their number is $\Omega = M$, giving an entropy $S = k \ln M$. We wish to define a way of counting the states when the probabilities $p_1, \dots, p_M$ are *not* necessarily equal. More generally, we seek a method of counting the number of entities in a set in which those entities all occur along with an attached weighting. For example, we know that a rectangle has four sides. Imagine snipping a tiny piece off the lower-right corner to make a very small fifth side:



To what extent should we treat the new polygon as having five sides when one of those sides barely exists? How small must that side be before we can usefully ignore it? Or could the polygon be said to have a fractional number of sides?

First, let's address a related question that might appear trivial, yet we'll answer it in what seems to be a complicated way. Consider $M = 3$ states for the sake of argument. Write them as $1, 2, 3$. How many states are there here? Call this number Ω (which, of course, equals 3), because it's entirely analogous to the number of states we would have been counting up until now; we have simply replaced the notion of a state with the notion of a digit. Now suppose we build a sequence of N digits, in which each digit is equally likely to be either 1, 2, or 3, and we'll set N to be a multiple of 3 (this will be

needed later).¹⁰ How many different sequences are possible? We will write them down, in a logical order:

$$\begin{array}{ccccccc}
 1 & 1 & 1 & \dots & 1 & 1 & 1 \\
 1 & 1 & 1 & \dots & 1 & 1 & 2 \\
 & & & & & & \vdots \\
 1 & 2 & 1 & \dots & 1 & 1 & 1 \\
 & & & & & & \vdots \\
 1 & 3 & 2 & \dots & 2 & 1 & 1 \\
 & & & & & & \vdots \\
 3 & 1 & 2 & \dots & 2 & 1 & 3 \\
 & & & & & & \vdots \\
 3 & 3 & 3 & \dots & 3 & 3 & 3
 \end{array} \tag{5.148}$$

Each position in each sequence can be taken by any of the three numbers, and hence there must be 3^N sequences in total. Now suppose that we *define* Ω such that Ω^N is the number of sequences, to match $\Omega = 3$ in this case. Thus, we might count the number of digits Ω by doing something apparently quite contrived: we count the number of sequences, set this number equal to Ω^N , and solve for Ω . If the sequences were to be constructed by some random process, then they would all be equally likely. This reminds us of counting the equally likely states of an isolated system.

Instead of counting the sequences correctly (to arrive at 3^N), we might attempt to count the sequences listed above by following an economical but slightly wrong procedure. We count only the most common sequences: the ones composed of one-third 1s, one-third 2s, and one-third 3s. We expect many sequences to be of this type, and so perhaps we'll still get an acceptable answer by limiting the counting to these. After all, the sequence with, say, all 1s only occurs once; most sequences do have a fairly even distribution of 1s, 2s, and 3s.

We count these most common sequences by labelling them in a particular economical way that is widely used in tasks involving counting. Each sequence maps to a set of numbers placed into three bins. For example, the sequence “3 1 2 1 2 3” can be described as “the digit 1 occurs in positions 2 and 4, the digit 2 occurs in positions 3 and 5, and the digit 3 occurs in positions 1 and 6”. So, we create three bins: the first bin holds the indices of all occurrences of 1 in increasing order, the second holds the indices of all occurrences of 2 in increasing order, and in the third bin are the indices of all occurrences of 3

¹⁰ How does N relate to M ? Each sequence of N digits in (5.148) represents an ensemble of N copies of the system whose states are 1, 2, 3. This might help you to think of this problem in terms of ensembles, but we don't really need ensemble language in this discussion.

in increasing order:

$$\text{“3 1 2 1 2 3”} \longleftrightarrow \underbrace{2, 4}_{\text{bin 1}} \quad \underbrace{3, 5}_{\text{bin 2}} \quad \underbrace{1, 6}_{\text{bin 3}}. \quad (5.149)$$

Note that the indices should be written in increasing order. If we didn’t enforce that, we could write another set of bin contents as

$$\underbrace{4, 2}_{\text{bin 1}} \quad \underbrace{3, 5}_{\text{bin 2}} \quad \underbrace{1, 6}_{\text{bin 3}}. \quad (5.150)$$

But these two sets of bins describe the *same* sequence “3 1 2 1 2 3”, and we only require one description of each sequence. Hence, we demand that the indices in each bin are written out in increasing order. (In other words, we are dealing with *combinations* here, not *permutations*.) It might appear that all we have done is convert the sequence “3 1 2 1 2 3” to another sequence of the same length, “2 4 3 5 1 6”; but it will turn out that we gain by doing this.

Of course, we can describe *all* the sequences (rows) of (5.148) with this notation, not just the most common ones. Here are two examples, two sequences from (5.148) written in this indexed way:

$$\begin{array}{ccc} 1, 4, 5, \dots & 3, 7, 8, \dots & 2, 6, 9, \dots \\ \hline \underbrace{1, 4, 9, \dots}_{\text{bin 1}} & \underbrace{2, 3, 46, \dots}_{\text{bin 2}} & \underbrace{5, 6, 7, \dots}_{\text{bin 3}} \end{array} \quad (5.151)$$

We plan to count only the most common sequences in (5.148). These have $N/3$ indices in each of the three bins. Consider an approach of “deliberate over-counting”: we write down all $N!$ permutations of the indices, and then “realise” that we have written the sequence represented by, say, the first row of (5.151) too many times. Each bin’s list of indices for that row’s sequence appears $(N/3)!$ times instead of just once. That means we have “overdone the listing” by a factor of $(N/3)!$ for each bin. It follows that the number $N!$ must be divided by those three factors to count the number of ways (combinations!) of placing $N/3$ indices in each of the three bins. We conclude that the number of sequences in (5.148)—which should approximate the total number of sequences Ω^N —must be

$$\frac{N!}{(N/3)!(N/3)!(N/3)!}. \quad (5.152)$$

Recall that we have used an approximation here because we counted only the most common sequences. Now, we are going to make a second approximation: we’ll approximate $N!$ by the simplified Stirling’s rule $N^N e^{-N}$ from (1.27). This is by no means as accurate as the fuller version (1.25) of the rule, but it *is* slightly simpler to use than that fuller version. In that case, we have

$$\frac{N!}{(N/3)!^3} \simeq \frac{N^N e^{-N}}{[(N/3)^{N/3} e^{-N/3}]^3} = \frac{N^N e^{-N}}{(N/3)^N e^{-N}} = 3^N. \quad (5.153)$$

We expect this count to be approximately Ω^N . If we now *define* Ω by setting Ω^N equal to the final number in (5.153), then it follows that $\Omega = 3$. Surprisingly, the correct value for the *total* number of sequences has emerged—despite our employing a bad counting procedure! After all, our procedure used two approximations: (a) we counted only the most common sequences, and (b) we used Stirling’s rule, and a poor form of Stirling’s rule at that. These two “wrongs” cancelled each other out to produce the correct answer of $\Omega = 3$ for the number of digits that we set out to count in the original set “1, 2, 3”.

Oblivious to the fact that the correct result “ $\Omega = 3$ ” emerged only by happy chance, suppose that we now use this deficient procedure to tackle the *real* counting question that we set out wanting to answer. If the numbers 1, 2, 3 are produced in some random way in which they are not necessarily equally likely, then would we still say there are 3 of them? If the chance of 1 appearing was only 10^{-100} , surely we would be only concerned with analyses that involved 2 and 3. In fact, we do just that every time we flip a coin: although the chance that the coin will land on its edge is non-zero, we always discount this and consider just two possibilities, heads and tails.

With this view in mind, work with probabilities p_1, p_2, p_3 of 1, 2, and 3 being randomly produced, respectively, where these the probabilities must, of course, sum to 1. There are still 3^N different sequences possible, but now some sequences are more probable than others. Again, we count only those with a reasonable chance of occurring: we’ll *define* Ω such that Ω^N is the number of sequences in which 1, 2, and 3 appear in the proportions of p_1, p_2, p_3 . Repeating the discussion of a few paragraphs up (which had $p_1 = p_2 = p_3 = 1/3$), but now using unspecified p_1, p_2, p_3 , this number of “common” sequences is

$$\Omega^N \equiv \frac{N!}{(Np_1)!(Np_2)!(Np_3)!}. \quad (5.154)$$

Again, apply the inaccurate version of Stirling, $N! \simeq N^N e^{-N}$, to write this as

$$\begin{aligned} \Omega^N &\simeq \frac{N^N e^{-N}}{(Np_1)^{Np_1} e^{-Np_1} (Np_2)^{Np_2} e^{-Np_2} (Np_3)^{Np_3} e^{-Np_3}} \\ &= \frac{N^N e^{-N}}{N^{Np_1} N^{Np_2} N^{Np_3} p_1^{Np_1} p_2^{Np_2} p_3^{Np_3} e^{-N}} \\ &= \frac{N^N}{N^N p_1^{Np_1} p_2^{Np_2} p_3^{Np_3}} = \frac{1}{p_1^{Np_1} p_2^{Np_2} p_3^{Np_3}}. \end{aligned} \quad (5.155)$$

It follows that

$$\Omega \simeq \frac{1}{p_1^{p_1} p_2^{p_2} p_3^{p_3}}. \quad (5.156)$$

As a check, for the case of $p_1 = p_2 = p_3 = 1/3$, this becomes

$$\Omega \simeq \frac{1}{[(1/3)^{1/3}]^3} = 3, \quad (5.157)$$

which matches (5.153) and the discussion after it, as expected.

Equation (5.156) doesn't give Ω exactly; but, because we are defining a new way of counting in a statistical manner (so to speak), we are free to define Ω in any *useful* way, as long as that definition produces $\Omega = 3$ for the case of $p_1 = p_2 = p_3 = 1/3$. And indeed, here, Ω *can* be arranged to exactly equal 3, provided we change (5.156) to become the *definition* of Ω . This is, in fact, what has been done in statistical mechanics: recall that the three digits we have been using here were really pseudonyms for three *states*. Thus, we have established a procedure for counting the states of a non-isolated system, which are not necessarily equally likely. When there are M states rather than 3, their number is defined as being

$$\Omega \equiv \frac{1}{p_1^{p_1} p_2^{p_2} \dots p_M^{p_M}}. \quad (5.158)$$

This is (5.147) again! It is a remarkable result: it extends the idea of counting “whole” objects to things that are, in a sense, almost not there, such as the edge of a flipped coin. Realise that (5.158) holds exactly for a normal “2-state” coin that has an equal chance of landing heads or tails: its “effective” number of occupiable states Ω equals M exactly:

$$\Omega = \frac{1}{0.5^{0.5} 0.5^{0.5}} = 2. \quad (5.159)$$

But suppose a thick coin has a 1% chance of landing on its edge. Heads and tails (corresponding to digits 1 and 2) now each occur 49.5% of the time, and the edge (digit 3) occurs the remaining 1% of the time. What value does (5.158) give for the effective number of digits (i.e., states) here?

$$\Omega = \frac{1}{0.495^{0.495} 0.495^{0.495} 0.01^{0.01}} \simeq 2.1. \quad (5.160)$$

We can credibly say that this coin has 2.1 states able to be occupied.

Now that we have a way of counting states whose probabilities of occurrence are not necessarily all equal, we again define the system's entropy in the usual way:

$$\begin{aligned}
S/k &\equiv \ln \Omega = \ln \prod_{n=1}^M p_n^{-p_n} = \sum_n \ln (p_n^{-p_n}) \\
&= -\sum_n p_n \ln p_n.
\end{aligned} \tag{5.161}$$

We can check that (5.161) gives the usual result that the entropy of an isolated system with M states is $S = k \ln M$. The fundamental postulate of statistical mechanics says that the M states of an isolated system in equilibrium are all equally likely. In that case, $p_n = 1/M$ for all n , and (5.161) becomes

$$S/k = -\sum_{n=1}^M \frac{1}{M} \ln \frac{1}{M} = -M \times \frac{1}{M} \ln \frac{1}{M} = \ln M, \tag{5.162}$$

as expected. So, the above analysis—which was actually an *ensemble* picture—is consistent with the definition of the entropy of an isolated system.

The Gibbs expression for entropy that applies to non-isolated as well as isolated systems,

$$S/k = -\sum_n p_n \ln p_n, \tag{5.163}$$

is well known in statistical mechanics. The above way of defining it via a counting argument is not fully rigorous, as it uses two approximations that fortuitously cancel each other out to yield fully correct results for simple systems. In fact, some approaches to statistical mechanics judge a system interacting with an environment to be more fundamental than an isolated system, and (somewhat opaquely) simply *define* entropy to have the form (5.163). But there is no obvious prior reason for why the quantity $-\sum_n p_n \ln p_n$ should have anything *remotely* to do with entropy—which, at its heart, is the phenomenon behind the inexorable spreading of an ink drop in a bathtub.

5.11 The Brandeis Dice

Following early work by Gibbs, the statistical physicist E.T. Jaynes related the Gibbs expression for entropy (5.163) to the Boltzmann distribution, through the following question that he posed in his 1962 lectures at Brandeis University. We will do something similar in Section 7.9, thus making it worthwhile to study Jaynes' argument here.

A possibly biased die is thrown many times, and the results are summarised in a single statement: “The mean number showing on the top face is 5.” What can we say about the probabilities of getting each of the numbers 1 to 6 on the next throw? The mean of the numbers obtained by a great many throws of an *unbiased* die will be 3.5, and so we presume that the above die is biased. From this little information, we can certainly begin to make an

educated guess of the unknown probabilities. With p_n being the probability of obtaining number n on a throw, we estimate that p_1 is small, whereas p_5 and p_6 are large.

Jaynes defined the best estimates of the probabilities p_n to be the values of the “blandest” probability distribution consistent with the constraints of

$$\sum_{n=1}^6 p_n = 1 \quad \text{and} \quad \sum_{n=1}^6 p_n n = 5. \quad (5.164)$$

Why? Because we hardly expect anything else. There is some small possibility that the probability distribution has an eye-catching peak: the die might be unbiased, but with “5” printed on *all* of its faces, leading to $p_5 = 1$ and all other $p_n = 0$; but that does not seem to be a sensible guess at the probability distribution. (If such a die is not allowed, and we know that all numbers 1 to 6 *are* present, this spiked distribution would still arise when the die has an extreme bias that forces the number 5 always to appear.)

Consider constructing estimates of this set of probabilities $p_1, \dots, p_6$ in the following way. We enlist a team of monkeys to spend a day constructing a “three-dimensional metallic” bar graph by dropping a huge number N of coins into six vertical slots numbered 1 to 6. At the end of the day, the monkeys have dropped a total of n_i coins into the i^{th} slot. We run this experiment for many days, recording the set of values $n_1, \dots, n_6$ at the end of each day, and then removing the coins and starting from scratch the next day. Jaynes defined the “blandest probability distribution” to be the most common distribution of coins that resulted from this procedure. This means that some set $n_1, \dots, n_6$ will be the blandest one possible if it maximises $\Omega(n_1, \dots, n_6)$, the number of ways of obtaining $n_1, \dots, n_6$. The sought-after probabilities will then be $p_i = n_i/N$.

Suppose, more generally, that the monkeys drop the N coins into M slots, where $M = 6$ for a die. Then, referring to (1.12), the number of ways that some given set of numbers $n_1, \dots, n_M$ can occur is¹¹

$$\Omega = \frac{N!}{n_1! n_2! \dots n_M!}. \quad (5.165)$$

Consider that maximising Ω is equivalent to maximising $\ln \Omega$. Use the rough-and-ready version of Stirling’s rule ($x! \simeq x^x e^{-x}$), along with $p_i = n_i/N$, to write

$$\begin{aligned} \ln \Omega &= \ln N! - \sum_i \ln n_i! \\ &\simeq N \ln N - N - \sum_i (n_i \ln n_i - n_i) \end{aligned}$$

¹¹ Remember that we are dealing with combinations, not permutations, in each slot.

$$\begin{aligned}
&= N \ln N - \sum_i N p_i (\ln N + \ln p_i) \\
&= N \ln N - N \ln N - N \sum_i p_i \ln p_i \\
&= -N \sum_i p_i \ln p_i.
\end{aligned} \tag{5.166}$$

Maximising $\ln \Omega$ is then equivalent to maximising $-\sum_i p_i \ln p_i$. (If we had used the more precise version of Stirling's rule, $x! \simeq x^{x+1/2} e^{-x} \sqrt{2\pi}$, we would have reached the same conclusion with a little more effort.) Jaynes made this the entry point for a new approach to statistical mechanics, one that gave pre-eminence to the expression $-\sum p_i \ln p_i$. We have seen that this expression equals S/k for a system contacting a thermal bath, but one might loosely ignore the k and simply refer to $-\sum p_i \ln p_i$ as entropy in the above experiment involving monkeys.

Suppose we generalise the rolled-die example further by making the number on face n not necessarily n , but some E_n . If the average number thrown is E , what are Jaynes' estimates of the p_n ? Here, we are required to maximise $-\sum_n p_n \ln p_n$ subject to

$$\sum_{n=1}^M p_n = 1 \quad \text{and} \quad \sum_{n=1}^M p_n E_n = E. \tag{5.167}$$

Extremising an expression subject to constraints is commonly accomplished by the method of *Lagrange multipliers*. These multipliers are unknowns, with one multiplier allocated to each constraint. The Lagrange-multiplier approach demands that the following holds for each variable p_n :

$$\frac{\partial}{\partial p_n} \left[\begin{array}{c} \text{expression} \\ \text{to extremise} \end{array} \right] = \sum_M (\text{multiplier } M) \times \frac{\partial}{\partial p_n} \text{constraint } M. \tag{5.168}$$

This unlikely looking equation is the heart of the method of Lagrange multipliers—and it is not supposed to be obvious! (You can find the method described in calculus books.) For the generalised die, write (5.167) as

$$\text{constraint 1} = \sum_{n=1}^M p_n - 1, \quad \text{constraint 2} = \sum_{n=1}^M p_n E_n - E, \tag{5.169}$$

where both of these expressions are understood to be required to equal zero. For the two constraints in (5.167), call the multipliers α and β (being “multiplier 1” and “multiplier 2”). Then, (5.168) becomes

$$\frac{\partial}{\partial p_n} \left(-\sum_i p_i \ln p_i \right) = \alpha \frac{\partial}{\partial p_n} \sum_i p_i + \beta \frac{\partial}{\partial p_n} \sum_i p_i E_i. \tag{5.170}$$

Evaluating the partial derivatives for each n produces

$$-\ln p_n - 1 = \alpha + \beta E_n, \quad \text{for all } n. \quad (5.171)$$

Solving for p_n yields

$$p_n = \underbrace{e^{-1-\alpha}}_{\text{normalisation}} e^{-\beta E_n} = \frac{e^{-\beta E_n}}{\sum_n e^{-\beta E_n}} \equiv \frac{e^{-\beta E_n}}{Z}, \quad (5.172)$$

where we have written the normalisation as $1/Z$ to match its use in the Boltzmann distribution.

For the case of the die that we began with in (5.164), $E_n = n$, $E = 5$, and $M = 6$. Equation (5.172) gives us

$$p_1 = \frac{e^{-\beta}}{Z}, \quad p_2 = \frac{e^{-2\beta}}{Z}, \quad \dots, \quad p_6 = \frac{e^{-6\beta}}{Z}. \quad (5.173)$$

To find β , apply the constraints of (5.164). Set $x \equiv e^{-\beta}$ for shorthand, and write

$$\frac{x}{Z} + \frac{x^2}{Z} + \dots + \frac{x^6}{Z} = 1, \quad \text{and} \quad \frac{x}{Z} + \frac{2x^2}{Z} + \dots + \frac{6x^6}{Z} = 5. \quad (5.174)$$

The first constraint says that $Z = x + x^2 + \dots + x^6$. Multiplying the second constraint by Z then yields

$$x + 2x^2 + 3x^3 + 4x^4 + 5x^5 + 6x^6 = 5(x + x^2 + x^3 + x^4 + x^5 + x^6). \quad (5.175)$$

Collecting terms gives us

$$-4x - 3x^2 - 2x^3 - x^4 + x^6 = 0. \quad (5.176)$$

Clearly, x (defined as being $e^{-\beta}$) cannot equal zero; so, divide through by it to produce

$$-4 - 3x - 2x^2 - x^3 + x^5 = 0. \quad (5.177)$$

This has one real root, found numerically: $x \simeq 1.87681$. We can now write

$$Z = x + x^2 + \dots + x^6 = \frac{x - x^7}{1 - x} \simeq 91.4068. \quad (5.178)$$

Substituting these values of x and Z into (5.173) returns the sought-after best estimates of the probabilities:

$$p_1 \simeq 0.02, \quad p_2 \simeq 0.04, \quad p_3 \simeq 0.07, \quad p_4 \simeq 0.14, \quad p_5 \simeq 0.25, \quad p_6 \simeq 0.48. \quad (5.179)$$

As expected, the probabilities suggest that mostly 5s and 6s will be thrown on the die, with a few 4s. This is consistent with the initial single observation that the mean number showing is 5.

What about the case of $E = 3.5$ and all $E_n = n$? This case represents the mean for a standard unbiased die. Equation (5.174) is now replaced by

$$\frac{x}{Z} + \frac{x^2}{Z} + \cdots + \frac{x^6}{Z} = 1, \quad \text{and} \quad \frac{x}{Z} + \frac{2x^2}{Z} + \cdots + \frac{6x^6}{Z} = 3.5. \quad (5.180)$$

Follow the same procedure as above: $Z = x + x^2 + \cdots + x^6$, and so the second constraint in (5.180) becomes

$$x + 2x^2 + 3x^3 + 4x^4 + 5x^5 + 6x^6 = 3.5(x + x^2 + x^3 + x^4 + x^5 + x^6). \quad (5.181)$$

This simplifies to

$$-2.5x - 1.5x^2 - 0.5x^3 + 0.5x^4 + 1.5x^5 + 2.5x^6 = 0. \quad (5.182)$$

Again, we can divide by $x = e^{-\beta}$, writing

$$-2.5 - 1.5x - 0.5x^2 + 0.5x^3 + 1.5x^4 + 2.5x^5 = 0. \quad (5.183)$$

This has one real root: x equals exactly 1. It follows that $Z = 6$, and we arrive at

$$p_1 = p_2 = \cdots = p_6 = 1/6, \quad (5.184)$$

as expected. The blandest die that shows an average of 3.5 is an unbiased one.

In line with Jaynes' argument, we might envisage the time evolution of a gas of many particles to be equivalent to a die being thrown at each instant of time for each particle. This die has an infinite number of faces, and on each face is written some value of energy that the particle can have.¹² The number density of the particles with energy E is proportional to the probability of an energy value E showing on the die; and Jaynes' argument shows that this probability is proportional to $e^{-\beta E}$ for some β . We see that Jaynes' "Brandeis dice" generate the energy exponential in the Boltzmann distribution. (We cannot go further to relate β to temperature, because temperature is a thermodynamic quantity, and so is not readily introduced with a non-thermodynamic argument such as throwing a die. In particular, an unbiased die corresponds to a system at infinite temperature!)

¹² Whether we postulate that the infinite number of faces is *countable* or *uncountable* in a mathematical sense impacts the argument mathematically, but that is a level of detail that we don't need to pursue here.

5.12 Entropy and Data Transmission

The Gibbs expression for entropy, $S = -k \sum p_i \ln p_i$, was known to physicists long before it was rediscovered in a new context by Claude Shannon in the 1940s. Shannon was one of the pioneers of the field known today as *information theory*. We'll end this chapter with a brief look at this topic.

Consider a language whose alphabet has just two letters, “i” and “w”. (I have used these letters because I will soon be likening “i” to an ink molecule and “w” to a water molecule.) The letter “i” occurs, on average, about 10% of the time, and “i” usually appears before “w”. Also, all words in this language have approximately 20 letters. An example set of words might be¹³

$$\begin{aligned}
 & \overset{\cdot}{i} \text{wwwwwwwwwwwwwwwwww} \\
 & \overset{\cdot}{i} \text{w} \overset{\cdot}{i} \text{wwwwwwwwwwwwwwwwww} \\
 & \overset{\cdot}{i} \text{wwwwwwwwwwwwwwwwww} \\
 & \overset{\cdot}{i} \text{w} \overset{\cdot}{i} \text{wwwwwwwwwwwwwwww} \\
 & \overset{\cdot}{i} \text{wwwwwwwwwwwwwwww.}
 \end{aligned} \tag{5.185}$$

A language can be written with different alphabets, but I will consider this “iw” language to be synonymous with its alphabet.¹⁴ You can see, from (5.185), that the “iw” language is very limited. If “i” appears, the chance of it being followed by another “i” is high until about two have appeared, and then the chance of another “i” appearing becomes low. If “w” appears, the chance of it being followed by another “w” is very high. The words in (5.185) are reminiscent of the drop of ink placed in a bathtub in Chapter 1, if we set “i” to denote an ink molecule and “w” a water molecule. Just as a drop of ink placed in a bathtub initially makes an ink–water configuration of very low entropy, so too the above “iw” language could be said to possess the same very low entropy—which goes hand in hand with the language being fairly useless. In contrast, if we modify the language by removing the correlation between appearances of “i” and appearances of “w”, then many more words can be made, such as:

$$\begin{aligned}
 & \overset{\cdot}{i} \text{wwww} \overset{\cdot}{i} \text{wwwwwwwwwwwwwwww} \\
 & \text{wwww} \overset{\cdot}{i} \text{wwww} \overset{\cdot}{i} \text{wwwwwwwwww} \\
 & \text{ww} \overset{\cdot}{i} \text{wwwwwwww} \overset{\cdot}{i} \text{wwwwwwww}
 \end{aligned}$$

¹³ I have made the “i” larger here so that it stands out from the sea of “w” characters.

¹⁴ On a linguistic note, I can’t resist but point out that “iw” is a modern transliteration of the ancient Egyptian verb  “to be”, and it was often used in hieroglyphic to commence a new sentence. What would a physics book be without such an observation?

$$\begin{array}{c}
 \dot{w} \dot{l} w w w w w w w w w \dot{l} \dot{w} \dot{l} w w w \\
 w w w w w w \dot{l} w w w w w \dot{l} w w w w w w \dot{l}
 \end{array} \quad (5.186)$$

In other words, the drop of ink has now dispersed: like the dispersed ink in the bathtub, this modified language possesses a higher entropy than the old version. The words of the higher-entropy modified language in (5.186) already seem to be more capable of transmitting a message than do the words of the low-entropy old form in (5.185). That is, the higher the entropy of the language, the more capable it is of transmitting a message using a minimal number of letters. High-entropy languages are efficient for transmitting data; they do not have the wastefulness of the unmodified “iw” language above, in which long monotonous strings of “w” are regularly transmitted that have almost no instances of a following “i”. The ability of a language, or alphabet, to form useful words increases in tandem with its ability to look like a random jumble of letters.

This idea that “the entropy of a language is a measure of how efficiently it transmits data” forms the core of the modern field of information theory. This use of the word “information” is conventional, and denotes an ability to transmit data; but the choice of the word is perhaps unfortunate because it has given rise to much debate, since it carries no implication that the data is *useful* or *meaningful*, which is the everyday meaning of “information”. Information in the everyday sense of “useful or interesting facts” is a subjective concept with no really quantifiable definition. To reinforce the fact that the field centres on ways in which data might be transmitted efficiently, I will use the term *data-transmission theory* instead. It should be added that efficiency in data transmission is not solely about eliminating redundancy in the language used to transmit the data. In practice, some redundancy (such as “check sums”) is added to strings of transmitted data to help detect and correct transmission errors.

We can think of the words in the languages above as being transmitted letter by letter along, say, an electronic line to a receiver. If the receiver is ever 100% certain of what the next received letter will be, then, from a pure efficiency point of view, nothing is gained by our sending that next letter. The question is: given a letter or perhaps a string that the receiver has just received, what is the chance that the next letter to be received will be, say, “i”? In the unmodified “iw” language, we saw that if an “i” is received, then the chance is high that another “i” will be received next; and similarly, if a “w” is received, the chance is high that another “w” will follow. We cannot say the chance that “i” occurs is always 10%, because it depends on what has gone before; after all, the unmodified *and* modified forms of the “iw” language both have about 10% occurrences of “i”. But, given the string of symbols that has been received so far, we can certainly estimate the chance of the next letter being “i”.

The idea of data-transmission theory is to consider that if an “i” is highly expected to be next in a transmission, then we can imagine an ensemble of “next-received” symbols to resemble pure ink. And if “w” is highly expected to be next in a transmission, then we can imagine that ensemble to resemble water. In both cases, the entropy of the ensemble of next-received symbols is close to zero; it has no randomness at all.

Things are slightly different for the modified form of the language. There, regardless of what has been received, we can only say that the chance of the next letter being “i” is about 10%. The ensemble of next-received symbols now looks like a well-mixed blend of 10% ink and 90% water. This mixture has a high entropy. We might calculate this entropy by counting microstates in the way of Chapter 1, but data-transmission theory takes a different approach. In essence, it treats the next reception of a symbol to be, in this case, like a 2-state system coupled to a thermal bath. In the modified “iw” language, the chance of the state “i” being occupied is always 10%, and, of course, the chance of the state “w” being occupied is always 90%. The entropy of the ensemble of “next-received” symbols is then written using a modified form of (5.163): Boltzmann’s constant k is dropped, and the natural logarithm is usually replaced by the base-2 logarithm:

$$\begin{aligned} S &= -p_i \log_2 p_i - p_w \log_2 p_w \\ &= -0.1 \times \log_2 0.1 - 0.9 \times \log_2 0.9 = 0.46900. \end{aligned} \quad (5.187)$$

(Base 2 is ubiquitous in the theory of computing, because processors currently use only zeroes and ones as their “machine language”, due to the very well-defined on/off way of representing those symbols in electronic circuits.) S is called the *Shannon entropy* of the language. Notice that the Shannon entropy can be written as

$$S = -\sum_n p_n \log_2 p_n = \langle -\log_2 p \rangle. \quad (5.188)$$

Informally, $-\log_2 p$ quantifies the surprise felt by the receiver when an event (the reception of a symbol) of probability p occurs.¹⁵ Events that are absolutely certain are not surprising: for them, $p = 1$ and $-\log_2 p = 0$. An event that is very rare is very surprising to see: specifically, when $p \rightarrow 0$, $-\log_2 p \rightarrow \infty$. This measure of surprise is plotted in Figure 5.7. The Shannon entropy associated with the next letter transmitted can then be viewed as the surprise the receiver feels on seeing that letter, averaged over all possibilities of that letter.

The Shannon entropy, or average surprise, in (5.188) can be shown to be maximal when all the p_i are equal; we’ll demonstrate that below for a 2-symbol alphabet. A high Shannon entropy means that each letter is being

¹⁵ The same could be said of the logarithm to any other base. The use of $\log_2$ is purely conventional.

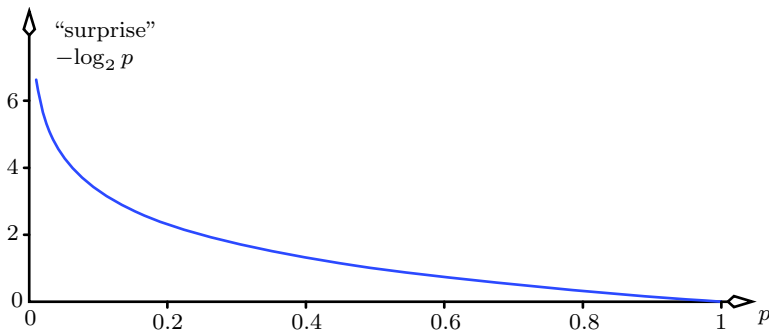


Fig. 5.7 We can think of $-\log_2 p$ as a measure of our surprise when an event of probability p occurs. When $p = 1$, we feel no surprise at all: $-\log_2 p = 0$. As $p \rightarrow 0$, the value of $-\log_2 p$ tends toward infinity

well used. Such an alphabet has a high data-transmitting ability:

$$\begin{aligned} \text{data-transmitting ability of alphabet} &\equiv \text{average surprise} \\ &\equiv \text{Shannon entropy } S. \end{aligned} \quad (5.189)$$

To demonstrate, what is the data-transmitting ability of a general alphabet of two symbols, for which there is no restriction on when any particular symbol can be used? This question relates to the *modified* form of the “iw” language above, in which we can assign fixed probabilities to the occurrences of each letter. (The unmodified form cannot be treated in this simple way, because the occurrences are correlated. In that case, we must calculate a kind of running entropy, by estimating the ever-changing probabilities as each new letter is received.) Symbol 1 appears with probability p_1 and symbol 2 appears with probability $p_2 = 1 - p_1$. There is only one free variable here: choose it to be p_1 . Equation (5.188) then says the Shannon entropy is¹⁶

$$\begin{aligned} S(p_1) &= -p_1 \log_2 p_1 - p_2 \log_2 p_2 \\ &= \frac{-1}{\ln 2} [p_1 \ln p_1 + (1 - p_1) \ln(1 - p_1)]. \end{aligned} \quad (5.191)$$

When $p_1 = 0$ or 1, the undefined expression “ $0 \ln 0$ ” occurs in (5.191). But the graph of $y = x \ln x$ has a *removable discontinuity* at $x = 0$; this means that the discontinuity is an isolated point on the function $x \ln x$ that is otherwise

¹⁶ Recall that for any a, b, c ,

$$\log_b a = \frac{\log_c a}{\log_c b} = \frac{\ln a}{\ln b}. \quad (5.190)$$

So, the logarithm to any base b equals $1/\ln b$ times the natural logarithm. This allows you to convert between $\ln$ and $\log_2$ with ease. It also means that, effectively, Shannon replaced (5.163)’s k with $1/\ln 2$.

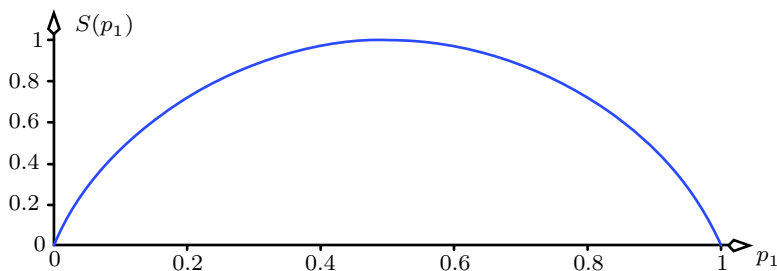


Fig. 5.8 The Shannon entropy $S(p_1)$ for a two-symbol language, using (5.191). The entropy is maximal when $p_1 = p_2 = 1/2$, meaning no one symbol is given any more prominence than the other

well behaved and continuous. In that case, we can easily replace $0 \ln 0$ with the appropriate limit, using L'Hôpital's "0/0" rule:

$$\lim_{x \rightarrow 0} x \ln x = \lim_{x \rightarrow 0} \frac{\ln x}{1/x} = \lim_{x \rightarrow 0} \frac{1/x}{-1/x^2} = \lim_{x \rightarrow 0} -x = 0. \quad (5.192)$$

Thus, we define $S(0) = S(1) \equiv 0$. Also,

$$S'(p_1) = \log_2(1/p_1 - 1). \quad (5.193)$$

This derivative is zero when $p_1 = 1/2$. A plot of $S(p_1)$ versus p_1 must then be everywhere concave down, rising from zero at the endpoints to a maximum of one at the midpoint $p_1 = p_2 = 1/2$, and symmetrical about that midpoint. This is plotted in Figure 5.8. We conclude that the data-transmitting ability (Shannon entropy) of this small alphabet is maximal when no one symbol is deliberately used more than the other. And, of course, this data-transmitting ability is zero when only one symbol is allowed to appear.

The method of Lagrange multipliers can be used to show that the same conclusion holds for an alphabet of any length—say, N symbols. To do this, we wish to extremise the entropy $-\sum_i p_i \log_2 p_i$ subject to $\sum_i p_i = 1$. This single constraint calls for a single Lagrange multiplier α . Use the natural log for simplicity: extremising $-\sum_i p_i \log_2 p_i$ is equivalent to extremising $-\sum_i p_i \ln p_i$. Referring to (5.168), we must solve

$$\frac{\partial}{\partial p_n} \left(-\sum p_i \ln p_i \right) = \alpha \frac{\partial}{\partial p_n} \sum p_i, \quad \text{for } n = 1, \dots, N. \quad (5.194)$$

Evaluating the partial derivatives for each n gives (5.171) again but without the β term:

$$-\ln p_n - 1 = \alpha, \quad \text{for all } n. \quad (5.195)$$

It follows that all of the N probabilities p_n are equal, in which case they must all equal $1/N$. It's clear that this single extremum is, in fact, a maximum,

because setting any one of the p_n equal to one and the rest to zero results in zero entropy—but entropy is always non-negative. So, the data-transmitting ability, or Shannon entropy, of an alphabet of N symbols is again maximal when each letter tends to be used equally often. Recalling (5.188), this entropy is then

$$S = \langle -\log_2 1/N \rangle = \log_2 N. \quad (5.196)$$

Efficient Transmission, Disorder, and Efficient Storage

The use of the word “entropy” in the context of data transmission can be perplexing at first. High entropy equates to efficient transmission of information—and yet high entropy also means high disorder, such as the high entropy/disorder of the fully mixed ink in a tub of water that we analysed in Chapter 1. So, does high transmission efficiency really go hand in hand with high disorder? Not really. Seen as a whole, a string of symbols that is being used to transmit information efficiently is evenly mixed. If this string had really been generated randomly, it would indeed be highly disordered. But it has *not* been generated randomly if it really is carrying a message. The even occurrences of letters will give the appearance of randomness, but they are quite the opposite of random. Hence, high transmission efficiency is not really about true disorder.

The bottom line is that although the word “entropy” stands for the expression $-\sum p_i \log p_i$ in both data-transmission theory and statistical mechanics, those probabilities p_i denote different things in these two different fields.

But additionally, information theory is not just about *transmitting* information; it’s also about *storing* information. Computer scientists continue to investigate the smallest number of *bits* (binary digits) that are sufficient to store a given data set. A data set that is very ordered (e.g., 100 ones followed by 100 zeroes) can be stored with a very low number of bits, and so is said to have very low entropy, and such a data set tends to hold very little information. When a data set is indistinguishable from a truly random one, it probably holds a lot of information—but it cannot be compressed to any great degree, and is said to have a very high entropy. Here, low or high information content equates to low or high entropy of storage, respectively.

In summary, we can transmit a message efficiently using a *high*-entropy alphabet; yet, for efficient storage, we hope the message has *low* entropy—but whether it does or not, we certainly *try* for a small amount of storage, which tends to be referred to as one of low entropy.

The Shannon Entropy of English

Let’s use (5.188) to estimate the data-transmitting ability of written English. Take a representative book: we’ll use *A Christmas Carol* by Charles Dickens. Count the letters and punctuation symbols in this book (we’ll refer to them all as symbols), and use these as estimates of the probabilities of the appearance of each of these symbols in everyday English use. We won’t distinguish between upper and lower case, and will also keep tallies of spaces, dots, commas, semicolons, colons, left quotes, and right quotes, making 33 symbols in total. The percentage occurrence of each symbol appears in Table 5.3. The space is the most common, comprising 18.7% of the total. Next most common is “e” at 9.5%, and so on. These proportions form the set $p_1, \dots, p_{33}$. Equation (5.188) estimates the Shannon entropy of written English with its alphabet as

$$S = -\sum p_i \log_2 p_i \simeq 4.23 .$$

(5.197)

Of course, we need not have sampled the entire book to arrive at this estimate; a page or two would’ve sufficed!

Sample from another book, say, Tolstoy’s “War and Peace”. The probabilities are roughly unchanged, and our estimate of the Shannon entropy of written English is now 4.17. Most English books of a similar era will yield similar estimates of the Shannon entropy of *written English*, precisely because they use the same set of rules for the appearance of each symbol. But be aware that this analysis treats the letters as occurring *randomly*: it makes no attempt to analyse their correlations. Clearly, a simple way to reduce the number of letters needed to transmit English is to replace all “qu” with “q” (ignoring the occasional mis-encoding of certain names that this produces), and yet this strong correlation between “q” and “u” does not appear in the above analysis of probabilities. To take correlations into account, we must recalculate the probabilities as each letter is transmitted. After all, if “q” is transmitted, then the chance that the next letter will be “u” is very high, which makes the entropy of transmitting this letter close to zero. That is, the data-transmitting ability of the language is temporarily reduced almost

Table 5.3 Percentage occurrences of symbols in Charles Dickens’ book *A Christmas Carol*

A, a = 6.0	J, j = 0.072	S, s = 5.1	space = 18.7
B, b = 1.2	K, k = 0.66	T, t = 7.0	dot = 0.91
C, c = 1.9	L, l = 2.9	U, u = 2.1	comma = 1.8
D, d = 3.6	M, m = 1.8	V, v = 0.66	semicolon = 0.24
E, e = 9.5	N, n = 5.1	W, w = 2.0	colon = 0.045
F, f = 1.6	O, o = 6.2	X, x = 0.084	left quote = 0.44
G, g = 1.9	P, p = 1.4	Y, y = 1.5	right quote = 0.44
H, h = 5.4	Q, q = 0.062	Z, z = 0.054	
I, i = 5.3	R, r = 4.5		

to zero. Whenever this is the case, we might choose to ignore transmitting the relevant letter entirely.

It's crucial to realise that “information theory” deals with a flow of symbols that encode data: it seeks only to describe how economically we can transmit information. It does not try to *define* information. Scrambling the letters of any book does not change the value of $-\sum_i p_i \log_2 p_i$, but it certainly does tend to destroy the information content of the book.

Suppose that the English alphabet were to be replaced by a new alphabet in which each symbol was equally likely to appear. Of course, we can always use an alphabet with two symbols: this is just what computers do. But we ask the question: how many symbols are needed to match the data-transmitting ability (Shannon entropy) of written English? In other words, how many symbols will give the same average surprise that we feel on seeing each of a stream of symbols in English? Call this number of symbols N . Then, since all of the new p_n are equal, they all equal $1/N$. So, we refer to (5.196) to say that the new alphabet's Shannon entropy is $\log_2 N$. This is required to equal the English-alphabet value of 4.2, and so we infer that $N = 2^{4.2} \simeq 18$. That is, the new alphabet would need just 18 symbols. Naturally, the language would need to change to make use of the new set of probabilities, making it no longer English as we know it.

Although 18 symbols suffice for this new version of English, this does not imply that English *should* be pared down to 18 symbols. Redundancy in data flow is useful for correcting errors in transmission. Plus, humans are not computers, and building redundancy into a language gives the listener or reader time to process—and savour!—the message.

In this chapter, we have travelled far down a path opened up by the Boltzmann distribution. The distribution lies at the core of statistical mechanics, because it predicts the behaviour of a *real* system: one that is not isolated from the rest of the world. As we have seen, the distribution predicts the behaviour of quantised systems to explain how, for example, a system's heat capacity undergoes relatively abrupt changes with temperature, as shown in Figure 5.6; this is a phenomenon of quantisation that classical physics was unable to explain. In the next chapter, we'll apply the Boltzmann distribution to classical gases, to investigate the motions of the gas particles. That analysis will extend the simplified view of gases that we began this book with, where we assumed that all the gas particles have the same energy. In practice, the particles do *not* all have the same energy, and the Boltzmann distribution is the key to predicting the more precise details of their motions.



Chapter 6

The Motion of Gas Particles, and Transport Processes

In which we study the velocity and speed distributions of particles in a gas. We use these to examine the temperature gradient in our atmosphere, and the composition of planetary atmospheres. We find out how to relate viscosity, thermal conductivity, and heat capacity using an atomic view of matter. We finish by describing the energy–momentum tensor, which has a key role in Einstein’s theory of gravity.

Why does the air around us cool when we climb a mountain? Climbing in the mountains does, of course, take us a minuscule distance closer to the Sun; but air derives almost no direct warmth from the Sun: the interaction between solar photons and air molecules is very weak. Rather, the Sun’s radiation interacts strongly with the ground, and the warmed ground then heats the air.¹ With mountain peaks all around, the Sun can rise later or set earlier in high country. This results in longer shadows, and thus fewer hours of sunlight to warm the ground and then the air. Also, winds powered by Earth’s rotation and possibly coming from cold shadowed regions around the mountains lower our skin temperature, especially if we are sweating while climbing. This *perceived* temperature drop is a product of a climber’s physiology and the temperature of the wind as measured by a thermometer. But even with plenty of strong sunlight, the simple fact is that, when we feel a strong surge of wind, that wind is *cold*, even if we have no sweat to create a chill factor.

Another reason we feel cold arises from the amount of air around us, rather than its temperature. We lose heat by conduction and radiation. Air is actually a very good insulator, provided it doesn’t move: hence, the use of double-glazed windows in cold climates, which use a vertical sheet of air trapped between panes of glass for insulation to keep a room warm. The same effect is created when insulating air is deliberately trapped around the fibres of garments, so that wearing several layers of such garments keeps us warm. The density of our atmosphere drops exponentially with altitude, and thus less air is present in the mountains to insulate us. This has the effect of lowering our skin temperature at high altitudes, and certainly, we will freeze if exposed to the upper atmosphere. On the other hand, in Section 4.3.1 we

¹ It’s nonetheless curious that we seem to feel far less *direct* heat from the Sun in winter than in summer. Compare a winter morning with its cold ground to a summer morning with its hot ground: the Sun might be, say, at an elevation of 30° on both days, and yet the direct summer sunlight certainly feels hotter—and it reduces quickly when a cloud passes in front of the Sun. Perhaps the summer’s hotter ground causes us to perceive the Sun’s rays to be hotter.

saw that the altitude at which the density drops to half its sea-level value is about 5.6 km; and yet we notice a drop in temperature even while ascending just a few hundred metres. Over a rise of 300 metres in altitude, the density drops by a factor of $2^{300/5600}$, which is only about a 4% drop. (See also the discussion of this in Section 3.6.2.)

Human physiology aside, a thermometer does register a lower temperature at higher altitudes. But thermometers also radiate, and so—like humans—they too are affected by the lower density of the air around them.

The above discussion concerned wind-chill factors and radiating objects, without addressing the question of whether an atmosphere's temperature actually decreases with altitude. Temperature is a measure of the random speeds of particles. If an atmosphere were created on an empty planet by introducing a layer of gas particles at its surface that all had *exactly* the same speed, and then we allowed that gas to expand (as gases do when not confined), then the particles that reached the mountain tops would have reduced speeds, because some of their kinetic energy had been converted to gravitational potential energy during their journey to those peaks. Their reduced speeds would imply a reduced temperature in the mountains. But consider that Section 5.8 has already suggested that air molecules have a *range* of speeds that is a function of their temperature via the equipartition theorem. For the sake of argument, suppose that most—but not all—molecules at Earth's sea level are moving at 500 m/s. Some move faster than this and others move slower. Gravity ensures that the slower molecules do not reach the mountain peaks, and it forces the faster molecules to slow down near those peaks. So, perhaps the majority of molecules at a high altitude are *also* moving at 500 m/s: they are the ones that were moving faster than 500 m/s at sea level. This suggests that the temperature of the reduced-density air in the mountains might just be the same as at sea level.

But recall that the argument we followed to define temperature in Section 3.5 assumed that no background potential was present to “soak up” energy along a gradient. Gravity *is* present in a real atmosphere, and so we cannot say a priori that no temperature gradient can exist in an atmosphere that is fully in equilibrium. We'll investigate this question of a temperature gradient in this chapter, after we determine the spread of speeds that air molecules really have.

Classically, we can view the air molecules of an idealised atmosphere (one without rain, sun/night temperature changes, and circulation of winds) as tiny ball bearings that have been tipped out of a huge bucket from a great height. Having dropped onto the warm ground at sea level, they are now bouncing incessantly, as they interact like billiard balls with the jiggling atoms of the warm ground. Earth radiates energy continuously, and if its surface were not warmed by the Sun and by the decay of radioactive elements underground, this radiating would cause it to cool toward absolute zero. Our atmosphere would gradually lose energy to the cold ground, and air molecules would then settle onto the ground as a solid. But Earth's surface *is* contin-

uously warmed by the Sun and by radioactive decay. This ensures that the molecules of the atmosphere keep bouncing up and down. Our atmosphere remains a gas. But it does not radiate energy into outer space: its “emissivity”, introduced in Section 9.7.2 ahead, is very low. Hence, it does not act as a conductor that passes energy from the ground outward.

Like a tennis ball thrown upward, each air molecule feels the pull of gravity, and so even in the absence of collisions with its neighbours, it cannot climb arbitrarily high. As air molecules climb, they lose speed. If all molecules had the same speed at sea level, this drop in their speed high up in the mountains would certainly produce a lower air temperature at those heights.

Suppose that’s the case: we’ll assume all air molecules have identical speeds at sea level. We can easily estimate this speed, as well as form a simple picture of the way in which an idealised atmosphere’s temperature then decreases with altitude. Imagine a single molecule in isolation, which has bounced from the ground at sea level and now climbs unimpeded until all of its kinetic energy has been converted to gravitational potential energy. At this point, it stops, and then falls back down. It will reach sea level with roughly the same speed that it had when it was last there. We assume this speed is much the same as the speed it would have gained in falling in the presence of other molecules.² Sometimes, collisions with those molecules will increase its speed, and other times, they will slow it down; so, we expect its speed at sea level to be largely unaffected by whether those molecules are present. A particle that falls from rest with constant acceleration g through a distance s will acquire a speed $v = \sqrt{2gs}$. Most air molecules are lower than about 15 km, so we set s to this value, and $g \simeq 9.8 \text{ m/s}^2$ for Earth. In that case, the falling particle acquires a speed at sea level of

$$v = \sqrt{2 \times 9.8 \times 15,000} \text{ m/s} \simeq 542 \text{ m/s}. \quad (6.1)$$

How does this value compare with the prediction of statistical mechanics? The speeds of the air molecules are not all the same, because in colliding with each other at various relative velocities, each can gain much energy or lose all of it. The numerical values of their speeds then spread out to form almost a continuum. Statistical mechanics treats the air in this classical way as a system in contact with a heat bath that is the ground. We found the rms speed of the molecules in (5.99):

$$v_{\text{rms}} \equiv \sqrt{\langle v^2 \rangle} = \sqrt{3RT/M_{\text{mol}}}. \quad (6.2)$$

The average temperature of the ground is about 15°C or $T = 288 \text{ K}$, and air has a molar mass of $M_{\text{mol}} = 29.0 \text{ g}$. The molecules’ rms speed is then

² Of course, this argument can be questioned, since it doesn’t apply to a feather, which certainly feels a continuous drag from the air as it falls.

$$v_{\text{rms}} = \sqrt{\frac{3 \times 8.314 \times 288}{0.0290}} \text{ m/s} \simeq 498 \text{ m/s}. \quad (6.3)$$

This is not dissimilar to the coarse estimate in (6.1).

At what rate does the temperature of the above idealised atmosphere decrease with altitude? Equation (6.2) rearranges to yield

$$T = \frac{M_{\text{mol}} \langle v^2 \rangle}{3R}. \quad (6.4)$$

Picture a generic air molecule that starts at sea level (altitude $z = 0$), where it's moving up with speed v_0 . We assume that its continual collisions act to make all of its kinetic energy available for climbing. It climbs to altitude z , where its speed has dropped to v_z . Equation (6.4) says that the temperatures at altitudes z and 0, respectively, are

$$T(z) = \frac{M_{\text{mol}} \langle v_z^2 \rangle}{3R}, \quad T_0 \equiv T(0) = \frac{M_{\text{mol}} \langle v_0^2 \rangle}{3R}. \quad (6.5)$$

Hence,

$$T(z) - T_0 = \frac{M_{\text{mol}} \langle v_z^2 - v_0^2 \rangle}{3R}. \quad (6.6)$$

The standard kinematic expression “ $v^2 - u^2 = 2as$ ”, for constant acceleration $a = -g$ and displacement $s = z$, becomes

$$v_z^2 - v_0^2 = -2gz. \quad (6.7)$$

It follows that

$$T(z) = T_0 - \frac{M_{\text{mol}} 2gz}{3R}. \quad (6.8)$$

The temperature drop with altitude is then

$$\frac{-dT}{dz} = \frac{2/3 M_{\text{mol}} g}{R}. \quad (6.9)$$

For our atmosphere, this is

$$\frac{-dT}{dz} \simeq \frac{2/3 \times 0.0290 \times 9.8}{8.314} \text{ K/m} \simeq 0.023 \text{ K/m}, \quad (6.10)$$

or about 23 kelvins per kilometre. The measured value for dry air (called our atmosphere's *dry lapse rate*) is about 10 K/km for the first 10 km of our atmosphere, the region known as the *troposphere*, where most of the air lies. Our ball-bearing model has correctly predicted three things: the speeds of molecules at sea level, the existence of a linear temperature gradient, and a value of that gradient that is close to the measured value.

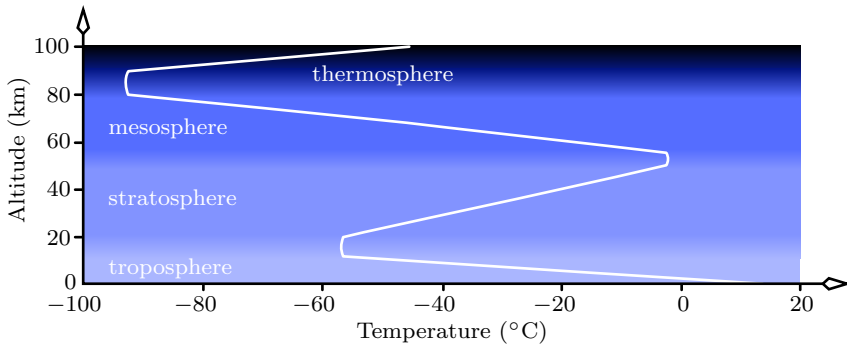


Fig. 6.1 The altitude variation of temperature in our atmosphere defines its named layers. Most of the atmosphere lies below 15 km: the troposphere

Our atmosphere also has higher layers, whose temperatures depart from the tropospheric behaviour. As shown in Figure 6.1, at higher altitudes, the temperature begins to climb, then falls again, and finally climbs once more to the outer reaches of the atmosphere. In the stratosphere, this behaviour is caused by sunlight interacting with ozone; but our discussion will remain well and truly in the troposphere.

So much for a *representative* speed of the air molecules. In this chapter, we'll go further to write down the precise form of their speed distribution, and begin to understand something of the atmospheric makeup of Earth and the other planets. We will also fine-tune the analysis of the slow-down of molecules as they climb in Earth's gravity, to see whether that implies a temperature drop. As we said earlier, the molecules' *mean* speed is a function of their temperature, but we have yet to see whether the mean speed—and thus temperature—of a set of molecules whose number dwindles with altitude is *independent* of altitude.

To begin to analyse the molecules' velocities, consider a “small box” (meaning gravity's gradient can be ignored) of ideal gas that is held at a fixed temperature by contacting a heat source. The Boltzmann distribution guarantees its particles to have a spread of energies, and hence a spread of speeds; being a gas, that also implies a spread in velocities. We ask two questions:

1. Given some velocity $\mathbf{v}$, how many particles are expected to have velocities in the vicinity of $\mathbf{v}$, meaning within a range from $\mathbf{v}$ to $\mathbf{v} + d\mathbf{v}$? (That is, these particles have similar speeds and move in similar directions.) This number is treated as infinitesimal, and so must be expressed as a *density* $\mathcal{N}_{\text{vel}}(\mathbf{v})$ times the infinitesimal volume $dv_x dv_y dv_z$ in velocity space. We will abbreviate $dv_x dv_y dv_z$ to d^3v , and so write

$$\mathcal{N}_{\text{vel}}(\mathbf{v}) d^3v \equiv \left[\begin{array}{l} \text{number of particles with velocity} \\ \text{in the range } \mathbf{v} \text{ to } \mathbf{v} + d\mathbf{v} \end{array} \right]. \quad (6.11)$$

The density function $\mathcal{N}_{\text{vel}}(\mathbf{v})$ is the *Maxwell velocity distribution*, and we wish to determine it. We can also treat a single component of velocity, say, v_x , by defining

$$\mathcal{N}_x(v_x) dv_x \equiv \left[\begin{array}{l} \text{number of particles with } x \text{ component} \\ \text{of velocity in the range } v_x \text{ to } v_x + dv_x \end{array} \right], \quad (6.12)$$

and similarly for v_y and v_z . The probability that a given gas particle has a velocity in the range $\mathbf{v}$ to $\mathbf{v} + d\mathbf{v}$ can be factored into the individual probabilities that it has velocities in the appropriate range for each dimension. In that case,

$$\frac{\mathcal{N}_{\text{vel}}(\mathbf{v}) d^3v}{N_{\text{tot}}} = \frac{\mathcal{N}_x(v_x) dv_x}{N_{\text{tot}}} \times \frac{\mathcal{N}_y(v_y) dv_y}{N_{\text{tot}}} \times \frac{\mathcal{N}_z(v_z) dv_z}{N_{\text{tot}}}. \quad (6.13)$$

If we multiply both sides of (6.13) by N_{tot} and integrate over v_y and v_z , we obtain

$$\mathcal{N}_x(v_x) dv_x = \iint_{v_y, v_z = -\infty}^{\infty} \mathcal{N}_{\text{vel}}(\mathbf{v}) d^3v. \quad (6.14)$$

2. Given some speed $v = |\mathbf{v}|$, how many particles are expected to have speeds in the range v to $v + dv$? (That is, all directions of motion are allowed. Note that v must be positive, unlike v_z above.) This number is treated as infinitesimal, and so is expressed as a density $\mathcal{N}_{\text{sp}}(v)$ times an infinitesimal interval width dv in speed space:

$$\mathcal{N}_{\text{sp}}(v) dv \equiv \left[\begin{array}{l} \text{number of particles with speed} \\ \text{in the range } v \text{ to } v + dv \end{array} \right]. \quad (6.15)$$

The density function $\mathcal{N}_{\text{sp}}(v)$ is the *Maxwell speed distribution*, and is also to be found.

For now, we omit any background force such as gravity from the discussion. Our analysis is confined to a box of gas in the lab.

6.1 The Maxwell Velocity Distribution

Suppose we set about drawing a bar graph of the numbers of particles in a room versus the x components of their velocities. In a first analysis, divide the total number N_{tot} particles into three roughly defined sets: half are moving up or down (and so have approximately zero x velocity), a quarter are moving left (negative x velocity), and a quarter are moving right (positive x velocity). Represent these numbers with a bar of height $1/2 N_{\text{tot}}$ at $v_x = 0$, followed by two bars each of height $1/4 N_{\text{tot}}$ at equal distances somewhere to the left and

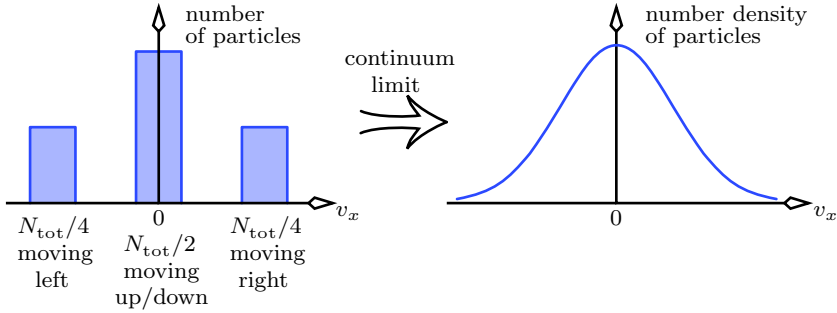


Fig. 6.2 Left: A very coarse first attempt at a bar graph showing the spread in x velocities of particles in a gas. **Right:** The limit of a continuous spread of velocity bins requires the number *density* of particles to be plotted

right of $v_x = 0$, with all bars having equal widths. This bar graph is shown at the left in Figure 6.2. We see that even with this coarsest of grainings, a symmetrical function that peaks at $v_x = 0$ is beginning to form. In the limit where velocity is truly continuous, the bin widths shrink to zero, and we can no longer plot the number of particles in each bar because those numbers shrink to zero. Instead, we plot the number *density* as a function of v_x . The result is called a *histogram*, shown at the right in Figure 6.2. The area under the curve between any two values of v_x gives the number of particles whose x velocities lie between those two values. In a similar fashion, we expect $\mathcal{N}_{\text{vel}}(\mathbf{v})$ to be symmetrical in the velocity components v_x, v_y, v_z .

We made the above plot using the coarsest counting of velocity vectors, binning them into left, right, up, and down. To count velocity vectors that point in all directions, refer to a representative set of these vectors in Figure 6.3. We place the tails of the vectors at the origin of *velocity space* and set about counting them. The number density $\mathcal{N}_{\text{vel}}(\mathbf{v})$ of the vectors at velocity $\mathbf{v}$ equals the number of vectors that have their heads in the interval $\mathbf{v}$ to $\mathbf{v} + d\mathbf{v}$ divided by the volume $d^3v = dv_x dv_y dv_z$ of that infinitesimal box in velocity space. The gas particles each have mass m , and we will treat them as distinguishable. The number in this infinitesimal box, $\mathcal{N}_{\text{vel}}(\mathbf{v}) d^3v$, divided by the total number N_{tot} , is the probability that any particular particle will be found with some velocity in $\mathbf{v}$ to $\mathbf{v} + d\mathbf{v}$:

$$\begin{aligned} \frac{\mathcal{N}_{\text{vel}}(\mathbf{v}) d^3v}{N_{\text{tot}}} &= \left[\begin{array}{c} \text{probability that particle} \\ \text{has } x \text{ velocity in } v_x \text{ to} \\ v_x + dv_x \end{array} \right] \times \left[\begin{array}{c} \text{same for} \\ y \text{ velocity} \end{array} \right] \times \left[\begin{array}{c} \text{same for} \\ z \text{ velocity} \end{array} \right] \\ &= \left[\begin{array}{c} \text{probability particle is in a state with} \\ E = 1/2 m v_x^2 + \cdots + 1/2 m v_z^2 = 1/2 m v^2 \end{array} \right] \times \left[\begin{array}{c} \text{number of states} \\ \text{in } E \text{ to } E + dE \end{array} \right]. \end{aligned} \quad (6.16)$$

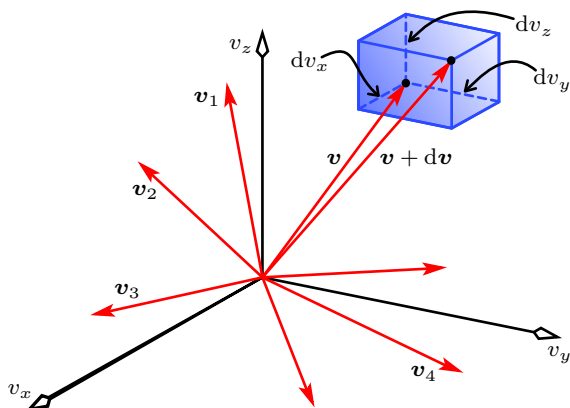


Fig. 6.3 All velocity vectors of the gas particles at any moment lie in velocity space, with cartesian axes that define an infinitesimal box extending from $\mathbf{v}$ to $\mathbf{v} + d\mathbf{v}$

The probability that the particle is in a state with energy $E = \frac{1}{2}mv^2$ is, according to Boltzmann, proportional to $\exp \frac{-mv^2}{2kT}$: the pressure/volume term in (5.5) is not relevant here.

The number of states in the energy range E to $E + dE$ is found with the approach of Sections 2.4 and 2.5. Recall that this number of states is $d\Omega_{\text{tot}} = g(E) dE$, where $g(E)$ is the density of states. We could calculate $\Omega_{\text{tot}}(E)$, the number of states in the energy range 0 to E , and then find the derivative $g(E) = \Omega'_{\text{tot}}(E)$. Alternatively, focus on $d\Omega_{\text{tot}}$ using (2.24), and refer to (6.16) to be reminded that we are analysing a *single* particle moving in three dimensions; thus, $D = 3$ and $N = 1$:

$$d\Omega_{\text{tot}} = \text{number of (micro)states in } E \text{ to } E + dE \propto \int_{\substack{\text{all space,} \\ \text{constant velocity}}} dx^3 dp^3. \quad (6.17)$$

The spatial part of (6.17) integrates to be the gas's volume, a constant. When momentum equals mass times velocity,³ the term dp^3 denotes the product of infinitesimal momentum intervals:

$$dp^3 \equiv dp_x dp_y dp_z = m dv_x m dv_y m dv_z = m^3 d^3v. \quad (6.18)$$

So, $d\Omega_{\text{tot}} \propto d^3v$. This allows (6.16) to be written as

$$\mathcal{N}_{\text{vel}}(\mathbf{v}) d^3v \propto \exp \frac{-mv^2}{2kT} d^3v, \quad \text{or } \mathcal{N}_{\text{vel}}(\mathbf{v}) = C \exp \frac{-mv^2}{2kT} \quad (6.19)$$

³ Momentum usually equals mass times velocity, but see the footnote on canonical momentum in Section 2.3.

for some normalisation C . Determine C by counting the particles [that is, integrating $\mathcal{N}_{\text{vel}}(\mathbf{v}) d^3v$], knowing that their total number is N_{tot} :

$$N_{\text{tot}} = \int_{\text{all velocities}} \mathcal{N}_{\text{vel}}(\mathbf{v}) d^3v = C \int_{\text{all velocities}} \exp \frac{-mv^2}{2kT} d^3v. \quad (6.20)$$

Each integral sign in (6.20) is really a triple integral that ranges over all velocity components from $-\infty$ to ∞ . (Such velocities don't accord with relativity, and we would not be able to make this analysis relativistic by simply changing the limits to the speed of light. Instead, we would have to consider the appropriate relativistic expression for energy, since $\frac{1}{2}mv^2$ is a non-relativistic expression.)

The integral in (6.20) is easy to evaluate, because $v^2 = v_x^2 + v_y^2 + v_z^2$:

$$N_{\text{tot}} = C \iiint_{-\infty}^{\infty} \exp \frac{-mv_x^2 - mv_y^2 - mv_z^2}{2kT} dv_x dv_y dv_z$$

$$\stackrel{(1.115)}{=} C \left(\frac{2\pi kT}{m} \right)^{3/2}. \quad (6.21)$$

Solving this for C and substituting the result into (6.19) gives us the *Maxwell velocity distribution*:

$$\boxed{\mathcal{N}_{\text{vel}}(\mathbf{v}) = N_{\text{tot}} \left(\frac{m}{2\pi kT} \right)^{3/2} \exp \frac{-mv^2}{2kT}.} \quad (6.22)$$

Note that while (6.22) is a *velocity* distribution, it requires knowledge only of *speed* $v = |\mathbf{v}|$. After all, symmetry dictates that the number density (in velocity space) of particles moving at approximately $v = 500$ m/s east should equal the number density of particles moving at approximately 500 m/s north-west-up: so, only this speed v should be required, and not the particles' direction.

The exponential part of (6.22) factors into three similar gaussian exponentials for v_x, v_y, v_z . Unlike v , each of these variables can be either positive or negative. Each separate gaussian is then an even function. Then, as expected, $\mathcal{N}_{\text{vel}}(\mathbf{v})$ is indeed symmetric in the velocity components v_x, v_y, v_z , just as we saw in Figure 6.2.

A measure of the width of the gaussian in (6.22) is the corresponding standard deviation σ . Compare (6.22) with (1.103), to write

$$\frac{1}{2\sigma^2} = \frac{m}{2kT}, \quad (6.23)$$

which implies $\sigma = \sqrt{kT/m}$. This makes good sense: it shows that the distribution is broadened by higher temperatures and less massive gas particles.

6.1.1 Alternative Derivation of the Velocity Distribution

Here is a derivation of the velocity distribution that doesn't depend on knowledge of the Boltzmann distribution. Instead, it uses the idea of the exponential fall-off of particle-number density in an atmosphere. In Section 5.1.1, we pointed out that we had derived the exponential fall-off (5.8) of the particle-number density with altitude z in an atmosphere with a *single temperature throughout* in three different ways, only one of which used the Boltzmann distribution. Let's now apply this exponential fall-off (5.8), which—for the sake of argument—did not rely on prior knowledge of the Boltzmann distribution.

Recall (6.13), which relates the Maxwell velocity distribution to the individual velocity distributions for each dimension. Molecular interactions ensure that the velocity distributions for motion in each direction at a single altitude, $\mathcal{N}_x(v_x)$, $\mathcal{N}_y(v_y)$, and $\mathcal{N}_z(v_z)$, have the same form. Focus, then, on $\mathcal{N}_z(v_z)$, for which we can bring in knowledge of the particle-number density's exponential fall-off with altitude z .

Define the altitude-dependent probability density $\mathcal{N}_z(z, v_z)$ by⁴

$$\frac{\mathcal{N}_z(z, v_z) dz dv_z}{N_{\text{tot}}} \equiv \left[\begin{array}{l} \text{probability that a particle has an} \\ \text{altitude of } z \text{ and a } z \text{ velocity of } v_z \end{array} \right]. \quad (6.24)$$

In an atmosphere with a single temperature throughout, it's reasonable to assume that the probability that a particle has a given z velocity is independent of its altitude. The probability in (6.24) can then be factored:

$$\begin{aligned} \frac{\mathcal{N}_z(z, v_z) dz dv_z}{N_{\text{tot}}} &= \left[\begin{array}{l} \text{probability that} \\ \text{a particle has an} \\ \text{altitude of } z \end{array} \right] \times \left[\begin{array}{l} \text{probability that a particle} \\ \text{has a } z \text{ velocity of } v_z \end{array} \right] \\ &\propto \exp \frac{-mgz}{kT} dz \times \frac{\mathcal{N}_z(v_z) dv_z}{N_{\text{tot}}}. \end{aligned} \quad (6.25)$$

Focus on sea level $z = 0$ by writing $\mathcal{N}_z(0, v_z) = \mathcal{N}_z(v_z)$. It follows from this and (6.25) that

$$\mathcal{N}_z(z, v_z) = \mathcal{N}_z(v_z) \exp \frac{-mgz}{kT}. \quad (6.26)$$

Now picture a small set of air molecules with z velocity v_z at altitude z . In a time dt , this z velocity carries the molecules to a new altitude $z + v_z dt$, after which gravity (causing an acceleration $-g$) has changed their z velocity to $v_z - g dt$. The first configuration has evolved to become the second; so, in equilibrium, the numbers of molecules in each configuration must be equal. Hence,

$$\mathcal{N}_z(z + v_z dt, v_z - g dt) = \mathcal{N}_z(z, v_z). \quad (6.27)$$

⁴ As introduced in (6.12), the subscript z here denotes that $\mathcal{N}_z(z, v_z)$ is a density for z velocities. The first argument z refers to the altitude z .

Taylor-expand the left-hand side of (6.27), remembering that what at first appears to be a first-order approximation is actually *exact*, because we are using infinitesimals:

$$\cancel{\mathcal{N}_z(z, v_z)} + \frac{\partial \mathcal{N}_z(z, v_z)}{\partial z} v_z dt - \frac{\partial \mathcal{N}_z(z, v_z)}{\partial v_z} g dt = \cancel{\mathcal{N}_z(z, v_z)}. \quad (6.28)$$

A slight rearranging produces

$$v_z \frac{\partial \mathcal{N}_z(z, v_z)}{\partial z} = g \frac{\partial \mathcal{N}_z(z, v_z)}{\partial v_z}. \quad (6.29)$$

Refer to (6.26) to calculate the partial derivatives. We obtain

$$\frac{\mathcal{N}'_z(v_z)}{\mathcal{N}_z(v_z)} = \frac{-mv_z}{kT}. \quad (6.30)$$

This integrates to

$$\ln \mathcal{N}_z(v_z) = \frac{-mv_z^2}{2kT} + \text{constant}, \quad (6.31)$$

and so

$$\mathcal{N}_z(v_z) \propto \exp \frac{-mv_z^2}{2kT}. \quad (6.32)$$

Normalise this for a box of N_{tot} particles sited at sea level, by invoking $\int_{-\infty}^{\infty} \mathcal{N}_z(v_z) dv_z = N_{\text{tot}}$. The result is

$$\mathcal{N}_z(v_z) = N_{\text{tot}} \sqrt{\frac{m}{2\pi kT}} \exp \frac{-mv_z^2}{2kT}. \quad (6.33)$$

Now apply (6.13), and use the idea that $\mathcal{N}_x(v_x)$, $\mathcal{N}_y(v_y)$, and $\mathcal{N}_z(v_z)$ all have the same form at the same altitude:

$$\begin{aligned} \frac{\mathcal{N}_{\text{vel}}(\mathbf{v}) d^3v}{N_{\text{tot}}} &= \frac{\mathcal{N}_x(v_x) dv_x}{N_{\text{tot}}} \times \frac{\mathcal{N}_y(v_y) dv_y}{N_{\text{tot}}} \times \frac{\mathcal{N}_z(v_z) dv_z}{N_{\text{tot}}} \\ &\stackrel{(6.33)}{=} \sqrt{\frac{m}{2\pi kT}} \exp \frac{-mv_x^2}{2kT} dv_x \times \left[\begin{array}{c} \text{similarly} \\ \text{for } y \end{array} \right] \times \left[\begin{array}{c} \text{similarly} \\ \text{for } z \end{array} \right] \\ &= \left(\frac{m}{2\pi kT} \right)^{3/2} \exp \frac{-mv^2}{2kT} d^3v. \end{aligned} \quad (6.34)$$

Hence,

$$\mathcal{N}_{\text{vel}}(\mathbf{v}) = N_{\text{tot}} \left(\frac{m}{2\pi kT} \right)^{3/2} \exp \frac{-mv^2}{2kT}, \quad (6.35)$$

which is the Maxwell velocity distribution once more. This analysis assumes that all the particles share a common temperature T . We'll relax that assumption for our atmosphere in Section 6.5.

6.2 The Maxwell Speed Distribution

The Maxwell velocity distribution gives a detailed description of the particle motions, but in practice, we tend to be more interested in the *speeds* of the particles than in their directions of motion. The applicable distribution of the numbers of particles in the speed range v to $v + dv$ is the Maxwell speed distribution. Be prepared to find that the Maxwell speed distribution is called the Maxwell velocity distribution in some books. One explanation for this might be that both distributions turn out to depend on speed alone, and not velocity. Speed, of course, is the length of the velocity vector $\mathbf{v}$, and so is usually written as a non-bold v . Perhaps the use of this standard symbol is one reason for why “speed” is often erroneously called “velocity”, and not only by physicists.⁵

In (6.15), we defined $\mathcal{N}_{\text{sp}}(v) dv$ as being the infinitesimal number of particles found in the range of speeds from v to $v + dv$. This is the total number of particles in the corresponding velocity range with all directions of motion allowed, so we return to the velocity distribution and sum over all directions in space. Part of this approach is akin to (1.116): that equation’s radial coordinate r that applies to three spatial dimensions is replaced here by speed v , the radial coordinate in velocity space:

$$\begin{aligned} \mathcal{N}_{\text{sp}}(v) dv &= \int_{\text{all directions}} \mathcal{N}_{\text{vel}}(\mathbf{v}) d^3v = \int_{\text{all directions}} N_{\text{tot}} \left(\frac{m}{2\pi kT} \right)^{3/2} \exp \frac{-mv^2}{2kT} d^3v \\ &= N_{\text{tot}} \left(\frac{m}{2\pi kT} \right)^{3/2} \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta v^2 \exp \frac{-mv^2}{2kT} dv \\ &= N_{\text{tot}} \left(\frac{m}{2\pi kT} \right)^{3/2} \times 2\pi \times 2 \times v^2 \exp \frac{-mv^2}{2kT} dv. \end{aligned} \quad (6.36)$$

This simplifies to

$$\mathcal{N}_{\text{sp}}(v) = N_{\text{tot}} \sqrt{\frac{2}{\pi}} \left(\frac{m}{kT} \right)^{3/2} v^2 \exp \frac{-mv^2}{2kT}. \quad (6.37)$$

This is the *Maxwell speed distribution*. Compare it with the velocity distribution (6.22): apart from the different normalisation, the speed distribution has an extra factor of v^2 . This pushes its peak out to some nonzero value of speed that we’ll determine shortly.

⁵ Probably the main reason is that “velocity” sounds more erudite than the everyday word “speed”. Such misappropriation of jargon is probably common in all the sciences. See ahead the comment just after (6.92), and the associated footnote.

Naturally, verifying the Maxwell speed distribution requires some careful work, and it was not until the 1955 experiments of Miller and Kusch that the distribution was finally verified; until then, experiments tended to be less sensitive to lower particle speeds.

6.2.1 Alternative Derivation of the Speed Distribution

We found the speed distribution above by integrating the velocity distribution over all directions. But we should be able to produce the speed distribution from first principles by counting states, without referring to the velocity distribution at all. Here is how to do that.

Just as we did for the velocity distribution, we again calculate a density of velocity vectors; but this time, only their length matters: their direction doesn't interest us. Thus, instead of analysing the infinitesimal box in velocity space in Figure 6.3, we analyse a spherical shell of radius v and infinitesimal thickness dv in velocity space. The number of velocity vectors whose heads are in this shell, $\mathcal{N}_{\text{sp}}(v) dv$, divided by N_{tot} equals the probability that any particular particle will be found with some speed in v to $v + dv$:

$$\begin{aligned} \frac{\mathcal{N}_{\text{sp}}(v) dv}{N_{\text{tot}}} &= \text{probability that particle has speed in } v \text{ to } v + dv \\ &= \left[\begin{array}{c} \text{probability particle is in} \\ \text{a state with } E = \frac{1}{2}mv^2 \end{array} \right] \times \left[\begin{array}{c} \text{number of states} \\ \text{in } E \text{ to } E + dE \end{array} \right]. \end{aligned} \quad (6.38)$$

Boltzmann's distribution says the probability that the particle is in a state with energy $E = \frac{1}{2}mv^2$ is proportional to $\exp \frac{-mv^2}{2kT}$. And similarly to the velocity distribution in (6.17), for this shell, we write

$$\text{number of (micro)states in } E \text{ to } E + dE = d\Omega_{\text{tot}} \propto \int_{\substack{\text{all space,} \\ \text{constant speed}}} dx^3 dp^3. \quad (6.39)$$

The spatial part here integrates to be the gas's constant volume. Again dp^3 denotes $m^3 d^3v$, as in (6.18), but now it is being integrated over all velocities that correspond to the given speed v . This integral is proportional to the volume of the shell in velocity space, $4\pi v^2 dv$, because the velocity vectors can have their heads anywhere in this shell. Hence, $d\Omega_{\text{tot}} \propto v^2 dv$, and (6.38) becomes

$$\mathcal{N}_{\text{sp}}(v) dv \propto \exp \frac{-mv^2}{2kT} v^2 dv, \quad \text{or} \quad \mathcal{N}_{\text{sp}}(v) \propto v^2 \exp \frac{-mv^2}{2kT}. \quad (6.40)$$

This can be normalised to produce (6.37) again.

One final mathematical point is important to be aware of here. The expression

$$\mathcal{N}_{\text{sp}}(v) \propto v^2 e^{-mv^2/(2kT)} \propto E e^{-E/(kT)} \quad (6.41)$$

does *not* imply that the number density expressed as a function of energy, $\mathcal{N}_{\text{en}}(E)$, is simply $E e^{-E/(kT)}$. Remember that $\mathcal{N}_{\text{en}}(E)$ is defined via

$$\mathcal{N}_{\text{en}}(E) dE \equiv \mathcal{N}_{\text{sp}}(v) dv, \quad (6.42)$$

which means

$$\mathcal{N}_{\text{en}}(E) = \frac{\mathcal{N}_{\text{sp}}(v)}{E} \times \frac{dv}{dE}. \quad (6.43)$$

What is dv/dE ? Use $E = 1/2 mv^2$; then $v \propto \sqrt{E}$, and so $dv/dE \propto E^{-1/2}$. Equation (6.43) becomes

$$\mathcal{N}_{\text{en}}(E) \propto E e^{-E/(kT)} E^{-1/2}. \quad (6.44)$$

We see that $\mathcal{N}_{\text{en}}(E) \propto \sqrt{E} e^{-E/(kT)}$, instead of the $E e^{-E/(kT)}$ that might have been inferred naively from (6.41).

6.3 Representative Speeds of Gas Particles

Knowledge of the spread of particle speeds is useful, but we are often interested in a one-parameter representation of this spread: an average speed. Perhaps little physical insight is gained by investigating how many different types of average we might define for a system; but having said that, the exercise is useful mathematically, and it sheds light on which definitions are simpler than others. We list here four standard types of average that apply to the speeds of the particles in our gas. In order of increasing size, they are

1. the *most likely speed* $\hat{v}$,
2. the *median speed* v_{med} ,
3. the (*arithmetic*) *mean speed* $\bar{v}$ or $\langle v \rangle$, and
4. the *rms speed* v_{rms} .

Each of these is derived from the Maxwell speed distribution $\mathcal{N}_{\text{sp}}(v)$ in (6.37). Here, we'll declutter $\mathcal{N}_{\text{sp}}(v)$ by switching to a dimensionless variable u :

$$u \equiv \frac{v}{\sqrt{kT/m}}. \quad (6.45)$$

$\mathcal{N}_{\text{sp}}(u)$ is defined via $\mathcal{N}_{\text{sp}}(u) du \equiv \mathcal{N}_{\text{sp}}(v) dv$, in which case

$$\mathcal{N}_{\text{sp}}(u) = N_{\text{tot}} \sqrt{2/\pi} u^2 e^{-u^2/2}. \quad (6.46)$$

This change of variables amounts to a change of units to express speed: the physics is not changed. It follows that the different means expressed in terms of u are related to the corresponding means expressed in terms of v by (6.45).

1. **Most likely speed** $\hat{v}$: the speed at which the Maxwell speed distribution peaks. Find it by solving $N'(\hat{u}) = 0$. The straightforward differentiation of (6.46) leads to $\hat{u} = \sqrt{2} \simeq 1.4$. Thus, from (6.45),

$$\hat{v} = \hat{u} \sqrt{\frac{kT}{m}} = \sqrt{\frac{2kT}{m}} = \sqrt{\frac{2RT}{M_{\text{mol}}}} \simeq 1.4 \sqrt{\frac{RT}{M_{\text{mol}}}}. \quad (6.47)$$

Here, as usual, R is the gas constant and M_{mol} is the gas particles' molar mass.

2. **Median speed** v_{med} : half the particles are travelling slower than this speed, and half faster. Find it by solving

$$\int_0^{u_{\text{med}}} \mathcal{N}_{\text{sp}}(u) \, du = \frac{N_{\text{tot}}}{2}. \quad (6.48)$$

This becomes

$$\int_0^{u_{\text{med}}} u^2 e^{-u^2/2} \, du = \frac{1}{2} \sqrt{\frac{\pi}{2}}. \quad (6.49)$$

Now apply (1.96) to the left-hand side of (6.49), to arrive at

$$-\sqrt{\frac{2}{\pi}} u_{\text{med}} e^{-u_{\text{med}}^2/2} + \text{erf} \frac{u_{\text{med}}}{\sqrt{2}} = \frac{1}{2}. \quad (6.50)$$

This solves numerically to yield one root at $u_{\text{med}} \simeq 1.53817$. The median speed is then

$$v_{\text{med}} = u_{\text{med}} \sqrt{\frac{kT}{m}} \simeq 1.53817 \sqrt{\frac{kT}{m}} \simeq \sqrt{\frac{2.366 kT}{m}} \simeq 1.5 \sqrt{\frac{RT}{M_{\text{mol}}}}. \quad (6.51)$$

3. **Arithmetic mean speed** $\bar{v}$ or $\langle v \rangle$: the usual “arithmetic” definition of the mean sums the speeds and divides this by the total number of particles. As in (1.157), this is equivalent to summing the speeds weighted by the fraction of particles with each speed:

$$\bar{u} = \int_0^\infty u \times \text{probability}(u) \, du = \int_0^\infty u \frac{\mathcal{N}_{\text{sp}}(u) \, du}{N_{\text{tot}}} \stackrel{(6.46)}{=} \sqrt{\frac{2}{\pi}} \int_0^\infty u^3 e^{-u^2/2} \, du. \quad (6.52)$$

You can do the last integral by parts, writing it as $\int_0^\infty u^2 \times u e^{-u^2/2} \, du$. The result is $\bar{u} = \sqrt{8/\pi} \simeq 1.6$, or

$$\bar{v} = \bar{u} \sqrt{\frac{kT}{m}} = \sqrt{\frac{8kT}{\pi m}} \simeq 1.6 \sqrt{\frac{RT}{M_{\text{mol}}}}. \quad (6.53)$$

It's intriguing that π appears here—as it does unexpectedly in so many equations of physics. After all, on the face of it, why should the ratio of a circle's circumference to its diameter have anything to do with summing the speeds of the particles in a gas?

4. **RMS speed** v_{rms} : the rms value of any varying quantity is the “(square) root (of the) mean (of the) square” of that quantity:

$$u_{\text{rms}}^2 = \langle u^2 \rangle = \int_0^\infty u^2 \frac{\mathcal{N}_{\text{sp}}(u) du}{N_{\text{tot}}} = \sqrt{\frac{2}{\pi}} \int_0^\infty u^4 e^{-u^2/2} du. \quad (6.54)$$

Evaluate this integral by “differentiating (twice) under the integral sign”, as per (1.99). The result is $u_{\text{rms}} = \sqrt{3} \simeq 1.7$. Equation (6.45) then gives

$$v_{\text{rms}} = u_{\text{rms}} \sqrt{\frac{kT}{m}} = \sqrt{\frac{3kT}{m}} \simeq 1.7 \sqrt{\frac{RT}{M_{\text{mol}}}}. \quad (6.55)$$

This value makes good sense, since it implies that the mean value of a particle's energy is

$$\langle E \rangle = \langle \frac{1}{2} m v^2 \rangle = \frac{m}{2} \langle v^2 \rangle = \frac{m v_{\text{rms}}^2}{2} = \frac{m}{2} \frac{3kT}{m} = \frac{3}{2} kT. \quad (6.56)$$

We expect this result from the equipartition theorem: an ideal-gas particle's average energy per quadratic energy term is $\frac{1}{2} kT$, and with only translational kinetic energy here, the particle has 3 energy terms. This simple connection with the average energy makes the rms speed the most widely used representative speed of the particles—the more so because it can be calculated without any knowledge of the Maxwell speed distribution. We did just that in Section 5.8, by running (6.56) in reverse:

$$v_{\text{rms}}^2 = \frac{2}{m} \left\langle \frac{m v^2}{2} \right\rangle = \frac{2}{m} \times \frac{3}{2} kT = \frac{3kT}{m}. \quad (6.57)$$

Figure 6.4 shows the speed distribution with the above four types of average speed indicated, to scale. Its shape is independent of the total number of particles. Given that its full width at half maximum is a rather large 1.63 approximately, the difference between the various flavours of average ($u \simeq 1.4, 1.5, 1.6, 1.7$) is not significant; any of them could be used as a measure of the distribution's centre.

Let's calculate the peak speeds of three representative molecules: H_2 , O_2 , and the heavyweight UF_6 . At a room temperature of 298 K, equation (6.47) gives the peak speed as

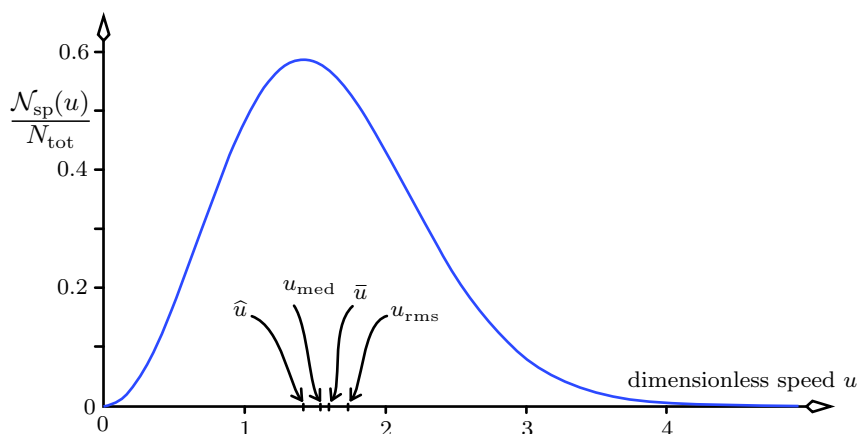


Fig. 6.4 Maxwell speed distribution with the four key speeds indicated, to scale

$$\hat{v} = \sqrt{\frac{2RT}{M_{\text{mol}}}} = \sqrt{\frac{2 \times 8.314 \times 298}{M_{\text{mol}}/(1 \text{ kg})}} \text{ m/s}. \quad (6.58)$$

The results are given in Table 6.1. Because the peak speed is inversely proportional to the *square root* of the molar mass, even the very heavy molecule UF_6 is no sluggard at room temperature.

How do the mean speed $\bar{v}$ and rms speed v_{rms} relate to each other? They are connected by the standard deviation of the speeds, σ_v . We did that calculation back in (1.47): just substitute v for x in that calculation, to arrive at

$$\sigma_v^2 = \langle v^2 \rangle - \bar{v}^2. \quad (6.59)$$

But since v_{rms}^2 is just another name for $\langle v^2 \rangle$, equation (6.59) rearranges to

$$v_{\text{rms}}^2 = \bar{v}^2 + \sigma_v^2. \quad (6.60)$$

That is, $\bar{v}$ and σ_v add “in quadrature” to produce v_{rms} : “in quadrature” refers to using Pythagoras’s theorem to show geometrically how the three quantities relate, as shown in Figure 6.5. We see here how v_{rms} , $\bar{v}$, and σ_v are related for *any* quantity v : it doesn’t have to be speed. v_{rms} will equal $\bar{v}$ if and only

Table 6.1 Peak speeds of three representative molecules at a room temperature of 298 K. The predominant isotope of uranium is ^{238}U (of molar mass 238 g). Fluorine atoms have a molar mass of 19 g

	H_2	O_2	UF_6
Molar mass:	2 g	32 g	352 g
$\hat{v}$:	1574 m/s	394 m/s	119 m/s

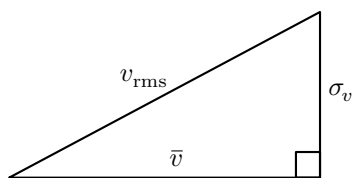


Fig. 6.5 Equation (6.60) gives a handy geometrical picture of how v_{rms} , $\bar{v}$, and σ_v are related for a set of values of *any* quantity v

if v has no spread σ_v in values. The greater the spread, the more the rms and mean values differ.

The characteristic factor of kT/m appearing in the above speeds matches what we expect: it says that the higher the temperature or the lower the particle mass, the faster the gas particles will move. This straightforward idea can be applied to the process of separating molecules of different masses from each other. It was used during World War II to separate the rare uranium isotope ^{235}U (needed for building an atomic bomb) from the much more abundant ^{238}U , which could not be used in the bomb because it isn't fissile. These isotopes have identical chemical properties, and so cannot be separated chemically. Instead, they must be separated by some physical process, such as one that exploits their differing masses. The raw uranium was converted into uranium hexafluoride gas, UF_6 . Picture the molecules of this gas bouncing around inside a chamber whose surface is punctured by a set of small holes. In a given time, the faster-moving $^{235}\text{UF}_6$ molecules explore the container more thoroughly than the slower $^{238}\text{UF}_6$ molecules, and so the $^{235}\text{UF}_6$ have more opportunities to escape through the holes. Over time, the gas inside the container loses proportionally more $^{235}\text{UF}_6$ molecules than $^{238}\text{UF}_6$ molecules. This physical process was indeed successful at separating these two uranium isotopes on a scale needed to build the first atomic bombs.

6.4 Doppler Broadening of a Spectral Line

Here is a more commonplace example of the Maxwell distribution. Low-pressure sodium lamps are very efficient at producing a bright yellow light, and so are widely used for illuminating large areas at night. This yellow light is produced by a de-excitation that occurs in vaporised sodium ions that have been excited by an electric current. Almost all of the light produced by such lamps is actually a *spectral doublet*: two closely spaced frequencies of 589.0 nm and 589.6 nm. The sodium ions in the lamp are really moving with a Maxwell spread of velocities, and each velocity has the effect of Doppler-shifting the light produced. We ask: what is the Doppler-shifted frequency distribution of the doublet produced by the sodium lamp?

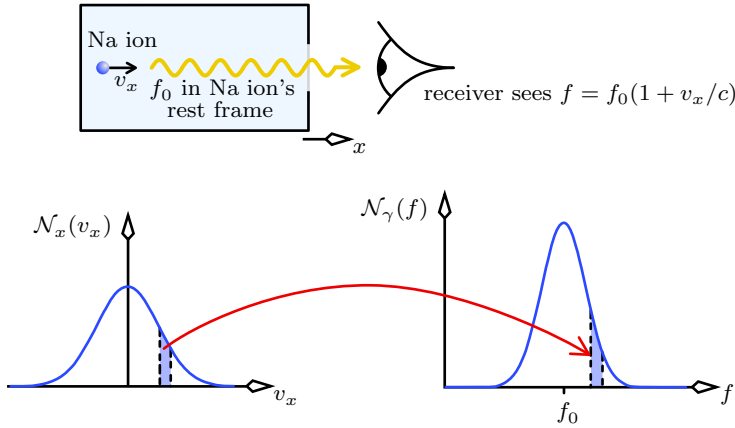


Fig. 6.6 Top: A sodium ion in a gas has some x component of velocity (v_x) in the direction of the receiver (the eyeball). It emits a single photon of frequency f_0 in its rest frame, which is Doppler-shifted to frequency f as seen by the receiver. **Bottom:** The top scenario as density plots. The left-hand plot shows the Maxwell distribution of v_x . Each atom with a velocity in the range v_x to $v_x + dv_x$ emits a photon that is Doppler-shifted to the frequency range f to $f + df$, shown in the right-hand plot

The setup is shown in Figure 6.6. A box contains a gas of sodium ions whose velocities are Maxwell distributed according to their temperature T . A sodium ion with x velocity v_x toward the receiver (the eyeball in the figure) emits a single photon of frequency f_0 in the ion's rest frame (that is, f_0 is either 589.0 nm or 589.6 nm). This photon is Doppler-shifted to frequency f when detected by the receiver. This non-relativistic Doppler shift is

$$f = f_0 \left(1 + \frac{v_x}{c} \right). \quad (6.61)$$

(It is more natural to work with frequency rather than wavelength here, because the distribution of Doppler-shifted frequencies will turn out to be gaussian.) Because each atom with an x velocity in the range v_x to $v_x + dv_x$ emits a single photon in the frequency range f to $f + df$, the frequency density of the Doppler-shifted photons is written as $\mathcal{N}_\gamma(f)$ (the “ γ ” denotes a photon), where

$$\mathcal{N}_\gamma(f) df \equiv \mathcal{N}_x(v_x) dv_x. \quad (6.62)$$

$\mathcal{N}_x(v_x)$ is easily found by replacing z with x in (6.33), where the sodium ions have mass m . Also, equation (6.61) says that $v_x = c(f/f_0 - 1)$. Then, (6.62) gives the photon frequency density as

$$\mathcal{N}_\gamma(f) = \mathcal{N}_x(v_x) \frac{dv_x}{df} = \frac{c}{f_0} N_{\text{tot}} \sqrt{\frac{m}{2\pi kT}} \exp \frac{-mc^2(f - f_0)^2}{2kT f_0^2}. \quad (6.63)$$

This is a gaussian centred on f_0 . Recalling (1.103), it has a characteristic spread of σ , where

$$2\sigma^2 = \frac{2kTf_0^2}{mc^2}, \quad \text{or} \quad \sigma = \sqrt{\frac{kT}{m}} \frac{f_0}{c} = \sqrt{\frac{kT}{m}} \frac{1}{\lambda_0}, \quad (6.64)$$

with λ_0 being the wavelength corresponding to frequency f_0 . We presume $\sigma \ll f_0$, meaning the gaussian function of frequency in this idealised model falls to zero at small frequencies, as shown in the right-hand plot in Figure 6.6. Is this correct? This presumed inequality can be written as $\sqrt{kT/m} f_0/c \ll f_0$, which is equivalent to $\sqrt{kT/m} \ll c$. And we certainly know that this last inequality holds, because

$$\sqrt{\frac{kT}{m}} \stackrel{(6.55)}{=} \frac{v_{\text{rms}}}{\sqrt{3}} \ll c, \quad (6.65)$$

since the speeds of the sodium ions are non-relativistic. So, it's certainly true that $\sigma \ll f_0$.

What is the frequency range of validity of (6.63), given that our scenario is non-relativistic? The non-relativistic Doppler expression in (6.61) assumes that $|v_x| \ll c$. Combining (6.61) with $-c \ll v_x \ll c$ leads to

$$-1 \ll f/f_0 - 1 \ll 1, \quad (6.66)$$

or

$$0 \ll f \ll 2f_0. \quad (6.67)$$

But almost all of the support⁶ of $\mathcal{N}_\gamma(f)$ is confined to the frequency range $[f_0 - 3\sigma, f_0 + 3\sigma]$; then, since $\sigma \ll f_0$, we see that effectively all of this support is confined close to f_0 , and certainly satisfies (6.67). Equation (6.63) thus holds for effectively all frequencies of interest.

The above analysis was more naturally carried out using frequency rather than wavelength, since the frequency distribution turned out to be gaussian; but our goal is find how much Doppler broadening the doublet of wavelengths undergoes. In other words, we wish to compare the wavelengths corresponding to the two frequencies f_0 and $f_0 + \sigma$. These wavelengths are, respectively, $c/f_0 = \lambda_0$ and $c/(f_0 + \sigma)$. These two wavelengths differ by

$$\sigma_\lambda \equiv \frac{c}{f_0} - \frac{c}{f_0 + \sigma} \simeq \frac{c\sigma}{f_0^2} = \frac{\lambda_0^2 \sigma}{c}. \quad (6.68)$$

We need σ :

$$\sigma = \sqrt{\frac{kT}{m}} \frac{1}{\lambda_0} = \sqrt{\frac{RT}{M_{\text{mol}}}} \frac{1}{\lambda_0}, \quad (6.69)$$

⁶ Meaning the values of f for which $\mathcal{N}_\gamma(f)$ is non-zero.

where R is the gas constant and $M_{\text{mol}} = 23.0$ g is sodium's molar mass. Then, with a sodium temperature of roughly 300°C ,

$$\sigma \simeq \sqrt{\frac{8.314 \times 573}{23.0}} \times \frac{1}{589} \text{ Hz} \simeq 7.7 \times 10^8 \text{ Hz}. \quad (6.70)$$

(Compare this with $f_0 = 2.998 \frac{8}{589} \text{ Hz} \simeq 5.1 \times 10^{14} \text{ Hz}$. Thus, $\sigma \ll f_0$, as expected.) Finally,

$$\sigma_\lambda = \frac{\lambda_0^2 \sigma}{c} \simeq \frac{(589)^2 \times 7.7}{2.998} \text{ m} \simeq 0.001 \text{ nm}. \quad (6.71)$$

This 0.001 nm is an indicator of the width of the wavelength distribution, so let us say that each line in the sodium doublet is broadened by several times this,⁷ giving a width of, say, 0.005 nm. This width is far smaller than the two lines' separation of 0.6 nm; hence, we expect the lines to be fully discernible as a doublet when light from a sodium lamp is put through a grating or prism. And indeed, the lines *are* fully discernible.

6.5 Temperature Gradient in a Weatherless Atmosphere

Throughout this book—apart from the introductory calculations of this chapter—we have used the simplifying assumption that an atmosphere's temperature is independent of altitude z . But the atmosphere derives its temperature mostly from Earth's warm surface, and the atmospheric particles at sea level slow down as they climb in Earth's gravitational field. Slower particles cannot climb to great heights, and faster particles turn into slower particles at these altitudes. Perhaps, then, the Maxwell distribution might turn out to have the same form at *all* altitudes—and thus give the same value of mean speed, and hence temperature, at all altitudes. To investigate, we must calculate the distribution at altitude z .⁸

The question of how an atmosphere's temperature might be a function of altitude was debated by three of the greatest physicist/chemists of the mid-to-late nineteenth century. On one side, Josef Loschmidt argued that a temperature gradient existed, but that this implied an ability to extract energy for free, by using the gradient to run a heat engine. On the other side,

⁷ That is, a gaussian distribution with standard deviation σ has a characteristic width of "several σ ".

⁸ The discussion in this section is not the same as the analysis given by Feynman in Volume 1 of his Lectures on Physics, Section 40-4. Feynman uses the simplifying assumption that temperature is independent of altitude to derive the Maxwell velocity distribution; in contrast, we begin with the Maxwell speed distribution at sea level and do not assume that temperature is independent of altitude.

Ludwig Boltzmann and James Clerk Maxwell turned this argument around to say that since we cannot expect ever to extract energy for free, an atmosphere cannot have a temperature gradient.

The idea of using a temperature gradient to run a heat engine is often pictured by running, say, a copper wire vertically in the atmosphere, and reasoning that the higher temperature at the base of the wire will create a heat flow up the wire that lasts forever—which clearly breaks energy conservation. But this picture is not physical. Consider an electrical analogy: an *electric* potential gradient certainly *does* exist in our atmosphere: it creates (or rather, is) an electric field of about 100 volts/metre pointing down; and yet, a permanent electric current certainly does not exist in the same wire. The reason is that the wire does not have access to an inexhaustible supply of electric charge. The free electrons in a copper wire will quickly arrange themselves to counter the field, and after this minuscule current dies down on a very short time scale, no more current can flow. (For the same reason, this electric field poses no threat to life, because the charges in our body are always quickly rearranging themselves to produce zero total field.)

Similar to that electrostatic situation, “heat flow” is really a flow of energy facilitated by the motion of particles; these particles pass energy via small motions along the wire. But particles are subject to Earth’s gravity, and they cannot pass energy upward without paying the gravitational tax man. Thus, this energy does not simply all pop out of the top of the wire. It follows that a temperature gradient can certainly exist without being exploitable to break energy conservation.⁹

Investigating whether such a temperature gradient exists has scope for the analysis of models of increasing complexity. We will make such a study using a simplified atmosphere: one that has no weather, so that no large-scale motion of air occurs (and thus no wind exists). We take the Maxwell distribution to hold at sea level, where the atmosphere contacts the ground at a shared temperature T_0 .

First, suppose that the atmosphere does not exist, and we start to create it by placing a thin horizontal slab of air on the ground. The N_{tot} particles in this slab are now allowed to escape by free expansion, moving upward, with the only work done being that against gravity. We suppose that a continuous mixing of horizontal and vertical velocities always occurs; hence, locally, the mean x , y , and z speeds are always equal. We will thus consider *speeds* of particles in the following analysis, as opposed to their z velocities.

Most particles of our atmosphere are diatomic molecules with five quadratic energy terms: three translational and two rotational. As they climb, some of their translational energy is converted to gravitational potential energy; also,

⁹ It is sometimes said that “Heat will flow up in such an atmosphere until the temperature is equalised throughout”. But thermodynamics taught us long ago that heat is not some nebulous thing that magically seeps upward without losing energy to a gravitational potential.

their continual interactions force some of their rotational energy to bleed into translational modes. Analysing this scenario is difficult, and we will instead examine the much simpler case of *monatomic* molecules, since the energy of these is purely translational.

To keep this discussion notationally independent of anything involving a Maxwell distribution, we introduce the following density function for our slab of particles:

$$f_z(v) dv \equiv \left[\begin{array}{c} \text{number of particles from slab arriving at} \\ \text{altitude } z \text{ with speeds in } v \text{ to } v + dv \end{array} \right]. \quad (6.72)$$

We require the speed density $f_z(v)$ of particles at height z : the number of particles arriving at z per unit speed. At sea level ($z = 0$), the number of particles $f_0(u) du$ in the speed interval $[u, u + du]$ of our horizontal slab is given by Maxwell's speed distribution (6.37):

$$f_0(u) du = \mathcal{N}_{\text{sp}}(u) du \stackrel{(6.37)}{=} \alpha u^2 \exp \frac{-mu^2}{2kT_0} du, \quad (6.73)$$

where, for this analysis, we don't need to write explicitly the normalisation

$$\alpha = N_{\text{tot}} \sqrt{2/\pi} [m/(kT_0)]^{3/2}. \quad (6.74)$$

We let these particles climb to altitude z . As they climb, they slow to form the speed interval $[v, v + dv]$. It follows that

$$f_z(v) dv = f_0(u) du. \quad (6.75)$$

The number of particles arriving at z per unit speed is thus

$$f_z(v) = f_0(u) \frac{du}{dv}. \quad (6.76)$$

The energy of one of these particles at sea level is $1/2 mu^2$. At altitude z , this (unchanged) total energy is $1/2 mv^2 + mgz$. Hence,

$$1/2 mu^2 = 1/2 mv^2 + mgz. \quad (6.77)$$

It follows that

$$u = \sqrt{v^2 + 2gz}, \quad \text{with} \quad \frac{du}{dv} = \frac{v}{u}. \quad (6.78)$$

Equation (6.76) then becomes

$$\begin{aligned} f_z(v) &= f_0(u) \frac{v}{u} \stackrel{(6.73)}{=} \frac{v}{u} \alpha u^2 \exp \frac{-mu^2}{2kT_0} \\ &= \alpha v \sqrt{v^2 + 2gz} \exp \left[\frac{-1}{kT_0} (1/2 mv^2 + mgz) \right]. \end{aligned} \quad (6.79)$$

We finally have the number of *monatomic particles* per unit speed arriving at height z :

$$f_z(v) = \alpha v \sqrt{v^2 + 2gz} \exp \frac{-1/2 mv^2 - mgz}{kT_0}. \quad (6.80)$$

This density has a Boltzmann factor that involves a particle's total energy $1/2 mv^2 + mgz$ at altitude z , as expected. Evident too is an exponential decrease with altitude z . The factor in front of the exponential is not the v^2 that occurs in the usual Maxwell speed distribution, but rather $v\sqrt{v^2 + 2gz}$. By how much does the factor $\sqrt{v^2 + 2gz}$ differ from v ? Referring to Table 6.1 for representative speeds in our atmosphere, set $v = 400$ m/s. Then, at an altitude of $z = 100$ metres,

$$\frac{\sqrt{v^2 + 2gz}}{v} \simeq \frac{\sqrt{400^2 + 2 \times 9.8 \times 100}}{400} \simeq 1.01. \quad (6.81)$$

At $z = 1$ km, the ratio is 1.06, and at $z = 10$ km, it is 1.5. It's apparent that the speed distribution (6.80) differs little from the sea-level Maxwell distribution throughout the lower part of our atmosphere. But this modified distribution peaks at a lower speed than the sea-level value of $\sqrt{2kT/m}$ in (6.47). We can calculate the location $\hat{v}$ of this peak by setting $f'_z(\hat{v}) = 0$. Hence, differentiating (6.80) and setting the result equal to zero produces

$$\frac{-m\hat{v}^4}{kT_0} + \hat{v}^2 \left(2 - \frac{2mgz}{kT_0} \right) + 2gz = 0. \quad (6.82)$$

This is a quadratic in $\hat{v}^2$, with one positive and one negative solution. We choose the positive one, since $\hat{v}^2$ is positive:

$$\hat{v}^2 = \frac{2kT_0}{m} - gz + \frac{(gz)^2}{2kT_0/m}. \quad (6.83)$$

This clearly reduces to the Maxwell value $\hat{v}^2 = 2kT_0/m$ at sea level, or for no gravity. For altitudes of around a kilometre in our atmosphere, the terms on the right-hand side of (6.83) are as follows:

$$\begin{aligned} 2kT_0/m &\simeq 166,000 \text{ m}^2/\text{s}^2, & gz &\simeq 10,000 \text{ m}^2/\text{s}^2, \\ \frac{(gz)^2}{2kT_0/m} &\simeq 600 \text{ m}^2/\text{s}^2. \end{aligned} \quad (6.84)$$

This suggests that we drop the last term in (6.83). Taking the square root then gives us

$$\hat{v} \simeq \left(\frac{2kT_0}{m} - gz \right)^{1/2} = \sqrt{\frac{2kT_0}{m}} \left(1 - \frac{gz}{2kT_0/m} \right)^{1/2}$$

$$\simeq \sqrt{\frac{2kT_0}{m}} \left(1 - \frac{gz}{4kT_0/m} \right). \quad (6.85)$$

For $z = 1$ km, this becomes [referring to (6.84)]

$$\hat{v} \simeq \sqrt{\frac{2kT_0}{m}} \left(1 - \frac{10,000}{332,000} \right) \simeq \sqrt{\frac{2kT_0}{m}} \times 0.97. \quad (6.86)$$

The peak of the distribution has shifted to a slightly lower speed. In this approximation of low altitudes, we define an equivalent temperature T at altitude z , such that the Maxwell distribution's peak for T equals the actual peak $\hat{v}$ in (6.85):

$$\sqrt{\frac{2kT}{m}} \equiv \sqrt{\frac{2kT_0}{m}} \left(1 - \frac{gz}{4kT_0/m} \right). \quad (6.87)$$

It follows that

$$T \simeq T_0 - \frac{mgz}{2k}. \quad (6.88)$$

The temperature decrease per unit height is then

$$\frac{-dT}{dz} = \frac{mg}{2k} = \frac{1/2 M_{\text{mol}} g}{R}. \quad (6.89)$$

For our atmosphere, this is

$$\frac{-dT}{dz} \simeq \frac{1/2 \times 0.0290 \times 9.8}{8.314} \text{ K/m} \simeq 17 \text{ K/km}. \quad (6.90)$$

Compare (6.89)'s $1/2 M_{\text{mol}} g/R$ with (6.9)'s $2/3 M_{\text{mol}} g/R$. These expressions are very similar, because they were both built on the idea of particles losing speed as they climb in Earth's gravitational field. But the expressions are not identical, because whereas (6.9) gives the temperature drop of a *single set* of particles in an atmosphere where all particles have the *same speed* at sea level, equation (6.89) refers to the speed difference of *two different sets* of particles evolving from a Maxwell distribution: those in the speed distribution's peak at sea level, and those in the speed distribution's peak at altitude z .

This simple model of “building an atmosphere” by laying down slabs of air molecules and allowing them to drift upward has predicted a linear temperature gradient of roughly the measured value. We reiterate that this temperature gradient cannot be used to power a heat engine to produce energy for free. Hoping to do so is akin to a mediaeval archer standing at the base of a very high castle, aiming his arrow upward with full string tension, and expecting it to do much damage when it arrives at the castle's top. As discussed at the start of this section, we cannot hold a copper wire vertically with its ends open to the atmosphere and expect heat to flow up through the wire.

Is this atmosphere, with its altitude-dependent temperature, in equilibrium? Yes: although its bulk parameters vary with altitude, they do not change with time. When we defined temperature in Section 3.5, we assumed there to be no background potential energy that would force particles in some particular direction. This assumption doesn't apply to our atmosphere, so we must modify the idea of thermal equilibrium to account for the presence of gravity. Even though the atmosphere gets colder with height, it is still in thermal equilibrium, because the temperature differences do not give rise to a heat current. Instead, the exchanges of energy that would normally constitute a heat current now occur within the gravitational potential.

The above analysis assumed that the particles of our atmosphere are monatomic. Of course, our real atmosphere has mostly diatomic molecules of N_2 and O_2 , and each of these has two rotational modes that it can store energy in. As those particles climb and lose translational energy to the gravity field, rotational energy is continually being passed into translational modes, with the result that the particles don't slow down with altitude as much as we calculated above. So, we expect the true value of $-dT/dz$ to be somewhat less than the value of 17 K/km calculated above. Also, the discussion in this section applied to a weatherless atmosphere; we expect the mixing effect of weather to flatten the temperature gradient still further. And indeed, as mentioned just after (6.10), the measured average temperature drop for dry air in our troposphere is about 10 K/km. For real air (which isn't always dry), the measured value is about 7 K/km.

Physicists have not reached a consensus on the question of a temperature gradient in a weatherless atmosphere. Most authors simply declare by fiat that temperature is independent of altitude. Analyses are rare and have varying degrees of fidelity, and so produce differing results.

Our real atmosphere is dominated by weather patterns arising from the Coriolis acceleration produced by Earth's spin, and these patterns induce heavy mixing. When a parcel of air is pushed up by bulk air movements, it encounters lower-pressure air that it expands into, doing work $PdV > 0$ in the process, and losing energy as a result. This lowers its temperature, and can produce steeper temperature gradients than calculated above. If this air is very moist, this sharp temperature drop can produce clouds and rainfall.

6.6 Gaseous Makeup of Planetary Atmospheres

For the idealised continuum of particles that the Maxwell distribution describes, the tail of the distribution extends to arbitrarily high values of speed. We know, of course, that the speed of light sets an upper limit on these values; but the value of $\mathcal{N}_{\text{sp}}(u)/N_{\text{tot}}$ in Figure 6.4 falls so quickly with increasing u that, for all intents and purposes, it is zero for speeds that are still a negligible fraction of the speed of light. We thus lose nothing by setting the highest

possible speed to infinity in any non-relativistic treatment of the Maxwell distribution.

Because particles in a gas have speeds around $\sqrt{kT/m}$, lighter molecules tend to move more quickly than massive ones. Maxwell's distribution of speeds can tell us whether very light molecules in Earth's atmosphere are moving so quickly as to leak away from Earth's atmosphere entirely. In the following analysis, we will use -20°C or about 253 K as a representative temperature of our current atmosphere, and will then use the standard Maxwell distribution throughout. This is simpler than using the modified form (6.80), while nonetheless still approximating that equation—which was only derived for monatomic particles anyway.

The fastest-moving gas will be the lightest: hydrogen. Its rms speed at this temperature is, from (6.55),

$$v_{\text{rms}}(\text{H}_2) = \sqrt{\frac{3RT}{M_{\text{mol}}}} \simeq \sqrt{\frac{3 \times 8.314 \times 253}{0.002}} \text{ m/s} \simeq 1.8 \text{ km/s}. \quad (6.91)$$

The rms speed of the 16-times heavier oxygen is a quarter of this:

$$v_{\text{rms}}(\text{O}_2) \simeq \sqrt{\frac{3 \times 8.314 \times 253}{0.032}} \text{ m/s} \simeq 444 \text{ m/s}. \quad (6.92)$$

How do these speeds compare with the *escape speed* from Earth, v_{esc} ? Escape speed is defined as the minimum speed required by a freely moving mass at Earth's surface for it to move arbitrarily far from Earth.¹⁰ To calculate v_{esc} , consider that a mass with speed v_{esc} at Earth's surface will only *just* reach infinity—meaning it has zero total energy when it “gets there”. Since energy is conserved, the mass must then have zero total energy everywhere, including at Earth's surface. The mass m 's potential energy at distance r from Earth's centre is $-GMm/r$, where G is the gravitational constant and M is Earth's mass: $GM \simeq 3.9860 \times 10^{14}$ SI units. Its total energy at Earth's surface—and hence everywhere else too—is then

¹⁰ Escape speed is usually called “escape velocity”, but it is undoubtedly a *speed*: the question of whether a particle can escape Earth's gravity is only one of total energy, and so the direction of the particle's motion has no bearing on the analysis. You might ask, if escape speed is a speed and not a velocity, why is there a “ v ” in the symbol v_{esc} for it? The reason is that speed is the length of a velocity vector, and since velocity is usually written with a bold-face $\mathbf{v}$, its length tends to be written as a non-bold v . This seemingly straightforward piece of notation can be a trap for young players in even simple calculations of kinematics. Physicists usually typeset a vector in bold ($\mathbf{v}$), and write it by hand using the same symbol, but non bold, with a tilde underneath or an arrow on top ($\vec{v}$ or $\vec{v}$); and its length (a positive number) is then written with the same symbol, but not bold and no tilde (v). But in one dimension this becomes slightly problematic, because a one-dimensional vector (which is coordinatised as a real number, positive *or* negative) tends to be written with a non-bold or non-tilded symbol (v)—which is then easily confused with its length, a *non-negative* number.

$$\frac{1}{2}mv_{\text{esc}}^2 - GMm/R_{\text{Earth}} = 0, \quad (6.93)$$

where $R_{\text{Earth}} \simeq 6370$ km is Earth's radius. The escape speed then follows as

$$v_{\text{esc}} = \sqrt{2GM/R_{\text{Earth}}}. \quad (6.94)$$

For Earth, this amounts to

$$v_{\text{esc}} = \sqrt{\frac{2 \times 3.9860 \frac{14}{\text{m}^3 \text{s}^2}}{6.37 \frac{6}{10^6} \text{m}}} \text{ m/s} \simeq 11.2 \text{ km/s}. \quad (6.95)$$

Recall, from (6.91) and (6.92), that hydrogen and oxygen have speeds of about one tenth this value. At first sight, it might then seem that both hydrogen and oxygen have insufficient speeds to escape Earth's gravity, making these gases remain in Earth's atmosphere indefinitely—barring chemical reactions that lock the gases into Earth's crust. But remember that the Maxwell speed distribution has a tail that extends to arbitrarily high values of speed. This means the fastest-moving particles can indeed escape Earth's surface. If we imagine all of these particles escaping at some given moment, then after some relaxation time Δt , the remaining particles will re-acquire a Maxwell distribution of speeds—and then some of *those* will now have speeds greater than v_{esc} , permitting *them* to leave. It seems that particles will indeed gradually leak away from Earth. The question is: given Earth's great age, should we be surprised that we still have an atmosphere?

This continuous leakage of lighter gas particles can be modelled in various ways. We will set the relaxation time to be that which represents a global rearrangement of the gas molecules, and so will set it to be the time required for the particles of an average speed in the gas to make one up/down trip in the atmosphere. It follows that $\Delta t/2$ is the time taken for one of these average-speed particles to travel from sea level to its maximum altitude. This upward motion involves the vertical (z) component of the velocity, and so we will use $v_{z,\text{rms}}$ as the required average speed. Earth's atmospheric layer is thin enough that the acceleration g due to gravity can be treated as a constant throughout. The standard expression for constant acceleration “ $v = u + at$ ” then becomes, for the ascending particle,

$$0 = v_{z,\text{rms}} - g\Delta t/2, \quad \text{or} \quad \Delta t = 2v_{z,\text{rms}}/g. \quad (6.96)$$

This is the up/down trip time for a particle of average speed. In this time, particles moving upward with speed v_z will make $v_z/v_{z,\text{rms}}$ trips to the top of the atmosphere; but we are interested particularly in particles whose speed is v_{esc} . Given a speed v , what is the average speed $\langle v_z \rangle$ of upward-moving particles? These particles are moving at an angle θ to the horizontal, where θ runs from zero to $\pi/2$. Hence,

$$\langle v_z \rangle = \langle v \sin \theta \rangle_{\theta=0 \rightarrow \pi/2} = \frac{v}{\pi/2} \int_0^{\pi/2} \sin \theta \, d\theta = \frac{2v}{\pi}. \quad (6.97)$$

Particles moving at speed v_{esc} with some upward motion will have an average v_z of $2v_{\text{esc}}/\pi$, and so will make $(2v_{\text{esc}}/\pi)/v_{z,\text{rms}}$ trips to the top of the atmosphere. The same can be said for downward-moving particles of speed v_{esc} . In the relaxation time Δt , we then expect there to be $(4v_{\text{esc}}/\pi)/v_{z,\text{rms}}$ “surfacing” of fast particles to the top of the atmosphere, where they can escape. Let one surfacing of a group of these fast particles happen in every period of time Δt_{surf} . Then,

$$\Delta t = \frac{4v_{\text{esc}}/\pi}{v_{z,\text{rms}}} \Delta t_{\text{surf}}. \quad (6.98)$$

Rearrange (6.98), giving

$$\Delta t_{\text{surf}} = \frac{v_{z,\text{rms}}}{4v_{\text{esc}}/\pi} \Delta t \stackrel{(6.96)}{=} \frac{v_{z,\text{rms}}}{4v_{\text{esc}}/\pi} \frac{2v_{z,\text{rms}}}{g} = \frac{\pi}{2g} \frac{v_{z,\text{rms}}^2}{v_{\text{esc}}}. \quad (6.99)$$

This use of a cycle of motion to model what is really a continuous process is not as true to reality as more sophisticated models would be. But it at least approximates reality while still being tractable.

Suppose that at each “surfacing” (that is, after each time Δt_{surf}), all of the particles with escape speed or higher either do escape, or are effectively “scheduled” to escape later. These form a fraction f of the total number of particles, and are in the tail of the Maxwell speed distribution. So, if we approximate $\mathcal{N}_{\text{sp}}(z, v)$ by $\mathcal{N}_{\text{sp}}(v)$, then

$$f = \int_{v_{\text{esc}}}^{\infty} \frac{\mathcal{N}_{\text{sp}}(v) \, dv}{N_{\text{tot}}}. \quad (6.100)$$

Imagine, for a moment, that $f = 1/10$ of the particles escape in each period of time Δt_{surf} . That does not imply that *all* of the particles will have escaped after a time $\Delta t_{\text{surf}}/f = 10\Delta t_{\text{surf}}$; but a significant number will have gone.

We can quantify just what is meant by “a significant number” in the following way. Suppose that after each period of time Δt_{surf} , the fraction f leaves all at once. We’ll abbreviate Δt_{surf} to “ T ” in the equations to follow. Now use a limit idea that will convert this process to one that occurs continuously: say that, after each “time step” T/n for some number n (that will eventually go to infinity), a fraction f_n leave. We’ll take n to infinity by first relating f_n to f . Write down the number N of particles remaining after each time step:

$$\begin{aligned} t : 0 &\rightarrow T/n && \rightarrow 2T/n && \rightarrow \dots \rightarrow nT/n = T \\ N : N_{\text{tot}} &\rightarrow N_{\text{tot}}(1 - f_n) && \rightarrow N_{\text{tot}}(1 - f_n)^2 && \rightarrow \dots \rightarrow N_{\text{tot}}(1 - f_n)^n. \end{aligned} \quad (6.101)$$

The number of particles remaining after T is $N_{\text{tot}}(1 - f_n)^n$; but by definition of f , this number remaining must equal $N_{\text{tot}}(1 - f)$:

$$N_{\text{tot}}(1 - f_n)^n = N_{\text{tot}}(1 - f). \quad (6.102)$$

It follows that

$$f_n = 1 - (1 - f)^{1/n}. \quad (6.103)$$

As n tends to infinity, write T/n as dt . Equation (6.103) becomes, with a binomial expansion,

$$f_\infty = 1 - (1 - f)^{dt/T} = 1 - \left(1 - \frac{dt}{T}f\right) = \frac{f dt}{T}. \quad (6.104)$$

Recall that in a time T/n , a fraction f_n leave. Thus, in a time $dt = T/n$ with $n \rightarrow \infty$, a fraction $f_\infty = f dt/T$ leave. But this fractional loss in particle number N is, by definition, equal to $-dN/N$:

$$\frac{-dN}{N} = \frac{f dt}{T}. \quad (6.105)$$

This integrates to

$$N = N_{\text{tot}} e^{-ft/T}. \quad (6.106)$$

We see that the number of particles remaining decreases exponentially with time.

This sort of exponential (6.106) is ubiquitous in the subject of radioactive decay, whose mathematics can be modelled in a similar way to the above argument. In the above case, $-dN$ particles are lost from the atmosphere in a time dt at each moment t , implying that these $-dN$ particles have survived for a “lifetime” of t . All of the other particles have lifetimes either less than or more than this. The average lifetime of all the particles is then calculated in the usual way of calculating a mean. For an example with simple numbers, if 3 particles each survive for 10 seconds and 2 particles each survive for 50 seconds, then their average lifetime is $(3 \times 10 + 2 \times 50)/(3 + 2)$ seconds. Similarly here, the particles’ mean lifetime of being in Earth’s atmosphere is

$$\text{mean lifetime} = \frac{1}{N_{\text{tot}}} \int_0^{N_{\text{tot}}} (-dN \times t). \quad (6.107)$$

Find $-dN$ from (6.105) or (6.106):

$$-dN = N_{\text{tot}} e^{-ft/T} f dt/T. \quad (6.108)$$

Substitute this into (6.107), to obtain

$$\begin{aligned} \text{mean lifetime} &= \frac{1}{N_{\text{tot}}} \int_0^\infty N_{\text{tot}} e^{-ft/T} \frac{f dt}{T} \times t \\ &= \frac{f}{T} \int_0^\infty t e^{-ft/T} dt = T/f. \end{aligned} \quad (6.109)$$

We see that the particles' mean lifetime is T/f . Alternatively, ask how long it would take for *all* of the particles to vanish *if* their rate of decay equalled their initial rate of decay in (6.106); specifically, find the intersection of the tangent to the curve of N versus t at $t = 0$ with the t axis. This is a simple exercise in calculus, and this length of time also turns out to be $t = T/f$. The bottom line is that we can consider a “significant number” of particles to have decayed after a time T/f . This is why we said, for the $f = 1/10$ example just after (6.100) above, that a significant number of particles will have escaped the atmosphere after a time $\Delta t_{\text{surf}}/f = 10\Delta t_{\text{surf}}$.

To re-iterate, the time required for significant depletion of atmospheric particles is $\Delta t_{\text{surf}}/f$. (We now abandon our shorthand T for Δt_{surf} .) We must calculate Δt_{surf} and f . We found Δt_{surf} in (6.99). That equation requires $v_{z,\text{rms}}$, which can be found from the equipartition theorem:

$$\frac{1}{2} m v_{z,\text{rms}}^2 = \frac{1}{2} m \langle v_z^2 \rangle = \langle \frac{1}{2} m v_z^2 \rangle = \frac{1}{2} kT. \quad (6.110)$$

It follows from this that

$$v_{z,\text{rms}} = \sqrt{kT/m}. \quad (6.111)$$

Alternatively—just to show that everything is self consistent—we can return to first principles, (1.157), to write

$$v_{z,\text{rms}}^2 \equiv \langle v_z^2 \rangle = \int_{-\infty}^{\infty} v_z^2 \frac{\mathcal{N}_z(v_z) dv_z}{N_{\text{tot}}}. \quad (6.112)$$

We wrote $\mathcal{N}_z(v_z)$ down in (6.33). Alternatively, we can produce $\mathcal{N}_z(v_z)$ by integrating over v_x and v_y in (6.22), using $v^2 = v_x^2 + v_y^2 + v_z^2$ [recall (6.14)]:

$$\begin{aligned} \mathcal{N}_z(v_z) dv_z &= \int \int_{v_x, v_y = -\infty}^{\infty} \mathcal{N}_{\text{vel}}(\mathbf{v}) dv_x dv_y dv_z \\ &\stackrel{(6.22)}{=} N_{\text{tot}} \left(\frac{m}{2\pi kT} \right)^{3/2} \int \int_{-\infty}^{\infty} \exp \frac{-mv^2}{2kT} dv_x dv_y dv_z \\ &= N_{\text{tot}} \left(\frac{m}{2\pi kT} \right)^{3/2} \left[\int_{-\infty}^{\infty} \exp \frac{-mv_x^2}{2kT} dv_x \right]^2 \exp \frac{-mv_z^2}{2kT} dv_z. \end{aligned} \quad (6.113)$$

Call on (1.91) to evaluate the integral in brackets in the last line above, obtaining

$$\mathcal{N}_z(v_z) \cancel{dv_z} = N_{\text{tot}} \sqrt{\frac{m}{2\pi kT}} \exp \frac{-mv_z^2}{2kT} \cancel{dv_z}. \quad (6.114)$$

$\mathcal{N}_z(v_z)$ in (6.114) matches (6.33), as expected. We now substitute this expression for $\mathcal{N}_z(v_z)$ into (6.112). Calling on (1.101) then produces (6.111) again.

Now that we have $v_{z,\text{rms}}$ in (6.111), the next step is to calculate f from (6.100). That equation is an integral over the Maxwell speed distribution $\mathcal{N}_{\text{sp}}(v)$ in (6.37):

$$f = \sqrt{\frac{2}{\pi}} \left(\frac{m}{kT} \right)^{3/2} \int_{v_{\text{esc}}}^{\infty} v^2 \exp \frac{-mv^2}{2kT} dv. \quad (6.115)$$

Call on (1.97) to write this as

$$f = \sqrt{\frac{2m}{\pi kT}} v_{\text{esc}} \exp \frac{-mv_{\text{esc}}^2}{2kT} + \text{erfc} \left(\sqrt{\frac{m}{2kT}} v_{\text{esc}} \right). \quad (6.116)$$

We now have $v_{z,\text{rms}}$ in (6.111) and f in (6.116). Hence, we can calculate the representative time $\Delta t_{\text{surf}}/f$ for depletion, by using (6.99) for Δt_{surf} , (6.94) for v_{esc} , and, of course, the gravitational acceleration:

$$g = GM/R_{\text{Earth}}^2. \quad (6.117)$$

Table 6.2 shows representative depletion times $\Delta t_{\text{surf}}/f$ for a range of temperatures T of Earth's atmosphere. A range is used because we wish also to study a young, rapidly evolving Earth with a hot atmosphere. We see that in an atmosphere that is, say, 200 K hotter than Earth's currently is, we can expect to lose all hydrogen in a short time compared with Earth's current age. And indeed, hydrogen is not found in our atmosphere—although that might have something to do with its flammability. We can expect to keep most of our helium. In fact, helium is currently extracted from our atmosphere in commercial quantities, although it's not especially cheap to buy. Finally, we can expect to retain essentially *all* of the oxygen in our atmosphere.

Table 6.2 Representative depletion times $\Delta t_{\text{surf}}/f$ for various temperatures T of Earth's atmosphere

T	H ₂	He	O ₂
300 K	$4.3 \overset{14}{=} \text{years}$	$9.5 \overset{35}{=} \text{years}$	$\gg 10^{200} \text{ years}$
500 K	$1.8 \overset{6}{=} \text{years}$	$7.5 \overset{18}{=} \text{years}$	$3.7 \overset{200}{=} \text{years}$
700 K	539 years	$4.2 \overset{11}{=} \text{years}$	$1.0 \overset{141}{=} \text{years}$
900 K	6.6 years	$4.3 \overset{7}{=} \text{years}$	$9.4 \overset{107}{=} \text{years}$

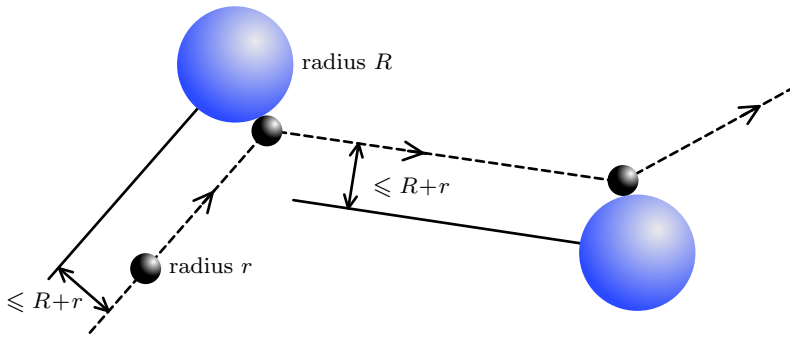


Fig. 6.7 A black particle of radius r collides repeatedly with large blue stationary particles of radius R , tracing out a crooked tube in the process

6.7 Mean Free Path of Gas Particles

A quantity of great importance in statistical mechanics is the *mean free path (length)* of a particle that collides repeatedly with others, whether they form a solid, liquid, or gas. Because the core idea here involves interacting *particles*, the theory of mean free path can make important predictions that bolster what is probably the most important theory in all of physics: the atomic theory of matter. Experiments that confirm these predictions have been crucial historically to the measurement of basic physical parameters, such as atomic sizes and Avogadro's number. And experiments that refuted these predictions have been just as important historically, because they have set bounds on how far classical atomic concepts can be pushed before physicists are obliged to introduce quantum concepts into the atomic model.

To lay out the ideas of mean free path and its stablemate, collision frequency, we will work with a gas comprising just one kind of distinguishable particle, for which the Maxwell velocity distribution will apply. These particles are very classical, modelled as tiny billiard balls. We wish to calculate λ , the particles' *mean free path*, being the mean distance that a particle travels between collisions. We assume the gas isn't too dense, so that each particle spends most of its time in free flight. This is a very good approximation for all manner of gases—even for pseudo-gases such as conduction electrons moving about in a metal, to be studied in Chapter 8.

Although we'll ultimately deal with just one type of particle, the scenario is easier to describe if we imagine a single “black” particle bouncing pinball-style through a lattice of stationary “blue” particles in Figure 6.7. Suppose the black particle travels with speed v for a time Δt . As it bounces off the other particles, it traces out the crooked tube of length $v\Delta t$ in the figure. The black particle's mean free path is

$$\lambda = \frac{\text{tube length } (= v\Delta t)}{\text{number of collisions in tube}}. \quad (6.118)$$

The blue particles are all at rest in the laboratory, and so are moving with relative speed v past the black particle. The number of collisions will then equal the number of blue particles whose centres are in this tube. For a blue particle density of ν particles per unit volume, we have

$$\begin{aligned} \text{number of collisions in tube} &= \text{number of blue particle centres in tube} \\ &= \nu \times \text{volume of tube} \\ &= \nu \times \text{tube cross-sectional area} \times \text{tube length} \\ &= \nu \sigma v \Delta t, \end{aligned} \quad (6.119)$$

where σ is the tube's cross-sectional area, also called the *collision cross section* for this scenario. For this area, consider that any blue particle whose centre is farther than $R + r$ from the centre of the black particle will not be struck by the black particle. The tube thus has radius $R + r$, and hence $\sigma = \pi(R + r)^2$ in this model. Equations (6.118) and (6.119) combine to give

$$\lambda = \frac{v\Delta t}{\nu \sigma v \Delta t} = \frac{1}{\nu \sigma}. \quad (6.120)$$

In practice, the blue particles are *not* at rest in the laboratory, and so they do *not* move past the black particle with a relative speed v . Suppose instead, that they all pass the black particle with a relative speed v_{rel} , and we will disregard the finer points of averaging over the various directions from which they came. Then, the number of collisions (meaning blue particles encountered) by the black particle is *as if* it were travelling at v_{rel} for a time Δt through stationary particles. Equation (6.119) becomes

$$\text{number of collisions} = \nu \sigma v_{\text{rel}} \Delta t. \quad (6.121)$$

The tube length in the laboratory is still $v\Delta t$. Hence, (6.118) combines with (6.121) to yield

$$\lambda = \frac{v\Delta t}{\nu \sigma v_{\text{rel}} \Delta t} = \frac{v}{\nu \sigma v_{\text{rel}}}. \quad (6.122)$$

What is v_{rel} ? When the particles have a Maxwell velocity distribution, we replace v and v_{rel} with their means, $\bar{v}$ and $\bar{v}_{\text{rel}}$. At the end of this section, we'll show that

$$\bar{v}_{\text{rel}} = \bar{v}\sqrt{2}. \quad (6.123)$$

Hence, in the Maxwell-distributed case with all speeds replaced with their means, (6.122) becomes

$$\lambda \text{ (average)} = \frac{\bar{v}}{\nu \sigma \bar{v}_{\text{rel}}} = \frac{1}{\nu \sigma \sqrt{2}}. \quad (6.124)$$

The stablemate of the mean free path is a particle's *collision frequency*:

$$\text{each particle's collision frequency} \equiv \frac{\text{number of collisions in } \Delta t}{\Delta t}. \quad (6.125)$$

Refer to (6.121), but again, replace v and v_{rel} with their means, $\bar{v}$ and $\bar{v}_{\text{rel}}$:

$$\text{each particle's collision frequency (average)} = \frac{\nu \sigma \bar{v}_{\text{rel}} \Delta t}{\Delta t} \stackrel{(6.123)}{=} \nu \sigma \bar{v} \sqrt{2}. \quad (6.126)$$

Crowding of Air Molecules in a Room

What are the mean free path and collision frequency for air molecules in a standard room? Use the ideal-gas law $PV = NkT$ to write their particle density as

$$\nu = \frac{N}{V} = \frac{P}{kT}. \quad (6.127)$$

There is now no distinction between the black and blue particles in Figure 6.7, and so the cross section $\sigma = \pi(R + r)^2$ becomes, with $R = r$,

$$\sigma = \pi(r + r)^2 = 4\pi r^2, \quad \text{where } r \simeq 10^{-10} \text{ m}. \quad (6.128)$$

The molecules have a temperature of, say, $T = 298 \text{ K}$ and a pressure of $P = 10^5 \text{ Pa}$. Their mean free path is, from (6.124),

$$\lambda = \frac{1}{\nu \sigma \sqrt{2}} = \frac{kT}{P 4\pi r^2 \sqrt{2}} = \frac{1.381 \stackrel{=23}{\times} 298}{10^5 \times 4\pi \times 10^{-20} \times \sqrt{2}} \text{ m} \simeq 0.23 \text{ } \mu\text{m}. \quad (6.129)$$

Each particle's collision frequency is, from (6.126),

$$\begin{aligned} \nu \sigma \bar{v} \sqrt{2} &\stackrel{(6.53)}{=} \frac{P}{kT} 4\pi r^2 \sqrt{\frac{8kT}{\pi m}} \sqrt{2} = 16Pr^2 \sqrt{\frac{\pi}{mkT}} \\ &= 16Pr^2 N_A \sqrt{\frac{\pi}{M_{\text{mol}} RT}}. \end{aligned} \quad (6.130)$$

Air has a molar mass $M_{\text{mol}} = 29.0 \text{ g}$, and so (6.130) yields a collision frequency of (using SI units)

$$16 \times 10^5 \times 10^{-20} \times 6.022 \stackrel{=23}{\times} \sqrt{\frac{\pi}{0.0290 \times 8.314 \times 298}} \text{ s}^{-1} \simeq 2.0 \times 10^9 \text{ s}^{-1}. \quad (6.131)$$

Each particle collides two thousand million times per second. This is an astoundingly large number, but it reinforces the validity of the idea that gas pressure arises from a teeming mix of particles. These particles are

constantly interacting to re-establish equilibrium, and smoothen out local disturbances to the pressure.

6.7.1 Excursus: The Proof of (6.123)

When the particle velocities follow a Maxwell distribution, the $\bar{v}_{\text{rel}}$ in (6.123) is a *mean relative speed* of the particles. The process of calculating its value is a useful exercise in multi-dimensional calculus, and we'll work through its details here. (Nothing that follows this proof depends on it, and so it can be skipped on a first reading.) We will have use for three variables: $\mathbf{v}_{\text{rel}}$, v_{rel} , $\bar{v}_{\text{rel}}$, which we'll respectively write as $\mathbf{u}$, u , $\bar{u}$ for shorthand.

We wish to calculate the mean relative speed $\bar{u}$ from an expression that resembles (1.157), but is for a relative velocity $\mathbf{u}$ and its length, speed u :

$$\bar{u} = \int u p(\mathbf{u}) d^3u, \quad (6.132)$$

where $p(\mathbf{u})$ is the probability density for the relative velocity $\mathbf{u}$. We know that the particles' laboratory velocities are Maxwell distributed. Take two such particles, with lab velocities $\mathbf{v}_1$ and $\mathbf{v}_2$. Define their relative velocity $\mathbf{u}$, and their centre-of-mass velocity $\mathbf{U}$:

$$\mathbf{u} \equiv \mathbf{v}_1 - \mathbf{v}_2, \quad \mathbf{U} \equiv (\mathbf{v}_1 + \mathbf{v}_2)/2. \quad (6.133)$$

This is a change of variables, and we'll need its inverse:

$$\mathbf{v}_1 = \mathbf{U} + \mathbf{u}/2, \quad \mathbf{v}_2 = \mathbf{U} - \mathbf{u}/2. \quad (6.134)$$

Equation (6.132) requires the density $p(\mathbf{u})$. We know that

$$p(\mathbf{u}) = \int p(\mathbf{u}, \mathbf{U}) d^3U, \quad (6.135)$$

and so this converts (6.132) into

$$\bar{u} = \int u p(\mathbf{u}, \mathbf{U}) d^3u d^3U. \quad (6.136)$$

This is a six-dimensional integral, and we can relate it to the probability density $p(\mathbf{v}_1, \mathbf{v}_2)$ —known from Maxwell's theory—via the change of variables (6.133). Recall that the Maxwell velocity distribution is (6.22):

$$p(\mathbf{v}) = \left(\frac{m}{2\pi kT} \right)^{3/2} \exp \frac{-mv^2}{2kT} \equiv A \exp [-\alpha v^2], \quad (6.137)$$

where the constants A and α are a convenient shorthand. Now observe that the combined probability density of the two velocities $\mathbf{v}_1, \mathbf{v}_2$ is

$$p(\mathbf{v}_1, \mathbf{v}_2) = p(\mathbf{v}_1) p(\mathbf{v}_2) = A^2 \exp [-\alpha(v_1^2 + v_2^2)]. \quad (6.138)$$

Here is a short reminder of the relevant change-of-variables theory of integration. To convert an integral over, say, two variables x, y to new variables u, v , write

$$\iint f(x, y) dx dy = \iint f(x, y) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv, \quad (6.139)$$

where the “jacobian determinant” is¹¹

$$\frac{\partial(x, y)}{\partial(u, v)} \equiv \begin{vmatrix} \partial x / \partial u & \partial x / \partial v \\ \partial y / \partial u & \partial y / \partial v \end{vmatrix}. \quad (6.140)$$

Consider using this jacobian idea for the six-dimensional integral (6.136). Begin with what we want, $\bar{u}$ from (6.136), switch momentarily to $\mathbf{v}_1, \mathbf{v}_2$ coordinates to introduce Maxwell’s (6.138), then switch back to $\mathbf{u}, \mathbf{U}$ coordinates using the jacobian, because it will turn out that the integrals are easier to evaluate in $\mathbf{u}, \mathbf{U}$ coordinates:¹²

$$\begin{aligned} \bar{u} &= \int u p(\mathbf{u}, \mathbf{U}) d^3u d^3U \\ &= \int u p(\mathbf{v}_1, \mathbf{v}_2) d^3v_1 d^3v_2 \quad (\text{by definition of the probability density}) \\ &= \int u p(\mathbf{v}_1, \mathbf{v}_2) \left| \frac{\partial(v_{1x}, \dots, v_{2z})}{\partial(u_x, \dots, U_z)} \right| d^3u d^3U. \end{aligned} \quad (6.141)$$

Now call on (6.138), writing its v_1^2, v_2^2 in terms of $\mathbf{u}, \mathbf{U}$ using (6.134):

$$\begin{aligned} v_1^2 &= \mathbf{v}_1 \cdot \mathbf{v}_1 = U^2 + \mathbf{U} \cdot \mathbf{u} + u^2/4, \\ v_2^2 &= \mathbf{v}_2 \cdot \mathbf{v}_2 = U^2 - \mathbf{U} \cdot \mathbf{u} + u^2/4. \end{aligned} \quad (6.142)$$

¹¹ Note that in (6.139), the pair of vertical bars $|\cdot|$ denotes an absolute value, whereas in (6.140), $|\cdot|$ denotes a determinant. No confusion will arise if you remain aware of what is between the bars.

¹² On reflection, we might have guessed in advance that the integrals would be easier to evaluate in relative/centre-of-mass coordinates $\mathbf{u}, \mathbf{U}$. After all, those coordinates were created in the early days of classical mechanics precisely because they simplified calculations. They allowed calculations to be separated into disconnected relative and centre-of-mass scenarios, which could then be treated independently.

It follows that

$$p(\mathbf{v}_1, \mathbf{v}_2) \stackrel{(6.138)}{=} A^2 \exp[-\alpha(2U^2 + u^2/2)]. \quad (6.143)$$

The jacobian determinant in (6.141) requires (6.134) to be written in component form:

$$\begin{aligned} v_{1x} &= U_x + u_x/2, & v_{2x} &= U_x - u_x/2, \\ \vdots & & \vdots & \\ v_{1z} &= U_z + u_z/2, & v_{2z} &= U_z - u_z/2. \end{aligned} \quad (6.144)$$

The jacobian determinant in (6.141) is then¹³

$$\frac{\partial(v_{1x}, \dots, v_{2z})}{\partial(u_x, \dots, U_z)} = \begin{vmatrix} 1 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1/2 \\ 1 & 0 & 0 & -1/2 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1/2 \end{vmatrix} = -1. \quad (6.145)$$

The last line of (6.141) now becomes (with all integrals from $-\infty$ to ∞ , since we are integrating over components of velocity)

$$\begin{aligned} \bar{u} &= \int u A^2 \exp[-\alpha(2U^2 + u^2/2)] |-1| d^3u d^3U \\ &= A^2 \int u \exp[-\alpha u^2/2] d^3u \times \int \exp[-2\alpha U^2] d^3U. \\ &= \underbrace{8\pi/\alpha^2: \text{ see (1.117), (1.118) }}_{= [\pi/(2\alpha)]^{3/2}: \text{ see (1.114), (1.115) }} = [\pi/(2\alpha)]^{3/2} \int u \exp[-\alpha u^2/2] d^3u \\ &= A^2 \sqrt{8} \pi^{5/2} \alpha^{-7/2}. \end{aligned} \quad (6.146)$$

Recollect, from (6.137), that

$$A^2 = \left(\frac{m}{2\pi kT} \right)^3, \quad \alpha = \frac{m}{2kT}. \quad (6.147)$$

We finally arrive at

$$\bar{u} = \left(\frac{m}{2\pi kT} \right)^3 \sqrt{8} \pi^{5/2} \left(\frac{m}{2kT} \right)^{-7/2} = \sqrt{\frac{16kT}{\pi m}} \stackrel{(6.53)}{=} \bar{v} \sqrt{2}. \quad (6.148)$$

¹³ When I first discovered determinants in secondary school, I would practise calculating them by hand, using row reduction and cofactors for matrices as large as 10×10 . This practical knowledge of the relevant manipulations turned out to be invaluable when I studied linear algebra some years later.

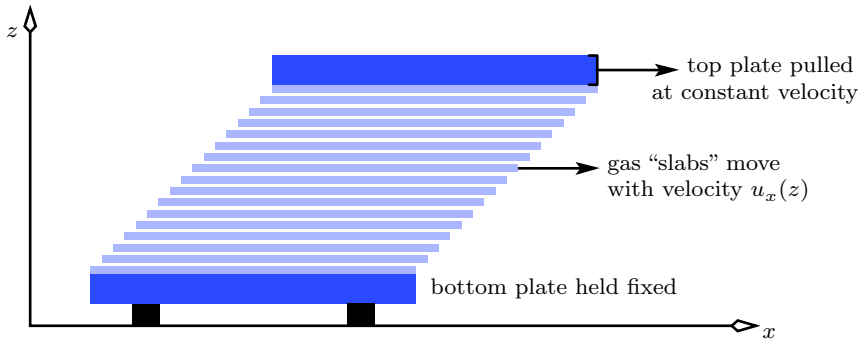


Fig. 6.8 The two-plate scenario for analysing viscosity. Seen here in profile, two solid plates are laid horizontally and separated in the z direction, with gas in between them. The top plate is pulled to the right, and this motion drags the viscous gas to the right

Now recall that $\mathbf{u}, u, \bar{u}$ were shorthand for $\mathbf{v}_{\text{rel}}, v_{\text{rel}}, \bar{v}_{\text{rel}}$, respectively, in this proof. Equation (6.148) then becomes $\bar{v}_{\text{rel}} = \bar{v}\sqrt{2}$. And that is just (6.123), which we set out to prove.

6.8 Viscosity and Mean Free Path

The above discussion of the mean free path of gas particles is an example of a *transport process*. It can be used, among other things, to analyse the microscopic nature of viscosity, thermal conductivity, and heat capacity. We discuss it here for the case of a gas.

A gas's viscosity quantifies the extent to which it resembles treacle. Picture a viscous gas confined between two horizontal solid plates in the xy plane, shown in profile in Figure 6.8. The bottom plate is held fixed. The top plate is dragged horizontally in the $+x$ direction at some fixed velocity, and viscous gas particles are dragged along immediately underneath it. Imagine the gas to be composed of a stack of “gaseous slabs”, each in the xy plane. The slab at a given value of z experiences a force from the slab above that drags it against internal friction along the $+x$ direction with velocity $u_x(z)$. The slab doesn't *accelerate*, because of the ever-present friction that comes with viscosity. As we pull on the top plate, x momentum is transferred by random particle motion down through the slabs, and this drags *them* in turn. The lower the slabs, the lesser the x velocity they inherit.

We wish to relate the force required to drag the top plate to some measure of the gas's viscosity. Relate this force applied to the whole slab to the general mechanical principle of force and momentum:

$$\text{force applied} = \frac{\text{momentum transferred}}{\text{time taken}}. \quad (6.149)$$

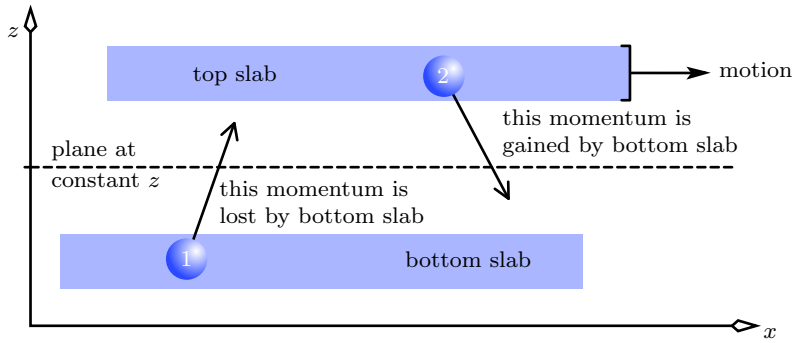


Fig. 6.9 Particle 1 carries its momentum away from the gas below the plane at constant z , while particle 2 carries its momentum into that gas below the constant- z plane

In particular, examine the flow of momentum carried by particles crossing the plane at constant z . Figure 6.9 examines the motion of two particles that cross this plane. Particles crossing up from below have less x momentum than those above this plane. When particle 1 below the plane crosses the plane upward, the gas below the plane loses x momentum equal to particle 1's mass m multiplied by its x velocity. Its x velocity has remained the same since the particle's last collision, which happened, on average, one mean free path length λ away, which we will say was at a z -value of approximately $z - \lambda$. Particle 1's momentum is then $mu_x(z - \lambda)$, where the parentheses in this last expression denote a functional dependence, not multiplication. Similarly, when particle 2 crosses the z plane downward from above, it carries a larger momentum of $mu_x(z + \lambda)$ into the gas below the plane.

The resulting continual injection of momentum into the gas below the z plane provides a force that overcomes the viscosity and drags the lower slabs of gas along. The force that must be applied to any slab to give it a constant velocity against internal friction must be proportional to the slab's area, since, if you pull individually on two adjoining tiles in one slab, you'll find that the force applied is the sum of the forces needed for each tile separately.¹⁴ So, focus on a unit area, and define a quantity T^{xz} :

$$T^{xz} \equiv \left[\begin{array}{l} x \text{ component of force needed to drag a unit area of} \\ \text{the slab of gas at constant } z \text{ with constant velocity} \end{array} \right]$$

$$\underline{\underline{(6.149)}} \quad \frac{p_x \text{ transferred to gas below unit-area plane at constant } z}{\text{time taken}}$$

¹⁴ Compare this comment about slab area to the discussion of the resistance of copper wire at the end of Section 1.9.1, where we showed that it makes no sense to define a “resistance per unit area” of, say, a metal, since resistances of adjacent unit areas don't add.

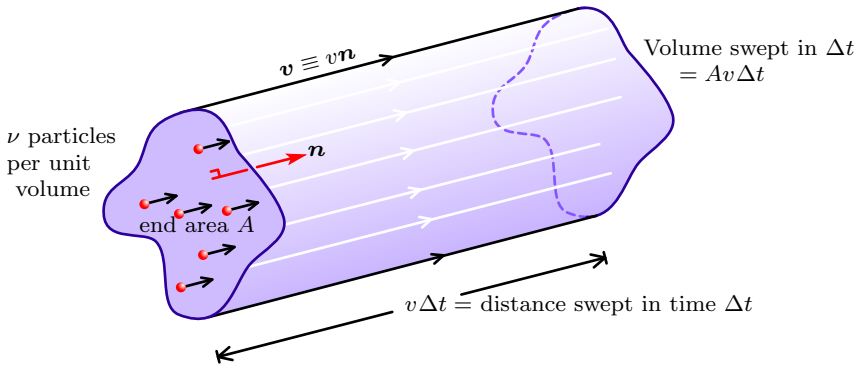


Fig. 6.10 A set of particles passing through, and normal to, a planar area A at speed v will travel a distance $v\Delta t$ in time Δt . Hence, they will sweep out a volume of $Av\Delta t$

$$\begin{aligned}
 &= (p_x \text{ per particle}) \times \left[\begin{array}{l} \text{number of particles transferred} \\ \text{down through unit-area } z \text{ plane,} \\ \text{per unit time} \end{array} \right] \\
 &\quad - (p_x \text{ per particle}) \times \left[\begin{array}{l} \text{number of particles transferred} \\ \text{up through unit-area } z \text{ plane,} \\ \text{per unit time} \end{array} \right]. \quad (6.150)
 \end{aligned}$$

How many particles cross a unit-area z plane per unit time? The amount or number of any quantity crossing a plane in some time interval is found from the *flux density* Φ of that quantity. If the particles have a common velocity $\mathbf{v} = v\mathbf{n}$, where $\mathbf{n}$ is a unit vector, then their flux density is defined as

$$\text{flux density} \equiv \left[\begin{array}{l} \text{the amount or number passing per unit} \\ \text{time through a unit-area plane with normal} \\ \text{vector } \mathbf{v} \text{ (or } \mathbf{n}), \text{ multiplied by } \mathbf{n} \end{array} \right]. \quad (6.151)$$

Figure 6.10 shows a set of such travelling particles. They form a tube whose cross-sectional area at right angles to the particles' velocity is A . The particles passing through one end of this tube in time Δt sweep out a volume of

$$\text{swept volume} = A \times \text{swept distance} = Av\Delta t. \quad (6.152)$$

With ν particles per unit volume, the number that sweep through A in this time is then $\nu Av\Delta t$. The flux density through the tube's end face is then

$$\text{flux density } \Phi \equiv \frac{\text{number through face}}{\text{area} \times \text{time}} \mathbf{n} = \frac{\nu Av\Delta t}{A\Delta t} \mathbf{n} = \nu v \mathbf{n} = \nu \mathbf{v}. \quad (6.153)$$

More generally, how many particles cross some planar area B in time Δt , when those particles have flux density Φ and their velocity is not necessarily

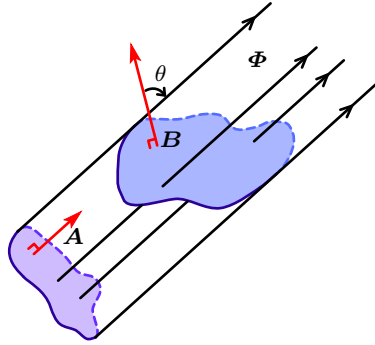


Fig. 6.11 Suppose the particles have flux density Φ . How many cross planar area B in time Δt ?

normal to the plane of area B ? Figure 6.11 shows this more general situation. Equation (6.153) tells us how Φ relates to the area A that *is* perpendicular to the particles' velocity. We also know that $A = B \cos \theta$, where θ is the angle between the normals to the two planes. Represent the area A by a vector $\mathbf{A}$ that has length A and is normal to A 's plane. Similarly, represent area B by a vector $\mathbf{B}$, and write $\Phi \equiv |\Phi|$. Then,

$$\begin{aligned} \left[\begin{array}{c} \text{number of particles} \\ \text{through } B \text{ in } \Delta t \end{array} \right] &= \left[\begin{array}{c} \text{number of particles} \\ \text{through } A \text{ in } \Delta t \end{array} \right] \\ \stackrel{(6.153)}{=} \Phi A \Delta t &= \Phi B \cos \theta \Delta t = \Phi \cdot \mathbf{B} \Delta t. \end{aligned} \quad (6.154)$$

Thus, the number of particles passing through *any* plane can be found using the flux density Φ .

Flux and Current

As used here, “flux” is synonymous with *current*. Flux can refer to the motion of anything, such as particles, mass, or electric charge. For the flow of a substance, (6.153) equates to the expression

$$\underbrace{\text{flux density}}_{\text{an areal density}} = \underbrace{\text{substance density}}_{\text{a volume density}} \times \text{substance velocity}. \quad (6.155)$$

Note that neither use of “density” in (6.155) refers to a unit time. The first “density” refers to the unit *area*: namely, flux density is flux per unit area, where flux means the amount of substance that flows per unit time. So, flux equals flux density times an area (not times a time!). The second “density” in (6.155) refers to the amount of substance per unit *volume*.

The idea that “flux = flux density $\times$ area” agrees with the standard use of flux in electromagnetism. There, for its use in Gauss’s law, flux is

defined as

$$\text{flux} \equiv \text{field strength} \times \text{area}. \quad (6.156)$$

Flux density is thus synonymous with field strength. But outside that subject, you will often find flux density simply called flux—which thus conflicts with electromagnetism. We will always distinguish between “flux density” and “flux”.

Suppose that our gas has ν particles per unit volume. Consider that $\nu/3$ of these have some motion in the z direction, with half of those going up with some representative speed taken from the Maxwell distribution (we’ll call it V for now), and the other half going down with the same speed. Then, referring to (6.153),

$$\left[\begin{array}{c} \text{number of particles transferred one way through} \\ \text{unit-area plane at constant } z, \text{ per unit time} \end{array} \right] \approx \nu/6 \times V. \quad (6.157)$$

Equation (6.150) now says the following. [Remember that the parentheses in the first two lines below of (6.158) denote functional dependence, not multiplication: $u_x(z \pm \lambda)$ denotes the value of u_x at $z \pm \lambda$.]

$$\begin{aligned} T^{xz} &= \underbrace{mu_x(z + \lambda) \times \nu V/6}_{\text{momentum added to lower slab}} - \underbrace{mu_x(z - \lambda) \times \nu V/6}_{\text{momentum lost from lower slab}} \\ &\simeq \frac{\nu V m}{6} [u_x(z) + u'_x(z)\lambda - u_x(z) + u'_x(z)\lambda] \quad (\text{a Taylor expansion}) \\ &= \frac{\nu V m \lambda}{3} \frac{\partial u_x}{\partial z} \equiv \eta \frac{\partial u_x}{\partial z}, \end{aligned} \quad (6.158)$$

where $\eta \equiv \nu V m \lambda/3$ is the gas’s *coefficient of viscosity*, and where we have written a partial derivative to show that u_x generally depends on y as well.

The value chosen for V is often the mean speed $\bar{v}$ in the Maxwell distribution, (6.53). But since we are discussing particles moving up and down, it might make better sense to choose the mean z component of their velocity—or, for the sake of simplicity, the rms value of that z component, (6.111). Compare these choices:

$$\bar{v} = \sqrt{\frac{8kT}{\pi m}}, \quad v_{z,\text{rms}} = \sqrt{\frac{kT}{m}}. \quad (6.159)$$

Their ratio is about 1.6, which is probably comparable to the impreciseness of the above discussion: for example, the distance λ above and below the plane at constant z is an extreme value, and it would be better to use some fraction f of this ($0 < f < 1$), turning λ into $f\lambda$ in the coefficient of viscosity in (6.158).

Likewise, we might set $V = g\sqrt{kT/m}$, where g is somewhat greater than 1; that is, $fg \approx 1$. Writing η in terms of these,

$$\eta \simeq \frac{\nu}{3} g \sqrt{\frac{kT}{m}} m f \lambda \stackrel{(6.124)}{=} \frac{fg}{3\sqrt{2}} \frac{\sqrt{kTm}}{\sigma} = \frac{fg}{3\sqrt{2}} \frac{\sqrt{RTM_{\text{mol}}}}{N_A \sigma}. \quad (6.160)$$

The gas's particle density ν has cancelled out; the surprising result is that viscosity is independent of this density (at a given temperature). This result was indeed derived, and also confirmed experimentally, by Maxwell.

A gas of close-packed particles, each with radius r , has an approximate mass density of

$$\text{mass density} \approx \frac{m}{r^3} = \frac{M_{\text{mol}}}{N_A r^3}. \quad (6.161)$$

This expression then relates the gas's mass density to its collision cross section $\sigma \simeq 4\pi r^2$ via their common factor r . Equations (6.160) and (6.161) enabled a measurement of a coefficient of viscosity η to provide the first values of r and N_A for Loschmidt, in 1885.

Equation (6.160) correctly predicts that the viscosity of a gas increases with temperature. In contrast, the viscosity of *liquids* actually decreases with temperature. For liquids, we must add something to the model: for example, their particles are so close together that particles in adjacent slabs must be modelled as interacting with each other, even though they are not crossing from one slab to the other.

6.9 Thermal Conductivity and Mean Free Path

In the previous section, we took a microscopic view of the transfer of momentum through successive “slabs” of a gas. We wish now to implement the same idea, but calculating the *energy* transfer in place of the momentum transfer. It might come as no surprise for you to find that this forms a good classical model of heat flow.

Begin by recalling Section 4.1.2, where we studied thermal conduction from a macroscopic viewpoint. Central to that discussion was the current (or flux) density $\mathbf{J}$ of the flow of energy that manifests as heat. We will model a medium as the set of slabs drawn in Figure 6.8, but we now focus on the flow of energy through those slabs. Figure 6.12 shows a side view of this new medium, again with a z axis pointing up. This is essentially a one-dimensional scenario: every xy plane has its own temperature $T(z)$. The current of energy flowing in the $+z$ direction across a unit area in the plane at height z is the z component of (4.58). Using an ordinary derivative for our one-dimensional scenario, this is

$$J_z = -\kappa \frac{dT}{dz}, \quad (6.162)$$

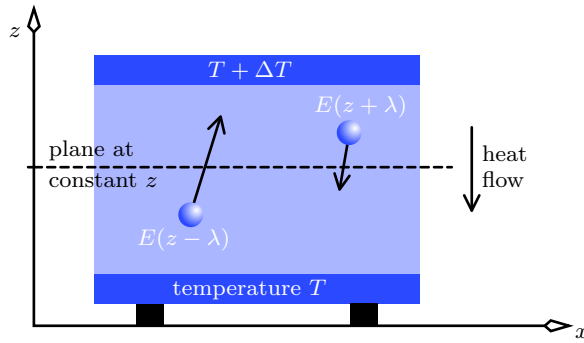


Fig. 6.12 A medium similar to that in Figures 6.8 and 6.9, again viewed from the side. Its top and bottom plates are held at different temperatures without moving them, and we wish to study the resulting heat flow between these plates

where κ is the thermal conductivity of the gas. We wish now to study J_z using the mean free path approach of Section 6.8. The particles at height z each have energy $E(z)$. Remembering that J_z measures the flow of energy toward $+z$, or *upward* in Figure 6.12, we have

$$\begin{aligned}
 J_z &= \frac{\text{energy transferred to gas above unit-area plane at constant } z}{\text{time taken}} \\
 &= (\text{energy per particle}) \times \left[\begin{array}{l} \text{number of particles transferred} \\ \text{up through unit-area } z \text{ plane,} \\ \text{per unit time} \end{array} \right] \\
 &\quad - (\text{energy per particle}) \times \left[\begin{array}{l} \text{number of particles transferred} \\ \text{down through unit-area } z \text{ plane,} \\ \text{per unit time} \end{array} \right].
 \end{aligned} \tag{6.163}$$

Equation (6.157) from Section 6.8 also applies here: the number of particles transferred per unit time one way through a unit-area z plane is approximately $\nu V/6$. The energy per particle that is moving up through the plane at z is $E(z - \lambda)$ [remember that the parentheses here mark functional dependence, not multiplication], and the energy per particle that is moving down through the plane at z is $E(z + \lambda)$. Equation (6.163) becomes

$$\begin{aligned}
 J_z &= \underbrace{E(z - \lambda) \times \nu V/6}_{\text{energy going up}} - \underbrace{E(z + \lambda) \times \nu V/6}_{\text{energy going down}} \\
 &\simeq \frac{\nu V}{6} \left[\cancel{E(z)} - E'(z)\lambda - \cancel{E(z)} - E'(z)\lambda \right] \quad (\text{a Taylor expansion}) \\
 &= \frac{-\nu V \lambda}{3} E'(z).
 \end{aligned} \tag{6.164}$$

Now equate (6.162) with (6.164):

$$-\kappa \frac{dT}{dz} = \frac{-\nu V \lambda}{3} \frac{dE}{dz}. \quad (6.165)$$

The thermal conductivity is then

$$\kappa = \frac{\nu V \lambda}{3} \frac{dE}{dT}. \quad (6.166)$$

At this point, recall the discussion of heat capacity in Section 4.1—but note: in that section, E was the energy of a number of particles, whereas here, E is the energy of a *single* particle. Equation (4.18) says

$$C_V^{\text{mol}} = \frac{d(\text{energy of one mole})}{dT}. \quad (6.167)$$

For our present use of E as the energy of one particle,

$$C_V^{\text{mol}} = \frac{d(N_A E)}{dT} = N_A \frac{dE}{dT} \stackrel{(6.166)}{=} N_A \frac{3\kappa}{\nu V \lambda}. \quad (6.168)$$

This rearranges to give the thermal conductivity κ , which we compare with the coefficient of viscosity η :

$$\kappa = \frac{\nu V \lambda}{3} \frac{C_V^{\text{mol}}}{N_A}; \quad \eta \stackrel{(6.158)}{=} \frac{\nu V m \lambda}{3}. \quad (6.169)$$

Clearly, the quantity κ/η doesn't depend on the particles' number density ν , their characteristic speed V , and their mean free path length λ :

$$\frac{\kappa}{\eta} = \frac{C_V^{\text{mol}}}{N_A m} = \frac{C_V^{\text{mol}}}{M_{\text{mol}}} \stackrel{(4.16)}{=} C_V^{\text{sp}}. \quad (6.170)$$

Equation (6.170) is the final fruit of the above analyses: it uses an atomic view of matter to relate thermal conductivity κ , viscosity η , and specific heat capacity C_V^{sp} . Experiments yield values of

$$\frac{\kappa}{\eta} \approx (1.5 \text{ to } 2.5) \times C_V^{\text{sp}}. \quad (6.171)$$

We can easily expect to be out by a factor of “1.5 to 2.5” in our calculations, because all of the foregoing arguments are based on heuristic models with a heavy reliance on averaging. Even so, the approximate agreement of theory and experiment forms a good justification for the validity of the kinetic/atomic models that we have been using. But although the above calculations agree well with experiments for gases, they fail (in particular) for metals. We'll use a quantum treatment for metals in Chapter 8.

6.10 Excursus: The Energy–Momentum Tensor

The two current (or flux) densities, T^{xz} in (6.158) and J_z in (6.163), belong to a set of similar quantities that can be used to construct an object that has a particular type of physical reality: it is independent of the frame in which it is described or quantified. Such an object is called a *tensor*. Any object that can be regarded as, in some sense, “real” has fundamental importance in physics, from a philosophical as well as a mathematical point of view.

The simplest tensor is a *scalar*, being a single number that all frames agree on, such as the temperature of any point in a room. This temperature’s value is not influenced by whether you are moving past the room or viewing the room upside down. Contrast this with your age, which is *not* a scalar, because not all frames agree on its value: the special theory of relativity says that anyone moving relative to you will state a number for your age that differs from the number that you say it is.¹⁵

The next simplest tensor is a *vector*: although different frames give the components of a vector generally different values, the vector itself is an arrow whose existence is independent of frame. For example, two gymnasts will agree that a javelin is quite real, even though one gymnast is hanging upside-down, and thus sees the javelin in a different orientation to that seen by the other gymnast.¹⁶ A vector thus has a physical reality, which is what makes it a tensor. Note that an ordered set of three numbers such as $[3, -5, 12]$ is *not* a vector; rather, it *might* be an ordered set of *components* of a vector (an arrow) in some frame using some coordinate choice; or it might just be three numbers that have no relation to any arrow, because no coordinate system has been specified. If these numbers are indeed the components of a vector, then clearly, we can index them as the first, second, and third elements of the vector. That procedure uses one index that counts those components using, say, the symbols “1, 2, 3” or perhaps “ x, y, z ”. These choices of symbols used to number the elements are immaterial; what counts is that only *one* index is needed to enumerate a vector’s elements. Because of that, a vector is called a “rank-1 tensor”. A scalar is a rank-0 tensor, because it needs no index to keep track of its single element.

The next-higher level of complexity is a tensor of rank 2: this has two indices to enumerate its elements. Along with scalars and vectors, rank-2 tensors encode physical laws in mathematically simple ways. Following the idea that a rank-1 tensor (a vector) can be coordinatised as a 1-dimensional array of elements, a rank-2 tensor can be coordinatised as a square 2-dimensional tableau of elements: a square matrix, which can typically have size 4×4 . This matrix could, of course, be unravelled into a 1-dimensional array of 16 num-

¹⁵ In fact, the word “scalar” is often used loosely to mean simply a single number, without any mention of whether its value depends on the choice of frame.

¹⁶ I use a javelin here in its visual sense of being an arrow that connects two points. But the important quantity is the *line* connecting any “point 1” to any “point 2”.

bers (and, in fact, this is the way computers tend to store matrices); but doing that does not produce a vector!

It's important to be aware that, just as an array of numbers such as $[3, -5, 12]$ is not a vector as such—and, at best, just represents a vector in some coordinate system, likewise, a matrix of numbers is not a rank-2 tensor as such; at best, it just represents such a tensor in some coordinate system. We first encountered this idea back in Section 1.9.3.

The current or flux densities T^{xz} and J_z turn out to be elements of a rank-2 tensor called the *energy–momentum tensor*. This tensor can be coordinatised relative to a time axis and three space axes, and the tensor can then be written as a 4×4 matrix. The relevant coordinates can be the cartesian set $\{t, x, y, z\}$.¹⁷ With this choice, J_z is called T^{tz} , and the full set of the tensor's cartesian elements consists of $T^{tt}, T^{tx}, \dots, T^{zz}$. In short:

- when the first index is time, it refers to “energy per particle”,
- when the first index is space, it refers to “momentum per particle”,
- when the second index is time, it refers to “particles’ volumetric density”,
- when the second index is space, it refers to “flux density through a plane”.

The space indices describe momentum and flow: because they are vectors, momentum and flow each need three spatial indices to coordinatise them. The energy–momentum tensor's elements are defined fully below. We write “ x momentum” to denote all momentum directed toward increasing x , and “ x plane” for the plane at constant x (the plane whose normal is the x axis). Remember that the phrase “passing through the y plane” means “passing through the y plane in the direction of positive y ”.

Elements of the Energy–Momentum Tensor

$$\begin{aligned} T^{tt} &\equiv [\text{energy per particle}] \times [\text{number of particles per unit volume}] \\ &= \text{energy density (a density in volume, not area).} \end{aligned}$$

$$\begin{aligned} T^{xt} &\equiv [x \text{ momentum per particle}] \times \left[\begin{array}{c} \text{number of particles} \\ \text{per unit volume} \end{array} \right] \\ &= x \text{ momentum density (a density in volume, not area).} \end{aligned}$$

$$\begin{aligned} T^{tx} &\equiv [\text{energy per particle}] \times \left[\begin{array}{c} \text{number of particles passing through} \\ x \text{ plane per unit area per unit time} \end{array} \right] \\ &= \left[\begin{array}{c} \text{energy flux density through } x \text{ plane} \\ \text{(a density in area, not volume)} \end{array} \right]. \end{aligned}$$

¹⁷ We could use other space coordinates, such as spherical polar. But the cartesian choice of x, y, z allows for a simple description of the resulting matrix, shown in (6.173).

$$\begin{aligned}
T^{xy} &\equiv [x \text{ momentum per particle}] \times \left[\begin{array}{c} \text{number of particles passing} \\ \text{through } y \text{ plane per unit area} \\ \text{per unit time} \end{array} \right] \\
&= \left[\begin{array}{c} \text{total } x \text{ momentum passing through } y \text{ plane} \\ \text{per unit area per unit time} \end{array} \right] \\
&= \left[\begin{array}{c} x\text{-momentum flux density through } y \text{ plane} \\ \text{(a density in area, not volume)} \end{array} \right]. \tag{6.172}
\end{aligned}$$

The matrix holding the cartesian components of the energy–momentum tensor can be written in the following block form. The “ $m \times n$ ” expression above each block is the size of the matrix comprising that block. In the bottom-right sub-matrix, the i and j refer to that sub-matrix’s ij^{th} element:

$$\left[\begin{array}{cc} \begin{array}{c} 1 \times 1 \\ \text{energy per unit volume} \end{array} & \begin{array}{c} 1 \times 3 \\ \left[\begin{array}{c} \text{energy passing through} \\ \{x, y, z\} \text{ planes per unit} \\ \text{area per unit time} \end{array} \right] \end{array} \\ \begin{array}{c} 3 \times 1 \\ \left[\begin{array}{c} \left\{ \begin{array}{c} x \\ y \\ z \end{array} \right\} \text{ momentum per} \\ \text{unit volume} \end{array} \right] \end{array} & \begin{array}{c} 3 \times 3 \\ \left[\begin{array}{c} \text{total } i \text{ momentum passing} \\ \text{through } j \text{ plane per unit} \\ \text{area per unit time} \end{array} \right] \end{array} \end{array} \right]. \tag{6.173}$$

Take note that (6.173) is not the energy–momentum tensor; it is the matrix holding the cartesian coordinates of the energy–momentum tensor. (Remember that a matrix is not a tensor; it simply holds the elements of the tensor over some basis.) Examining this matrix shows that its rows involve energy and momentum:

$$\left[\begin{array}{cc} \text{—— energy —— energy ——} \\ \text{—— momentum —— momentum ——} \end{array} \right]. \tag{6.174}$$

Its columns involve the two kinds of density, volumetric and areal:

$$\left[\begin{array}{cc} \begin{array}{c} | \\ \text{per unit volume} \\ | \end{array} & \begin{array}{c} | \\ \text{per unit area per unit time} \\ | \end{array} \\ \begin{array}{c} | \\ \text{per unit volume} \\ | \end{array} & \begin{array}{c} | \\ \text{per unit area per unit time} \\ | \end{array} \end{array} \right]. \tag{6.175}$$

In an inertial frame, an ideal gas with energy density ρ and pressure P (which, of course, acts equally in all directions) has an energy–momentum

tensor whose cartesian coordinates have the following simple form:

$$T^{tt} = \varrho, \quad T^{xx} = T^{yy} = T^{zz} = P, \quad \text{all other coordinates} = 0. \quad (6.176)$$

The dimensions of the elements of the energy–momentum tensor can be related using the dimensions of speed. If we let c stand for an arbitrary speed and let $[A]$ denote the dimensions of A , then it’s easy to see that

$$[T^{tt}] = [cT^{xt}] = [T^{tx}/c] = [T^{xy}] = [\text{pressure}] = \left[\frac{\text{energy}}{\text{volume}} \right]. \quad (6.177)$$

In the special theory of relativity, speed tends to be expressed as a (dimensionless) fraction of the speed of light as measured in a vacuum in an inertial frame: this speed of light is 299,792,458 m/s. Expressing speed relative to the vacuum-inertial speed of light results in the speed of light appearing in all manner of quantities, including the energy–momentum tensor. It follows from (6.177) that all elements of this tensor will then have the same dimensions: of pressure, or, equivalently, of energy density in space. The matrix of these elements then turns out to be symmetric. Thus, for example, $T^{xt} = T^{tx}$, and so the x momentum density (a density in volume, not area) equals the energy flux density through the x plane (a density in area, not volume). Likewise, $T^{xy} = T^{yx}$: the x -momentum flux density through the y plane equals the y -momentum flux density through the x plane (where both of these densities relate to area, not volume).

The central idea of special relativity is that time and space are intermixed, and so what one observer calls time, another observer calls a mixture of time and space, so to speak. Energy turns out to “mix” with momentum in precisely the same way that time “mixes” with space. It follows that what one observer calls energy, another calls a mixture of energy and momentum.¹⁸

But special relativity also tells us that mass possesses (or, in some sense, equates to) energy.¹⁹ This means that if we wish, as Einstein did, to build a relativistic theory of gravity, we might start with Newton’s idea of mass being the source of gravity, and then quickly realise that mass equates to energy—but energy is inextricably linked with momentum in relativity. This suggests that the energy–momentum tensor might act as the source of gravity, or at least play a fundamental role in any relativistic theory of gravity. And that is just how Einstein crafted his *general theory of relativity*, his relativistic theory of gravity. The curvature of spacetime turns out to be a tensor, and Einstein postulated that a certain function of this curvature tensor, called the *Einstein*

¹⁸ That is, in special relativity, energy and momentum pair up to transform between frames with a Lorentz transform. The electrostatic potential Φ and the magnetic vector potential $\mathbf{A}$ also pair up to obey the Lorentz transform. The same can be said for a light wave’s angular frequency ω and its wave vector $\mathbf{k}$.

¹⁹ That sentence must remain imprecise, because physics has not yet determined what mass and energy are—or even whether they *can* be determined at all.

tensor, is proportional to the energy–momentum tensor that describes how energy and momentum flow within that spacetime:

$$\underbrace{\text{Einstein tensor}}_{\text{describes curvature of spacetime}} \propto \underbrace{\text{energy–momentum tensor}}_{\text{describes energy and momentum flow}}. \quad (6.178)$$

This makes for a somewhat esoteric theory. Spacetime curvature determines how objects move: it replaces Newton’s gravitational field in Einstein’s theory. In general, to calculate this curvature, we must know how energy and momentum flow to be able to build the energy–momentum tensor. But to know how energy and momentum flow, we must know the curvature! This difficulty is one reason for why Einstein’s governing equation of general relativity, (6.178), has been solved only for relatively simple models of spacetime.

One Solution of Einstein’s Equation (6.178)

One simple model for which Einstein’s equation (6.178) has been solved exactly is a universe that is empty, save for a single point mass that has remained unchanged forever into the past and the future. In this case, the solution of Einstein’s equation for spacetime’s curvature is called *Schwarzschild spacetime*.

Schwarzschild spacetime is curved in such a way that within what might be called a “radial distance” $2GM/c^2$ of the point mass (where G is Newton’s gravitational constant, M is the mass of the point, and c is the vacuum-inertial speed of light), time and space swap roles, and our intuition of physical laws breaks down. This region of Schwarzschild spacetime is called a *Schwarzschild black hole*. Its bizarre properties involving the roles of time and space all result from the extremely nonphysical simplicity of a universe that is empty save for a single point mass. In close analogy, similar strange properties arise in special relativity for a “uniformly accelerated observer”, one who has accelerated forever in the past and into the future. (Note that the standard representation of Schwarzschild spacetime, which uses *Kruskal–Szekeres coordinates*, resembles the standard representation of a uniformly accelerated frame very closely. We might expect as much by recalling Einstein’s “equivalence principle”; but that is another story.) If this observer’s acceleration is tweaked to become even the slightest bit realistic, by being, say, reduced to zero even a million years into the past and future, the bizarre relativistic consequences of his eternal acceleration go away. The moral of this story is that we should not take the strange properties of Schwarzschild black holes too seriously.

In fact, it can be shown that the spacetime *external* to any spherically symmetric *non-point* mass will also be described by the Schwarzschild solution. (This is called *Birkhoff’s theorem*, and is the general relativistic

analogy of the familiar case in newtonian physics, in which the gravitational field of a spherically symmetric non-point mass, external to the mass, is identical to the gravitational field of a *point* with the same mass.) For example, to a high level of accuracy, the Schwarzschild solution describes the spacetime external to (but not too far from) Earth's surface, and it is used in calculations performed by the ubiquitous satellite receivers that determine the locations of many vehicles on Earth today.

It's clear, then, that when a test object is measured to move in such a way that it indicates spacetime's curvature is given by the Schwarzschild solution, we must not infer that a Schwarzschild black hole is present nearby. Nonetheless, some astronomers do exactly that, when they study the motions of stars orbiting the centre of our galaxy: because these motions are apparently consistent with the Schwarzschild solution, astronomers are apt to infer the presence of a black hole at the galactic centre. But no such inference can logically be made. The best that can be said is that a super-massive object might exist at the galactic centre; but our current physical theories break down in such an extreme domain, and we cannot logically infer the presence there of anything as exotic—and yet naïvely simplistic—as a black hole.



Chapter 7

Introductory Quantum Statistics

In which we describe Einstein's and Debye's theories of heat capacity in a crystal, study the extent of a system's quantum nature, describe the two types of fundamental particle found in Nature, examine liquid helium, and count particle configurations.

In past chapters, we concentrated mostly on the classical idea that particles are tiny balls that interact according to the rules of classical mechanics. This view was fairly successful up until the start of the twentieth century, but not completely. In Section 5.6, we saw that a classical view could not explain the specific heat's stepped dependence on temperature, shown in Figure 5.6. Only the advent of quantum mechanics enabled such a discretisation of phenomena to be explained with any predictive power. But even that required “only” the concept of energy quantisation, which, for example, could allow a molecule seemingly to switch on a vibrational mode when the temperature was increased to some sufficient value.

Eventually, to make sense of newer and more subtle experimental results, it became necessary to build more advanced quantum ideas into statistical mechanics. In particular, the notion that two different types of fundamental particle appear in Nature is now well established in physics, given its success in many diverse experimental areas of physics and chemistry. We focus on these two types of fundamental particle in the remaining chapters. To introduce such concepts, we begin with Einstein's model of heat capacity, and show how its relative success produced a new way of imagining a crystal as being “occupied”, in some sense, by a gas of massless particles that obeyed the rules of the then-new subject of quantum mechanics. This gas altered the classical value of the crystal's heat capacity in a way that gave close agreement with experimental data.

7.1 Einstein's Model of Heat Capacity

In Section 5.6.2, we encountered the fact that at low temperatures, vibrational motion is effectively locked out of simple molecules. This locking out manifests in a more complex way in crystals. Recall that in Section 4.1, we derived the Dulong–Petit law, which says that the molar heat capacity of a crystal is

$C_V^{\text{mol}} = 3R$, where R is the gas constant ($8.314 \text{ JK}^{-1}\text{mol}^{-1}$). This law was originally found empirically, and the value of $3R$ agrees well with experiment at laboratory temperatures and above. But further experiments show that as its temperature is lowered, a crystal's heat capacity no longer remains constant, and instead reduces to zero. It seems that describing the crystal classically as a set of atoms that each have six quadratic energy terms due to their vibrating in three spatial dimensions is too simplistic. Just as we saw in Section 5.6.2 when discussing vibration of diatomic gas molecules, it seems that a quantum-mechanical view of vibration is necessary to explain the finer details of the heat capacity's temperature dependence in crystals.

Around 1907, Einstein put the new quantum concepts of Planck to use in a model of heat capacity that gave a good fit to C_V^{mol} data at low temperatures, where the Dulong–Petit law failed completely. Einstein's model can be described in modern quantum-mechanical language in the following way. He began by assuming that the crystal is composed of atomic oscillators, for which

- the oscillator energies are quantised,
- the oscillators all have the same frequency of vibration, and
- this frequency is the same along all three spatial dimensions.

Recall (4.7)'s expression for heat capacity at constant volume, which refers to a system's *total* internal energy E . With n moles of atoms present, equation (4.15) says

$$C_V^{\text{mol}} \equiv \frac{C_V}{n} \stackrel{(4.7)}{=} \frac{1}{n} \left(\frac{\partial E}{\partial T} \right)_{V,N} = \frac{N_A}{N} \left(\frac{\partial E}{\partial T} \right)_{V,N} = N_A \left(\frac{\partial \bar{E}}{\partial T} \right)_{V,N}, \quad (7.1)$$

where $\bar{E}$ is the average energy of *one* oscillator. We wish to calculate the crystal's molar heat capacity subject to the above assumptions, and so we work with the canonical ensemble: the crystal interacts with a heat bath, with no volume or particles exchanged with that bath. Work in one dimension for simplicity, knowing that the mean energy in three dimensions will be triple the one-dimensional value (denoted “1D”):

$$\bar{E} = 3\bar{E}_{1D}. \quad (7.2)$$

Recall the energy E_n of energy level n of a one-dimensional quantised oscillator, which we saw in (5.74):

$$E_n = (n + 1/2) hf, \quad \text{where } n = 0, 1, 2, \dots \quad (7.3)$$

A one-dimensional quantised oscillator's mean energy $\bar{E}_{1D}$ is a weighted sum over these energy levels:

$$\bar{E}_{1D} = \sum_{n=0}^{\infty} p_n E_n, \quad (7.4)$$

where p_n is the chance that the oscillator is found in energy level n . Each level contains just one state of vibration, so $\Omega_n = 1$ for all n . The probability p_n is then given by (5.32) with $\Omega_n = 1$. [Or, equivalently, p_n is given by (5.37) and (5.39), because, in this case, a level is the same as a state.]

We saw, in (5.106), that differentiating the partition function Z with respect to $\beta = 1/(kT)$ gives the average energy $\bar{E}_{1D}$ (in one dimension, since the energy levels in (7.3) are for one dimension of oscillation):

$$\bar{E}_{1D} = \frac{-1}{Z} \frac{\partial Z}{\partial \beta}, \quad \text{where } Z \stackrel{(5.39)}{=} \sum_n e^{-\beta E_n}. \quad (7.5)$$

We must calculate Z . Partly for convenience, we'll write the zero-point energy $1/2 hf$ in (7.3) as some fixed baseline energy ε whose value doesn't concern us; but also, doing so will show clearly that the final expression for heat capacity does not depend on ε . The partition function is then a geometric series with the usual sum of "first term over (1 minus ratio)":

$$Z = \sum_{n=0}^{\infty} e^{-\beta(nhf+\varepsilon)} = \frac{e^{-\beta\varepsilon}}{1 - e^{-\beta hf}}. \quad (7.6)$$

It follows that

$$\begin{aligned} \bar{E}_{1D} &= \frac{-1}{Z} \frac{\partial Z}{\partial \beta} = \frac{-\partial}{\partial \beta} \ln Z = \frac{-\partial}{\partial \beta} [-\beta\varepsilon - \ln(1 - e^{-\beta hf})] \\ &= \varepsilon + \frac{-e^{-\beta hf} \times -hf}{1 - e^{-\beta hf}} = \varepsilon + \frac{hf}{e^{\beta hf} - 1}. \end{aligned} \quad (7.7)$$

Observe that:

- **In the low-temperature limit** ($hf \gg kT$, i.e., $\beta hf \gg 1$), $\bar{E}_{1D} \rightarrow \varepsilon$. So, the average energy of an oscillator is just its zero-point energy ε .
- **In the high-temperature limit** ($hf \ll kT$, or $\beta hf \ll 1$),

$$\bar{E}_{1D} \rightarrow \varepsilon + \frac{hf}{1 + \beta hf - 1} = \varepsilon + kT. \quad (7.8)$$

By definition, the zero-point energy ε cannot be removed from the oscillator, effectively meaning it does not exist as far as thermal interactions are concerned. But even if that were not the case, since $\varepsilon = 1/2 hf \ll kT$, in this high-temperature limit, we see that kT dominates the average energy of a single oscillator. And that is precisely as expected: kT is the classical equipartition result arising from the two quadratic energy terms that this oscillator has (one kinetic and one potential) when high temperatures render its behaviour classical.

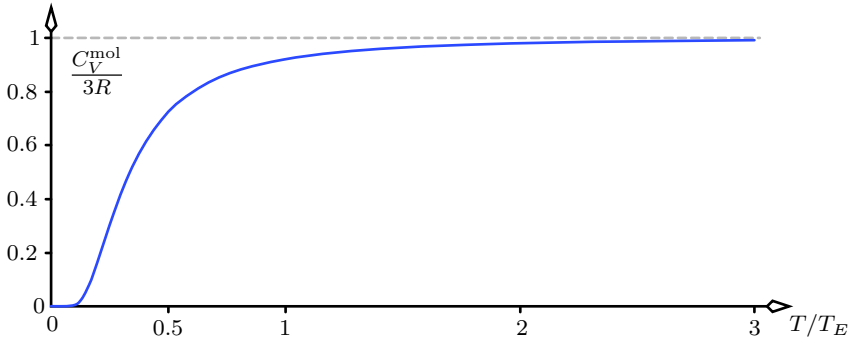


Fig. 7.1 Einstein's prediction of $C_V^{\text{mol}}/(3R)$ versus T/T_E , using (7.12). It is reasonably successful, but experimental values of $C_V^{\text{mol}}/(3R)$ for a variety of substances can depart from this curve by as much as 10%

The mean oscillator energy in the three-dimensional crystal is triple that of the one-dimensional oscillator:

$$\bar{E} = 3\bar{E}_{1D} \stackrel{(7.7)}{=} 3\varepsilon + \frac{3hf}{e^{\beta hf} - 1}. \quad (7.9)$$

Applying (7.1) to this expression results in

$$C_V^{\text{mol}} = N_A \left(\frac{\partial \bar{E}}{\partial T} \right)_{V,N} = 3R \left(\frac{hf}{kT} \right)^2 \frac{e^{\beta hf}}{(e^{\beta hf} - 1)^2}. \quad (7.10)$$

This is Einstein's expression for the molar heat capacity of a crystal. As might be expected, some conciseness is achieved when we introduce the “Einstein temperature” T_E via

$$kT_E \equiv hf. \quad (7.11)$$

This converts (7.10) into a one-parameter form that is convenient for comparing with experimental data:

$$C_V^{\text{mol}} = 3R \left(\frac{T_E}{T} \right)^2 \frac{e^{T_E/T}}{(e^{T_E/T} - 1)^2}. \quad (7.12)$$

This function is plotted in Figure 7.1. Experiments show that a variety of substances follow this curve fairly well, but their values can depart from Einstein's prediction by as much as 10%.

The plot in Figure 7.1 indicates that C_V^{mol} tends toward the Dulong–Petit result of $3R$ in the limit of high temperature ($T \gg T_E$). We can also show this analytically. At these comparatively high temperatures, (7.12) becomes

$$C_V^{\text{mol}} \rightarrow 3R \left(\frac{T_E}{T} \right)^2 \frac{1 + T_E/T}{(1 + T_E/T - 1)^2} \rightarrow 3R, \quad (7.13)$$

as expected.¹

In laboratory use, Einstein's expression (7.12) is fitted to experimental data by choosing a value for T_E . This is equivalent to letting the data tell us the value of the atoms' vibrational frequency f . Values of T_E tend to be around room temperature, corresponding to a vibrational frequency in the terahertz range:

$$f = \frac{kT_E}{h} \simeq \frac{1.381 \overset{-23}{\times} 298}{6.626 \overset{-34}{}} \text{ Hz} \simeq 6 \times 10^{12} \text{ Hz}. \quad (7.14)$$

A good choice of T_E results in (7.12) matching experimental heat-capacity data very well, even down to low temperatures, where C_V^{mol} departs from the Dulong–Petit value of $3R$ and reduces to zero. At the start of the twentieth century, the success of Einstein's explanation of the departure from the Dulong–Petit model gave an important boost to what was then the new theory of quantum mechanics.

But, at very low temperatures, Einstein's prediction (7.12) does *not* agree well with experimental data, since it falls to zero with temperature exponentially quickly, in disagreement with the polynomial behaviour shown in experiments. For crystals with no conduction electrons, the experimental result is $C_V^{\text{mol}} \propto T^3$; and for metals, experiments shows that $C_V^{\text{mol}} \propto T$.

In 1912, Debye refined Einstein's model by switching attention from the individual atomic oscillators to the *normal modes* of vibration of the crystal lattice as a whole, and his model does indeed predict $C_V^{\text{mol}} \propto T^3$ for crystals at low temperature. We'll analyse Debye's model in the next two sections by applying a quantised view of the normal modes. It will turn out that the mathematics of Einstein's and Debye's formulations are almost identical. The two approaches differ in their interpretation of a state, along with their choice of the density of states $g(E)$ that we introduced back in Section 2.5. Chapter 8 will show how the $C_V^{\text{mol}} \propto T$ dependence comes about for metals.

7.2 A Refinement of Einstein's Model of Heat Capacity

We showed, in the last section, that when Einstein analysed a crystal's heat capacity by modelling the crystal as a set of quantised oscillators that obey the Boltzmann distribution, the resulting expression (7.10) gave a much better fit to low-temperature data than did the Dulong–Petit prediction of $C_V^{\text{mol}} = 3R$. To pave the road to Debye's improvement of Einstein's model, we'll rephrase Einstein's model in a way that has become the basis for the modern “phonon” view of crystal heat capacity.

¹ In this high-temperature regime, we can simplify the analysis by referring to (7.8). Then, set $\bar{E} = 3\varepsilon + 3kT$, and now apply (7.1) to arrive again at $C_V^{\text{mol}} = 3R$.

Return to (7.4) for the mean energy of a one-dimensional (“1D”) oscillator:

$$\bar{E}_{1D} = \sum_{n=0}^{\infty} p_n E_n, \quad (7.15)$$

where n labels the energy level of a quantised oscillator: this oscillator is one atom or molecule of the crystal (we’ll just use the word “molecule”), which oscillates in one spatial dimension. These oscillators all have the same fundamental frequency f . Modern quantum mechanics gives an expression for E_n in (7.3) that includes the oscillator’s zero-point energy $1/2 hf$. But, because Einstein’s model preceded the idea of zero-point energy, we’ll generalise (7.3) to become $E_n = nhf + \varepsilon$ [just as we did in (7.6)], where ε is some fixed baseline energy whose value we may not know—and, in fact, do not *need* to know. Equation (7.15) becomes

$$\begin{aligned} \bar{E}_{1D} &= \sum_n p_n (nhf + \varepsilon) = hf \sum_n p_n n + \varepsilon \\ &= hf \bar{n} + \varepsilon, \end{aligned} \quad (7.16)$$

where $\bar{n}$ is the average of the energy levels n occupied by the oscillator (and so depends on temperature). That is, n and $\bar{n}$ are just numbers: n is a whole number and $\bar{n}$ is a real number.

Here, n indexes the energy levels, and $\bar{n}$ is its mean value in a population of possibly excited oscillators. But a different interpretation of n and $\bar{n}$ can be found. Energy level n with its energy $E_n = nhf + \varepsilon$ can be treated as a state of the oscillator in which n “massless non-interacting particles” are present in some “ghostly” quantum-mechanical sense. Each of these quantum particles has energy hf . Also, a background energy ε exists that cannot be removed, and so plays no pivotal role in the discussion.² Equation (7.16) says

$$\begin{aligned} \bar{n} &= \frac{\bar{E}_{1D} - \varepsilon}{hf} = \frac{\text{mean “non-background” energy of a 1D-oscillator}}{\text{quantum particle energy}} \\ &= \text{mean number of quantum particles present per 1D-oscillator,} \\ &\quad \text{and called the crystal’s } \textit{occupation number}. \end{aligned} \quad (7.17)$$

With this interpretation, Einstein’s model says that the three-dimensional crystal comprising N molecular oscillators that each have 3 spatial dimensions of oscillation has, on average, $3N \times \bar{n}$ massless non-interacting particles present. Each of these quantum particles has energy hf .

We can calculate the crystal’s occupation number $\bar{n}$ with a Boltzmann treatment. Start with

² This language harks back to the comment at the end of Section 5.3 that energy levels are often called states.

$$\bar{n} = \sum_{n=0}^{\infty} n p_n = \frac{1}{Z} \sum_n n e^{-\beta(nhf+\varepsilon)}, \quad \text{where } Z = \sum_n e^{-\beta(nhf+\varepsilon)}. \quad (7.18)$$

Write $\alpha \equiv -\beta hf$ for convenience:

$$\bar{n} = \frac{1}{Z} \sum_n n e^{n\alpha - \beta\varepsilon}, \quad Z = \sum_n e^{n\alpha - \beta\varepsilon}. \quad (7.19)$$

This resembles the calculations in Section 7.1. There, we calculated $\partial Z / \partial \beta$ to find a mean energy. That suggests that here, $\partial Z / \partial \alpha$ might give a mean particle number:

$$\frac{\partial Z}{\partial \alpha} \stackrel{(7.19)}{=} \sum_n e^{n\alpha - \beta\varepsilon} \times n \stackrel{(7.19)}{=} \bar{n} Z. \quad (7.20)$$

It follows that

$$\bar{n} = \frac{1}{Z} \frac{\partial Z}{\partial \alpha}. \quad (7.21)$$

But Z is a geometric series, and thus is easily summed, as we saw in (7.6):

$$Z \stackrel{(7.19)}{=} \sum_{n=0}^{\infty} e^{n\alpha - \beta\varepsilon} = \left[\frac{\text{“first term over”}}{\text{“1 minus ratio”}} \right] = \frac{e^{-\beta\varepsilon}}{1 - e^{\alpha}}. \quad (7.22)$$

Now apply (7.21) to the right-hand expression in (7.22), to obtain

$$\bar{n} = \frac{1}{e^{-\alpha} - 1} = \frac{1}{e^{\beta hf} - 1}. \quad (7.23)$$

The ever-present zero-point energy of one oscillator, ε , doesn't appear here.

Remember that we are calculating a heat capacity using, say, (7.1):

$$C_V^{\text{mol}} = N_A \left(\frac{\partial \bar{E}}{\partial T} \right)_{V,N}, \quad (7.24)$$

where $\bar{E}$ is the mean energy per crystal molecule. But we will now redefine $\bar{E}$ to exclude the zero-point energy. Also, recall that the set of N molecular oscillators that each have 3 spatial dimensions of oscillation is being treated as $3N$ 1D-oscillators. Then,

$$\begin{aligned} \bar{E} &= \frac{\text{total energy of oscillation, excluding zero-point energy}}{\text{number of crystal molecules present}} \\ &= \frac{\left[\begin{array}{l} \text{number of 1D-oscillators } (3N) \\ \times \text{ number of massless particles per 1D-oscillator } (\bar{n}) \\ \times \text{ energy per massless particle } (hf) \end{array} \right]}{\text{number of crystal molecules present } (N)} \end{aligned}$$

$$= 3\bar{n}hf \stackrel{(7.23)}{=} \frac{3hf}{e^{\beta hf} - 1}. \quad (7.25)$$

This is the same as (7.9) without the zero-point energy 3ε . The absence of zero-point energy shows that we have switched viewpoints from calculating a mean energy in Section 7.1 to calculating a mean number of massless quantum particles here. Substituting this new form of $\bar{E}$ into (7.24) returns Einstein's expression (7.10) for the molar heat capacity. Thus, we have converted Einstein's model of the crystal's heat capacity into a form that treats the energy of the crystal molecules' oscillations as though it were being carried by a shadowy gas of massless non-interacting quantum particles.

This reformulation of Einstein's model can be repackaged in a way that will allow it to evolve seamlessly into Debye's 1912 theory of specific heat. To see how, recall the comment just before (7.17), where we said that an oscillator's energy level n (with energy $E_n = nhf + \varepsilon$) is treated as a state of the oscillator that is occupied by n quantum particles, each of energy hf . Now alter the viewpoint: treat this state as being associated with a *single* energy hf —think of it as a box with “Energy hf ” written on it. This state is occupied by n quantum particles, and when a quantum particle occupies this state, that particle “is given” energy hf . This is an important change of viewpoint: instead of a state having energy nhf because each of its n particles has a set energy hf , we now picture a state as having energy hf , and the n particles that occupy it each have energy hf by virtue of being in that state.

Next, recall the idea of counting states introduced in Chapter 2, and in particular, the density of states $g(E)$ in Section 2.5, being the number of states per unit energy with some energy E . The total number of states available to the crystal is $\Omega_{\text{tot}} = 3N$. These are all piled onto a single value of energy $E = hf$. It follows that if we use a continuum view of the crystal's internal energy, its density of states spikes at $E = hf$. Its total number of states can be written as

$$\Omega_{\text{tot}} = \int_0^\infty g(E) dE = 3N, \quad (7.26)$$

where $g(E)$ is the density of states. Then, because each of the $3N$ states is associated with an energy hf , the density of states must be a delta function:

$$g(E) = 3N \delta(E - hf). \quad (7.27)$$

This is shown in Figure 7.2.

Now refer to (7.25) to write the total energy of oscillation as (with an explicit subscript “tot” here, since we are now using E as a variable)

$$E_{\text{tot}} = N\bar{E} = 3N\bar{n}hf = \bar{n}hf \int_0^\infty g(E) dE. \quad (7.28)$$

Recall (7.23) and put everything under the integral sign:

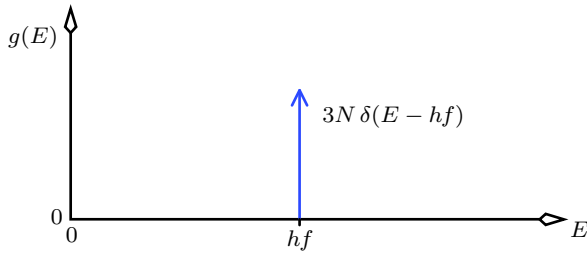


Fig. 7.2 The density of states for Einstein's model of the crystal consists of a single delta function

$$\begin{aligned}
 E_{\text{tot}} &= \int_0^{\infty} hf \times \frac{1}{e^{\beta hf} - 1} \times g(E) dE \\
 &= \int_0^{\infty} hf \times \frac{1}{e^{\beta hf} - 1} \times 3N \delta(E - hf) dE \\
 &= \int_0^{\infty} E \times \frac{1}{e^{\beta E} - 1} \times 3N \delta(E - hf) dE. \quad (7.29)
 \end{aligned}$$

Suppose that we allow the occupation number to be a function of energy along with temperature. (It was always a function of temperature, but we indicate this explicitly now.) Write

$$\bar{n}(E, T) \equiv \frac{1}{e^{\beta E} - 1}. \quad (7.30)$$

Equation (7.29) then becomes

$$E_{\text{tot}} = \int_0^{\infty} E \bar{n}(E, T) g(E) dE. \quad (7.31)$$

The various quantities above tie together in the following way:

$$\underbrace{\frac{\text{number of quantum particles in } E \text{ to } E + dE}{= dN = \frac{dN}{dE} dE}}_{\substack{\text{number of quantum} \\ \text{particles per unit} \\ \text{energy interval}}} = \underbrace{\bar{n}(E, T)}_{\substack{\text{mean number} \\ \text{of quantum} \\ \text{particles per} \\ \text{state}}} \times \underbrace{g(E) dE}_{\substack{\text{number of} \\ \text{states per unit} \\ \text{energy interval}}}. \quad (7.32)$$

On dividing (7.32) by dE , this becomes

$$\underbrace{\frac{dN}{dE}}_{\substack{\text{number of quantum} \\ \text{particles per unit} \\ \text{energy interval}}} = \underbrace{\bar{n}(E, T)}_{\substack{\text{mean number} \\ \text{of quantum} \\ \text{particles per} \\ \text{state}}} \times \underbrace{g(E)}_{\substack{\text{number of} \\ \text{states per unit} \\ \text{energy interval}}}. \quad (7.33)$$

Quantum systems are characterised by their occupation number $\bar{n}(E, T)$ [often written $F(E)$], and their density of states, $g(E)$. Their occupation number $\bar{n}(E, T)$ is simple and well-behaved: we'll calculate its general form for the two different types of quantum particle (*fermions* and *bosons*) in Section 7.6. The density of states $g(E)$ is another case entirely: it varies for different materials and is generally a complicated function of energy E . Its precise form for a given material is usually determined empirically.

Equation (7.32) says that the total number of massless quantum particles present in the crystal is

$$\left[\begin{array}{l} \text{total number of massless} \\ \text{quantum particles in crystal} \end{array} \right] = \int dN = \int_0^\infty \bar{n}(E, T) g(E) dE. \quad (7.34)$$

We'll use this equation in Section 7.7. Also, (7.31) and (7.32) combine as

$$E_{\text{tot}} = \int_0^\infty E dN. \quad (7.35)$$

This last equation makes perfect sense: the total energy of the oscillators is the aggregate of all quantum-particle energies, where the contribution from the dN quantum particles with energy E is $dN \times E$.

Take note of the various symbols above denoting numbers of massless quantum particles:

- n is the energy level of a one-dimensional oscillator, and also the number of quantum particles per state, where each state denotes one dimension of oscillation of one crystal molecule.
- $\bar{n}$, the occupation number, is the arithmetic mean of n : the mean number of quantum particles per state.
- dN/dE is the number of quantum particles per unit energy interval.
- N is the number of quantum particles with energies up to E . It could also be called $N(E)$.

7.3 Debye's Model of Heat Capacity

The above description of Einstein's model that used the language of occupation number $\bar{n}(E, T)$ and density of states $g(E)$ leads naturally to a description of Debye's 1912 model of heat capacity.

Debye's theory builds on the basic elements of Einstein's model by spreading the states out in energy, rather than having them all concentrated at energy hf . So, it begins with (7.31), but replaces Einstein's occupation number (7.23) with (7.30):

$$\bar{n} = \frac{1}{e^{\beta hf} - 1} \quad \text{becomes} \quad \bar{n}(E, T) = \frac{1}{e^{\beta E} - 1}. \quad (7.36)$$

The density of states $g(E)$ is changed by using a new definition of the crystal's states. Einstein's states were individual oscillators, three per crystal molecule. Debye changed the idea of state to encompass the entire crystal. After all, the crystal's molecules form a tightly coupled set of oscillators; they do not oscillate independently of each other. The motion of coupled oscillators can be complex, but this motion can always be decomposed into a linear combination of *normal modes*. A normal mode describes the whole set of oscillators in the special case of motion when the amplitude of each oscillator remains constant in time.³

For example, consider two identical pendula linked to each other with a very light spring, as in Figure 7.3. When set into motion with some random initial condition (and, as usual for such pendula, we assume the oscillations are small), the amplitude of each pendulum will usually change with time, *unless* the system occupies one of its two normal modes. The first mode results when both pendula are raised together and set into motion as one: clearly, they oscillate in phase at their natural frequency, since each is effectively unaware of the other because the spring simply goes along for the ride. The second mode results when the initial amplitudes are also equal but the pendula are released from opposite directions. They now oscillate 180° out of phase ("in antiphase"). They also oscillate at higher than their natural frequency. This is because their connecting spring is now being stretched and compressed periodically, and thus supplies its own restoring force.

Figure 7.4 shows three coupled identical pendula side by side, with each pair of neighbours connected by a light spring (with both springs identical). This set has three normal modes. The first mode results when all three pendula are set to oscillate in phase by supplying each with the same initial conditions. The second mode results when the two end pendula are given equal but opposite initial amplitudes and the middle pendulum is given zero amplitude: here, the end pendula swing in antiphase and the middle one never moves at all. The third mode occurs when the outer pendula are given the same initial amplitude and direction, and the middle pendulum is given twice

³ I use "amplitude" here in the sense that it has throughout physics and throughout this book, whereby the amplitude of the function $y = A \sin \omega t$ equals A . This tallies with its Latin root "amplus", meaning large or abundant. Some engineers and signal processors use "amplitude" to mean the instantaneous value of an oscillating quantity. They will then say $y = A \sin \omega t$ has the time-varying amplitude y . But all oscillating quantities already have specific names such as "displacement" and "voltage"; so, it makes more sense for "amplitude" to denote something other than y .

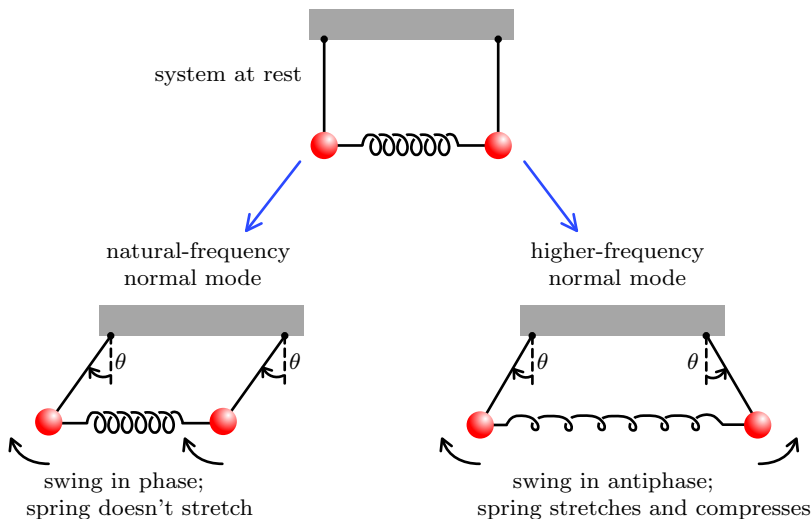


Fig. 7.3 The two normal modes available to a pair of identical pendula connected by a light spring. **Left:** When the pendula are set oscillating in phase with the same amplitude, the spring is effectively absent, and they oscillate together at their natural frequency. **Right:** When the pendula are set oscillating in antiphase with the same amplitude, the spring is alternately stretched and compressed, and the pendula continue to oscillate in antiphase, but at higher than their natural frequency

that amplitude in the opposite direction: now, the middle one always swings in antiphase to the outer pair, with twice their amplitude.

More generally, n coupled pendula give rise to n coupled equations of motion. These can always be separated into n non-coupled equations using linear algebra; what results can be encoded into an $n \times n$ matrix whose determinant must be found. This procedure shows that this set of pendula has n normal modes.

Return now to our crystal. Whereas Einstein treated the N -molecule crystal as having $3N$ states (one state per molecular oscillator per dimension of motion), Debye took each state to be a normal mode of oscillation of the entire crystal when it's treated as a set of $3N$ *coupled* oscillators. Thus, Debye's model allocates $3N$ states to the crystal too; but each of Debye's states corresponds not to one of Einstein's individual oscillators allied to a quantum particle of energy hf , but rather, to a single normal mode. Each of these normal modes is associated with a new quantum particle. These quantum particles have a *spread* of energies as they move back and forth through the crystal. Today, these massless quantum particles are called *phonons*.

We calculated the density of states of such massless particles in Section 2.6. Specifically, for a crystal of volume V , equation (2.115) is

$$g(E) = \frac{12\pi V E^2}{h^3 c^3}, \quad (7.37)$$

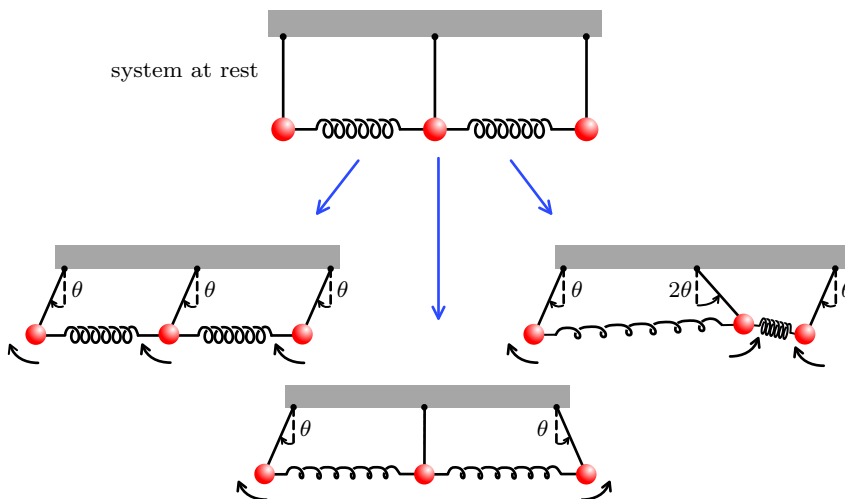


Fig. 7.4 The three normal modes available to three identical pendula connected with light springs. **Left:** Oscillating in phase with the same amplitude. **Middle:** Outer pendula oscillating in antiphase, middle pendulum stationary. **Right:** Outer pendula oscillating in phase, middle pendulum oscillating in antiphase with twice the amplitude of the outer pendula

where c is the appropriate mean of the speeds of the phonons' three available polarisations, as discussed in Section 2.6. We take c to be the mean *speed of sound*, since sound is carried by vibrations propagating through a crystal.

The theory that gave rise to (7.37) allowed for arbitrarily high phonon energies, which correspond to arbitrarily small wavelengths of these quantum particles. But a wavelength of oscillation that is a great deal smaller than an inter-molecular spacing makes no real sense. For example, we might jiggle a string at high frequency to produce waves on it with a small wavelength. But these waves will never be shorter than the spacing of the string's molecules, since there is nothing between the molecules to wave. Similarly, for the crystal, Debye realised that the energy E in (7.37) can be no higher than some maximum value E_D called the crystal's *Debye energy*, which is generally different for each crystal. The density of states for a real crystal is then

$$g(E) = \begin{cases} \frac{12\pi V E^2}{h^3 c^3} & E \leq E_D, \\ 0 & E > E_D. \end{cases} \quad (7.38)$$

On the left in Figure 7.5 is a plot of $\bar{n}(E, T)$ versus E [from (7.30)]. On the right in the figure is a plot of $g(E)$ for the Einstein case [the delta function in (7.27), reproduced from Figure 7.2], along with $g(E)$ for the Debye case [the parabola in (7.38) cut off at E_D].

Debye's approach now writes (7.31) as

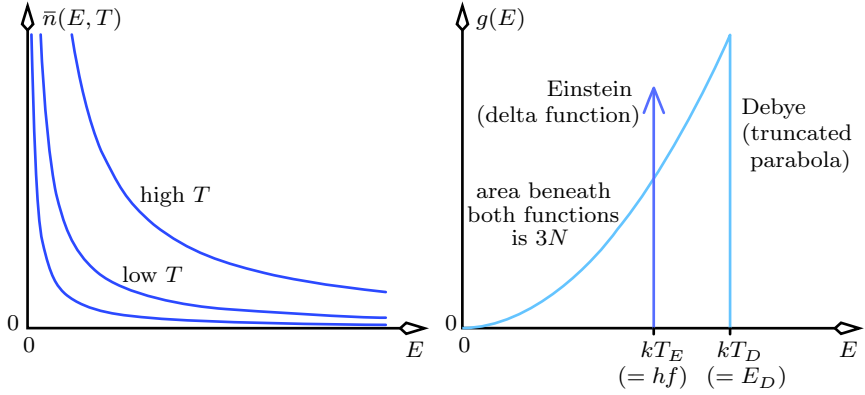


Fig. 7.5 **Left:** $\bar{n}(E, T)$ versus E , from (7.30), for a selection of temperatures. **Right:** $g(E)$ versus E for both Einstein and Debye. Einstein's $g(E)$, from (7.27), is a delta function at $E = hf = kT_E$ [see (7.11)]. Debye's version from (7.38) is a parabola truncated at $E = kT_D$. The Einstein temperature T_E and the Debye temperature T_D are typically similar in value. The area under each choice of $g(E)$ is the total number of states, which is $3N$ for both Einstein and Debye

$$\begin{aligned}
 E_{\text{tot}} &= \int_0^\infty E \bar{n}(E, T) g(E) dE \quad [\text{and now use (7.30) and (7.38)}] \\
 &= \int_0^{E_D} E \times \frac{1}{e^{\beta E} - 1} \times \frac{12\pi V E^2}{h^3 c^3} dE = \frac{12\pi V}{h^3 c^3} \int_0^{E_D} \frac{E^3 dE}{e^{\beta E} - 1}. \quad (7.39)
 \end{aligned}$$

We can simplify this integral by noting that the crystal has $3N$ states:

$$\int_0^\infty g(E) dE = 3N. \quad (7.40)$$

In other words, from (7.38),

$$\frac{12\pi V}{h^3 c^3} \int_0^{E_D} E^2 dE = 3N. \quad (7.41)$$

This easily becomes

$$\frac{12\pi V}{h^3 c^3} = \frac{9N}{E_D^3}. \quad (7.42)$$

This allows (7.39) to be written more conveniently as

$$E_{\text{tot}} = \frac{9N}{E_D^3} \int_0^{E_D} \frac{E^3 dE}{e^{\beta E} - 1}. \quad (7.43)$$

With a change of variables $x \equiv \beta E$, the temperature dependence disappears from the integrand (but not the integral):

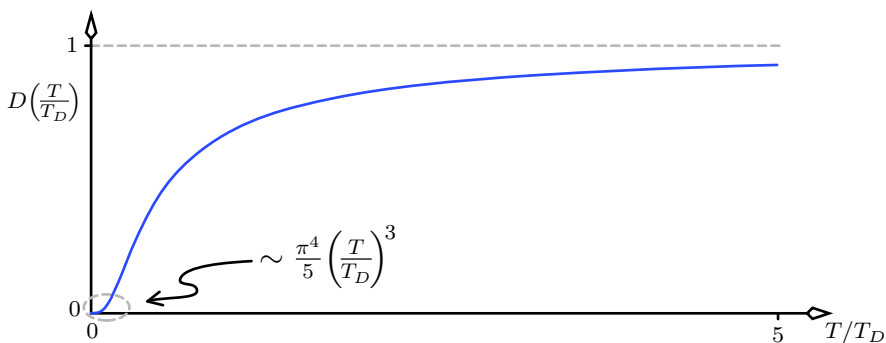


Fig. 7.6 The Debye function $D(T/T_D)$, from (7.47). Its high- and low-temperature limits are given in (7.48) and (7.50), respectively

$$E_{\text{tot}} = \frac{9N}{E_D^3 \beta^4} \int_0^{\beta E_D} \frac{x^3 dx}{e^x - 1}. \quad (7.44)$$

Remember that if we treat $3N$ oscillators classically using the equipartition theorem, they will each have two quadratic energy terms, corresponding to kinetic and potential energies. Each energy term has energy $kT/2$, giving the set a total energy of $E_{\text{tot}} = 3N \times 2 \times kT/2 = 3NkT$. To make a connection with this classical value (which, at high temperatures, leads to the Dulong–Petit law), take it out as a factor in front of (7.44), to produce

$$E_{\text{tot}} = 3NkT \times \frac{3}{(\beta E_D)^3} \int_0^{\beta E_D} \frac{x^3 dx}{e^x - 1}. \quad (7.45)$$

The two-fold appearance of $\beta E_D = E_D/(kT)$ in (7.45) suggests that we define the *Debye temperature* T_D via

$$E_D = kT_D. \quad (7.46)$$

Then, $\beta E_D = T_D/T$. This converts (7.45) into

$$E_{\text{tot}} = 3NkT \times 3 \left(\frac{T}{T_D} \right)^3 \underbrace{\int_0^{T_D/T} \frac{x^3 dx}{e^x - 1}}_{\equiv \text{“Debye function” } D(T/T_D)}. \quad (7.47)$$

Equation (7.47) defines the *Debye function*, shown in Figure 7.6. In general, this function must be calculated numerically, but we can treat the high- and low-temperature regimes analytically:

- **In the high-temperature limit**, $T_D/T \ll 1$. Hence, x in the integrand of (7.47) is always much less than 1. Thus,

$$\begin{aligned}
D(T/T_D) &\simeq 3 \left(\frac{T}{T_D} \right)^3 \int_0^{T_D/T} \frac{x^3 dx}{1+x-1} \\
&= 3 \left(\frac{T}{T_D} \right)^3 \times \frac{1}{3} \left(\frac{T_D}{T} \right)^3 = 1.
\end{aligned} \tag{7.48}$$

Equation (7.47) then becomes $E_{\text{tot}} = 3NkT$, which is the expected classical result mentioned just after (7.44). Remember that we are calculating the molar heat capacity from (7.1), where the E in that equation is the total energy, called E_{tot} here:

$$C_V^{\text{mol}} = \frac{N_A}{N} \left(\frac{\partial E_{\text{tot}}}{\partial T} \right)_{V,N} = \frac{N_A}{N} \times 3Nk = 3R, \tag{7.49}$$

where $R = N_A k$ is the gas constant. Hence, the high-temperature limit of Debye's theory gives the Dulong–Petit result, as it should.

- **In the low-temperature limit**, $T_D/T \rightarrow \infty$, and the value of the resulting integral is known:

$$D(T/T_D) \rightarrow 3 \left(\frac{T}{T_D} \right)^3 \underbrace{\int_0^\infty \frac{x^3 dx}{e^x - 1}}_{= \pi^4/15} = \frac{\pi^4}{5} \left(\frac{T}{T_D} \right)^3. \tag{7.50}$$

Here, (7.47) becomes

$$E_{\text{tot}} = 3NkT \times \frac{\pi^4 T^3}{5T_D^3} = \frac{3N\pi^4 k T^4}{5T_D^3}. \tag{7.51}$$

The molar heat capacity in (7.1) is now

$$C_V^{\text{mol}} = \frac{N_A}{N} \frac{\partial E_{\text{tot}}}{\partial T} = \frac{N_A}{N} \frac{12N\pi^4 k T^3}{5T_D^3} = \frac{12\pi^4 R}{5} \left(\frac{T}{T_D} \right)^3. \tag{7.52}$$

This indeed shows the experimentally observed dependence on T^3 in crystals with no conduction electrons, which was mentioned at the end of Section 7.1. A century ago, this result for the heat capacity made Debye's model of energy transport by phonons the premier model of heat capacity for such crystals. In the case of metals at low temperatures, the presence of conduction electrons leads to C_V^{mol} being proportional to T rather than T^3 . We'll find the explanation for this in Chapter 8.

Let's show (7.52)'s agreement with the experimentally determined molar heat capacity of copper, at a temperature that is much less than copper's Debye temperature, but is high enough that C_V^{mol} is still dependent on T^3 rather than T . We require copper's Debye temperature:

$$T_D = \frac{E_D}{k} \stackrel{(7.42)}{=} \left(\frac{3N}{4\pi V} \right)^{1/3} \frac{hc}{k}. \quad (7.53)$$

To find a numerical value for this, we need the number of copper atoms per unit volume, N/V , and the mean speed of sound in copper, c . We know that

$$N \times \text{mass of 1 atom} = \text{mass of volume } V = \varrho V, \quad (7.54)$$

where ϱ is copper's mass density. It follows that

$$N \times \frac{M_{\text{mol}}}{N_A} = \varrho V, \quad (7.55)$$

or

$$\frac{N}{V} = \frac{\varrho N_A}{M_{\text{mol}}}. \quad (7.56)$$

Copper's density is $\varrho = 8933 \text{ kg/m}^3$, and its molar mass is $M_{\text{mol}} = 63.5 \text{ g}$. Its atoms' number density is then

$$\frac{N}{V} = \frac{8933 \times 6.022 \stackrel{23}{\text{---}}}{63.5 \stackrel{3}{\text{---}}} \text{ m}^{-3} \simeq 8.47 \times 10^{28} \text{ m}^{-3}. \quad (7.57)$$

The mean speed of sound c of the phonons through copper is the cubic harmonic mean of three speeds, following the analysis in Section 2.6. The speed of longitudinal sound waves through copper is about 4400 m/s, and that of transverse sound waves is about 2240 m/s. The value of c is then the cubic harmonic mean of 4400 m/s and two lots of 2240 m/s (because both transverse polarisations have this speed). Equation (2.113) becomes

$$\frac{3}{c^3} = \frac{1}{4400^3} + \frac{2}{2240^3} \text{ SI units}, \quad (7.58)$$

leading to $c \simeq 2510 \text{ m/s}$. (Notice that the cubic harmonic mean favours lower speeds heavily.) Equation (7.53) becomes

$$\begin{aligned} T_D &= \left(\frac{3N}{4\pi V} \right)^{1/3} \frac{hc}{k} = \left(\frac{3}{4\pi} \times 8.47 \stackrel{28}{\text{---}} \right)^{1/3} \times \frac{6.626 \stackrel{-34}{\text{---}} \times 2510}{1.381 \stackrel{-23}{\text{---}}} \text{ K} \\ &\simeq 328 \text{ K}. \end{aligned} \quad (7.59)$$

The molar heat capacity of copper in the previously discussed limit of low temperature is then

$$\begin{aligned} C_V^{\text{mol}} \stackrel{(7.52)}{=} \frac{12\pi^4 R}{5T_D^3} T^3 &= \frac{12\pi^4 \times 8.314}{5 \times 328^3} \text{ SI units} \times T^3 \\ &\simeq 5.5 \times 10^{-5} \text{ J K}^{-4} \text{ mol}^{-1} \times T^3. \end{aligned} \quad (7.60)$$

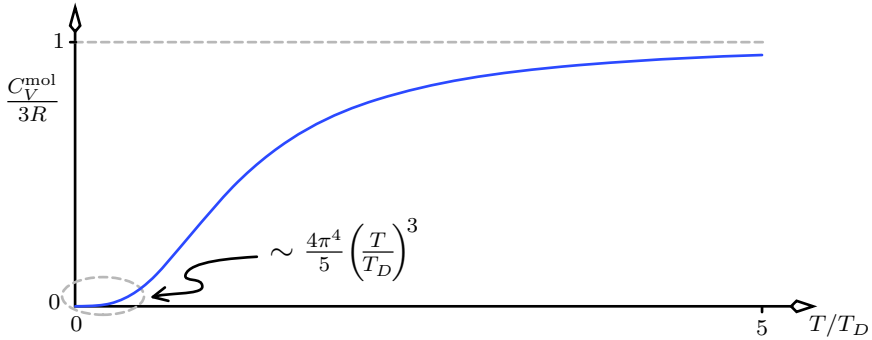


Fig. 7.7 The molar heat capacity (7.61) divided by $3R$, as predicted from Debye's theory. Its high- and low-temperature limits are given in (7.49) and (7.52) respectively. Experimental values of $C_V^{\text{mol}}/(3R)$ for a variety of substances lie on this curve up to an accuracy of its printed thickness. Compare this plot with Einstein's prediction in Figure 7.1

This coefficient of 5.5×10^{-5} SI units for copper agrees with the experimental value of about 5×10^{-5} SI units. This, together with similar agreements for other materials, was a resounding early success for Debye's theory.

The agreement with experiment of the above high- and low-temperature values gives us confidence to calculate Debye's prediction of the molar heat capacity for an arbitrary temperature. Do this by using (7.1) to write [and remember that E in (7.1) is called E_{tot} here]

$$\begin{aligned}
 C_V^{\text{mol}} &= \frac{\partial}{\partial T} \frac{N_A E_{\text{tot}}}{N} \stackrel{(7.47)}{=} \frac{d}{dT} \left[\frac{9RT^4}{T_D^3} \int_0^{T_D/T} \frac{x^3 dx}{e^x - 1} \right] \\
 &= \frac{36RT^3}{T_D^3} \int_0^{T_D/T} \frac{x^3 dx}{e^x - 1} + \frac{9RT^4}{T_D^3} \frac{(T_D/T)^3}{e^{T_D/T} - 1} \times \frac{-T_D}{T^2} \\
 &= 3R \left[4D(T/T_D) - \frac{3T_D/T}{e^{T_D/T} - 1} \right]. \tag{7.61}
 \end{aligned}$$

This function is shown in Figure 7.7. Experimental values of $C_V^{\text{mol}}/(3R)$ for a variety of substances match this curve very precisely, with experimental errors lying within the printed thickness of the curve.

In summary, the Einstein and Debye predictions can be compared via:

$$\begin{aligned}
 C_V^{\text{mol}} \text{ formula: } & \text{Einstein's (7.12),} \quad \text{and Debye's (7.61);} \\
 C_V^{\text{mol}} \text{ plot: } & \text{Einstein's Figure 7.1, and Debye's Figure 7.7.} \tag{7.62}
 \end{aligned}$$

But Debye's work uses T_D , whereas Einstein's uses T_E ; so, no direct comparison between the two can really be made. Even so, if we simply set $T_E = T_D$,

then

$$C_V^{\text{mol}}(\text{Debye}) > C_V^{\text{mol}}(\text{Einstein}), \quad (7.63)$$

and the ratio of these quantities has the following limits:

$$\lim_{T \rightarrow 0} \frac{C_V^{\text{mol}}(\text{Debye})}{C_V^{\text{mol}}(\text{Einstein})} = \infty, \quad \lim_{T \rightarrow \infty} \frac{C_V^{\text{mol}}(\text{Debye})}{C_V^{\text{mol}}(\text{Einstein})} = 1. \quad (7.64)$$

7.4 Gibbs' Paradox and Its Resolution

Historically, the core ideas of statistical mechanics were established well before the appearance of quantum mechanics. Quantum mechanics introduced the new idea that a set of particles can be *absolutely* identical, meaning that no particle has its own individual identity. This idea of complete indistinguishability turns out also to have a place in classical statistical mechanics. To see why, consider an ideal gas of point particles in a box. A central partition divides the volume into halves, as shown in Figure 7.8. On removing the partition, we expect that the gases in each half will mix. But, surely, removing the partition has not changed the nature of the gas, and so we wouldn't expect its entropy to increase as per the Second Law—or would we? Let's calculate this entropy before and after removing the partition.

Initially, each half of the box holds N *distinguishable* particles in a volume V . After the partition is removed, $2N$ particles are spread through a volume $2V$. Equation (3.145) gives the entropy of an ideal gas of distinguishable point particles:

$$S^{\text{dist}}(N, V) \simeq Nk \left[\ln V + \frac{3}{2} + \frac{3}{2} \ln \frac{2\pi mkT}{h^2} \right]. \quad (7.65)$$

The total entropy before removing the partition is the sum of the two halves, which is

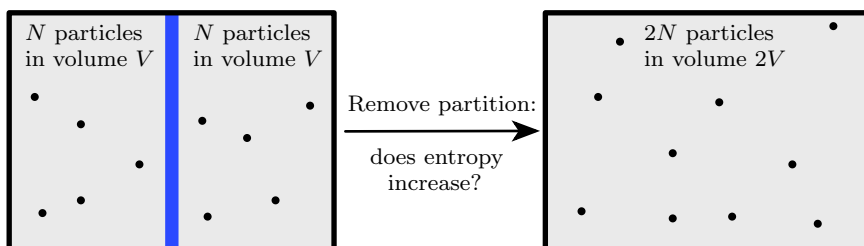


Fig. 7.8 **Left:** Before we remove the partition (in blue), the box contains two cells of an ideal gas of point particles that cannot mix with each other. **Right:** After the partition is removed, the two cells of gas are free to mix. Does the total entropy increase as a result?

$$2S^{\text{dist}}(N, V) \simeq 2Nk \left[\ln V + \frac{3}{2} + \frac{3}{2} \ln \frac{2\pi mkT}{h^2} \right]. \quad (7.66)$$

After the partition is removed, the total entropy is that of $2N$ particles in volume $2V$:

$$S^{\text{dist}}(2N, 2V) \simeq 2Nk \left[\ln(2V) + \frac{3}{2} + \frac{3}{2} \ln \frac{2\pi mkT}{h^2} \right]. \quad (7.67)$$

The entropy increase is then

$$S^{\text{dist}}(2N, 2V) - 2S^{\text{dist}}(N, V) = 2Nk \ln 2. \quad (7.68)$$

The entropy has increased because the particles are like little numbered billiard balls, and the number of ways in which such balls can be arranged increases when they are allowed to mix, just as we found in Chapter 1.

But can the particles of, say, oxygen gas really be treated as little numbered billiard balls? In the nineteenth century, Josiah Gibbs drew attention to the idea that it might not be reasonable to suppose that the entropy of, say, pure oxygen increases when the partition is removed. This situation is known as *Gibbs' paradox*. One way to resolve it is to take seriously the modern idea that identical particles really are fundamentally identical, just as each of the dollars in a bank account are identical. Suppose, then, that we treat the particles of a pure gas such as oxygen as identical classical. This requires reducing the number of states by a factor of $N!$, as discussed in Sections 1.1.1 and 2.4.2. Recall the entropy of an ideal gas of identical-classical point particles in (3.146):

$$S^{\text{ic}}(N, V) = Nk \left[\ln \frac{V}{N} + \frac{5}{2} + \frac{3}{2} \ln \frac{2\pi mkT}{h^2} \right]. \quad (7.69)$$

Use this entropy to re-analyse the scenario; we will even generalise the two compartments to be n compartments that each hold N identical-classical particles in a volume V . Now, the initial entropy is n times the entropy of one compartment holding N identical particles in a volume V :

$$\text{initial entropy} = nS^{\text{ic}}(N, V) = nNk \left[\ln \frac{V}{N} + \frac{5}{2} + \frac{3}{2} \ln \frac{2\pi mkT}{h^2} \right]. \quad (7.70)$$

The final entropy is that of one compartment holding nN identical particles in a volume nV :

$$\text{final entropy} = S^{\text{ic}}(nN, nV) \simeq nNk \left[\ln \frac{nV}{nN} + \frac{5}{2} + \frac{3}{2} \ln \frac{2\pi mkT}{h^2} \right]. \quad (7.71)$$

But this just equals the initial entropy! So, the total entropy does *not* increase when the partitions are removed, and Gibbs' paradox is resolved. It seems,

then, that we might take seriously the idea that identical particles really are identical in the deepest way.

7.5 The Extent of a System's Quantum Nature

The phonons of Debye's model revealed one aspect of this type of quantum particle: that it can be associated with excited energy levels of oscillators. Another aspect of the quantum nature or "quantumness" of a set of particles is the extent to which they can crowd together in phase space. The identical-classical particles introduced back in Section 1.1.1 had something of a quantum nature: they were identical, but were still spread widely enough in phase space so as to appear classical to all intents and purposes. For example, all electrons are quantum mechanically identical, and electrons bound to the same nucleus must be treated quantum mechanically by, say, invoking the Pauli exclusion principle. But two electrons in separate pieces of metal can be treated as distinguishable, because their wave functions have negligible overlap. These are identical-classical particles: technically identical, but still able to be treated (somewhat) classically.

The degree to which a set of identical particles must be treated using quantum mechanics can be estimated by comparing their de Broglie wavelength λ to their typical spacing from each other when treated purely classically. (Recall that a particle's de Broglie wavelength is $\lambda = h/p$, where h is Planck's constant and p its momentum.) If λ is much greater than this classical spacing, the particles must be considered as overlapping, and the system must be treated quantum mechanically. The classical particle separation results from envisaging each particle as being allocated, say, a cube of space. If N particles occupy a total volume V , then each particle can be imagined to lie at one corner of a cube with volume V/N . The particles' classical spacing is then this cube's side length of $(V/N)^{1/3}$. Hence, we can say

$$\lambda \gg (V/N)^{1/3} \iff \text{the particles are very quantum in nature,} \quad (7.72)$$

in which case quantum mechanics is required to analyse them. Some examples follow.

1. **Air:** Treat air as an ideal gas at room temperature. Assume, for simplicity, that all the molecules have the same momentum (while moving in all directions, of course). We relate this momentum to their kinetic energy, which is determined from the equipartition theorem with three translational quadratic energy terms per particle. A molecule's de Broglie wavelength is then

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2mE}} = \frac{h}{\sqrt{2m \times \frac{3}{2}kT}} = \frac{h}{\sqrt{3mkT}}. \quad (7.73)$$

The average mass of an air molecule is $29.0 \text{ g}/N_A$, or $4.8 \times 10^{-26} \text{ kg}$. At a room temperature of $T = 298 \text{ K}$, each particle's quantum extent is then (using SI units throughout)

$$\lambda \simeq \frac{6.626 \overset{-34}{\text{m}}}{\sqrt{3 \times 4.8 \overset{-26}{\text{kg}} \times 1.381 \overset{-23}{\text{J/K}} \times 298}} \simeq 0.03 \text{ nm}. \quad (7.74)$$

The classical volume per particle is, with P the air pressure (atmospheric, $101,325 \text{ Pa}$),

$$\frac{V}{N} = \frac{kT}{P} \simeq \frac{1.381 \overset{-23}{\text{J/K}} \times 298}{101,325} \text{ m}^3, \quad (7.75)$$

which leads to a classical spacing of

$$(V/N)^{1/3} \simeq 3 \text{ nm}. \quad (7.76)$$

The de Broglie wavelength of 0.03 nm is much smaller than the classical particle spacing of 3 nm . We conclude that this air can be treated as a *very* classical collection of particles.

2. **Helium:** We can produce a gas with a greater value of λ by reducing both the particle's mass m and the temperature. Helium gas has a representative de Broglie wavelength of

$$\lambda = \frac{h}{\sqrt{3mkT}} \simeq \frac{6.626 \overset{-34}{\text{m}}}{\sqrt{3 \times \frac{4 \overset{-3}{\text{kg}}}{6.022 \overset{23}{\text{mol}}}} \times 1.381 \overset{-23}{\text{J/K}} \times \frac{T}{1 \text{ K}}} \simeq \frac{1.3 \text{ nm}}{\sqrt{T/(1 \text{ K})}}. \quad (7.77)$$

At higher temperatures, where helium is a gas, its atoms' classical spacing at atmospheric pressure is

$$\left(\frac{V}{N}\right)^{1/3} = \left(\frac{kT}{P}\right)^{1/3} \simeq \left(\frac{1.381 \overset{-23}{\text{J/K}} \times T/(1 \text{ K})}{101,325}\right)^{1/3} \text{ m} \simeq 0.5 [T/(1 \text{ K})]^{1/3} \text{ nm}. \quad (7.78)$$

Helium liquefies at 4.22 K , and so below this temperature, we recall (7.56) to write

$$\left(\frac{V}{N}\right)^{1/3} = \left(\frac{M_{\text{mol}}}{\varrho N_A}\right)^{1/3} \simeq \left(\frac{4 \overset{-3}{\text{kg}}}{120 \times 6.022 \overset{23}{\text{mol}}}\right)^{1/3} \text{ m} \simeq 0.4 \text{ nm}. \quad (7.79)$$

Equations (7.77)–(7.79) allow values of λ and $(V/N)^{1/3}$ to be calculated for various low temperatures, as shown in Table 7.1. Clearly, liquid helium requires a full quantum-mechanical treatment below about 2 kelvins .

Table 7.1 A comparison of de Broglie wavelength λ with the representative particle spacing $(V/N)^{1/3}$ for helium, in both liquid and gaseous form. When the temperature drops below about 2 K, λ starts to become greater than $(V/N)^{1/3}$, and so quantum mechanics must be used to analyse the helium

T (K)	0.5	1	2	5	10	100
λ (nm)	1.8	1.3	0.9	0.6	0.4	0.1
$(V/N)^{1/3}$ (nm)	0.4	0.4	0.4	0.9	1.1	2.4

3. Conduction electrons in copper metal: When these electrons are treated as an ideal gas at room temperature, the usual de Broglie wavelength results:

$$\lambda = \frac{h}{\sqrt{3mkT}} \simeq \frac{6.626 \overset{-34}{\text{m}}}{\sqrt{3 \times 9.11 \overset{-31}{\text{kg}} \times 1.381 \overset{-23}{\text{J/K}} \times 298}} \simeq 6 \text{ nm}. \quad (7.80)$$

To calculate $(V/N)^{1/3}$, realise that each copper atom produces one conduction electron, and so the classical volume occupied by one electron equals the volume occupied by one copper atom. We calculated the reciprocal of this latter value in (7.57). It follows that

$$(V/N)^{1/3} \simeq \left(\frac{1}{8.47 \overset{-28}{\text{m}^3}} \right)^{1/3} \text{ m} \simeq 0.23 \text{ nm}. \quad (7.81)$$

A conduction electron's de Broglie wavelength of 6 nm is very much larger than the classical electron spacing of 0.23 nm, and so the conduction electrons must certainly be treated quantum mechanically. We'll do just that in Chapter 8, where we'll discover why treating the electrons as a gas of non-interacting particles works so well.

7.5.1 Average de Broglie Wavelength

In (7.73), we wrote the de Broglie wavelength of particles in a gas as

$$\lambda = \frac{h}{\sqrt{3mkT}}, \quad (7.82)$$

by assuming that the particles all have exactly the same momenta (while moving in different directions). In practice, their momenta follow a Maxwell speed distribution. For a gas of N particles following this distribution, their *mean* de Broglie wavelength $\langle \lambda \rangle$ is governed by their probability density of having speed v , which is $\mathcal{N}(v)/N$ from (6.37):

$$\begin{aligned}
\langle \lambda \rangle &= \left\langle \frac{h}{p} \right\rangle = \frac{h}{m} \left\langle \frac{1}{v} \right\rangle = \frac{h}{m} \int_0^\infty \frac{1}{v} \frac{\mathcal{N}(v)}{N} dv \\
&\stackrel{(6.37)}{=} \frac{h}{m} \sqrt{\frac{2}{\pi}} \left(\frac{m}{kT} \right)^{3/2} \int_0^\infty v \exp \frac{-mv^2}{2kT} dv \stackrel{(1.95)}{=} \frac{h\sqrt{2}}{\sqrt{\pi mkT}}. \quad (7.83)
\end{aligned}$$

We might also calculate those particles' rms de Broglie wavelength λ_{rms} :

$$\begin{aligned}
\lambda_{\text{rms}}^2 &\equiv \langle \lambda^2 \rangle = \frac{h^2}{m^2} \left\langle \frac{1}{v^2} \right\rangle = \frac{h^2}{m^2} \int_0^\infty \frac{1}{v^2} \frac{\mathcal{N}(v)}{N} dv \\
&\stackrel{(6.37)}{=} \frac{h^2}{m^2} \sqrt{\frac{2}{\pi}} \left(\frac{m}{kT} \right)^{3/2} \int_0^\infty \exp \frac{-mv^2}{2kT} dv \stackrel{(1.90)}{=} \frac{h^2}{mkT}, \quad (7.84)
\end{aligned}$$

which leads to

$$\lambda_{\text{rms}} = \frac{h}{\sqrt{mkT}}. \quad (7.85)$$

These quantities, the simplified λ along with $\langle \lambda \rangle$ and λ_{rms} , all have similar values, and using any one is fine for general calculations. For example, the entropies of distinguishable and identical-classical gases of point particles, (3.145) and (3.146), are written as

$$\begin{aligned}
S^{\text{dist}} &= Nk \left[\ln \frac{V}{\langle \lambda \rangle^3} + \frac{3}{2} - \ln 8 \right] \simeq Nk \left[\ln \frac{V}{\langle \lambda \rangle^3} - 0.58 \right], \\
S^{\text{ic}} &= Nk \left[\ln \frac{V}{N \langle \lambda \rangle^3} + \frac{5}{2} - \ln 8 \right] \simeq Nk \left[\ln \frac{V}{N \langle \lambda \rangle^3} + 0.42 \right], \quad (7.86)
\end{aligned}$$

provided the temperature doesn't tend toward zero. Evident here is the competition between volume and de Broglie wavelength. For particles to behave classically, (7.72) says that $(V/N) \gg \langle \lambda \rangle^3$, in which case

$$\ln \frac{V}{N \langle \lambda \rangle^3} > 0. \quad (7.87)$$

This suggests the following simple rule-of-thumb lower limits for the distinguishable and identical-classical entropies:

$$S^{\text{dist}} > Nk \ln N, \quad S^{\text{ic}} > Nk. \quad (7.88)$$

The mean de Broglie wavelength $\langle \lambda \rangle$ is sometimes *halved* to give a quantity that has acquired the name “thermal wavelength” or “thermal de Broglie wavelength”:

$$\frac{\langle \lambda \rangle}{2} \stackrel{(7.83)}{=} \frac{h}{\sqrt{2\pi mkT}}. \quad (7.89)$$

This expression appears widely in statistical mechanics. The fact that it is one *half* of the mean de Broglie wavelength is of no real significance, since it usually appears with other factors anyway, such as in (2.50). Even so, its widespread appearance might lead one to think that a more natural expression for the characteristic length of a wave is $\lambda/2$, rather than λ . That might well be true for discussions such as Section 2.2, where whole numbers of half wavelengths are being fitted between walls. On the other hand, consider that a sinusoid is naturally written in terms of its wave number k and angular frequency ω as

$$y = \sin(kx - \omega t) = \sin\left(\frac{2\pi x}{\lambda} - \frac{2\pi t}{T}\right). \quad (7.90)$$

In that case, if we argue that $\lambda/2$ is a more natural choice than λ for a wave's length, then we might also want to argue that $T/2$ is a more natural choice than T for a wave's period. And *that* is not such a clear argument to make.

7.6 Fermions and Bosons

When studying the Boltzmann distribution in Chapter 5, we focussed on a single particle such as a hydrogen atom, and calculated the chance that it could occupy any one of its available energy levels. This was equivalent to finding the mean number of hydrogen atoms occupying each level. This mean number was larger for low-energy levels and decreased for higher-energy levels, because thermal agitation will seldom give a hydrogen atom enough energy to occupy a higher-energy level.

In contrast, Einstein's theory introduces the massless non-interacting indistinguishable quantum particles described earlier in this chapter, and each energy level's quantum number [n in (7.3)] is re-interpreted as the number of these quantum particles present for each oscillator (that is, for each dimension of oscillation of each crystal molecule), or “per state”. In Einstein's theory, a state was no longer treated as an energy level of one oscillator: instead, a state was viewed as the oscillator being associated with n quantum particles. These particles came in a single variety—having energy hf —and the average number of particles over all states was $\bar{n}$ in (7.23).

In Debye's theory, this idea evolved into the average number of these quantum particles present per normal mode of oscillation of the entire crystal. A state could now be defined as a kind of “box” associated with each of these normal modes. This box was labelled with “Energy E ”, and would contain some number of quantum particles, each having energy E by virtue of being in that box. We introduced the occupation number $\bar{n}(E, T)$ in (7.30) as the average number of these quantum particles per box over all $3N$ boxes, mean-

ing the average number of these quantum particles per state at energy E and temperature T .

The different interpretations or definitions of a state here can certainly be confusing. To summarise:

- For an *ideal gas* with N particles, a state is a cell in $6N$ -dimensional phase space, with a cell “volume” h^{DN} , where D is the number of internal variables into which a particle can store its energy (which is *not* necessarily the number of quadratic energy terms: see Table 2.1).
- When using a *quantum description of a set of atoms*, each atom occupies a single quantum state labelled by a set of quantum numbers. In general, several states exist at each energy level.
- For a *more complicated quantum system*, such as a set of oscillators in Einstein’s model of heat capacity, a state is one dimension of oscillation of a single oscillator. A crystal of N atoms has $3N$ states. Each state can be occupied by some number n of massless quantum particles, each of which having energy hf , where h is Planck’s constant and f is the oscillators’ common frequency.
- In *Debye’s model of a crystal*, a state is a normal mode of oscillation of the entire crystal. A crystal of N atoms has $3N$ normal modes of oscillation, and thus has $3N$ states of various energies E . Each state can be occupied by some number n of massless quantum particles called phonons. Each phonon’s energy is given by the energy “label” E of the state that it occupies.

The phonons in Debye’s model (along with their prototype in Einstein’s model) are gregarious: there can be any number present in a crystal. Such particles are called *bosons*. These particular bosons are massless, but massive particles can also be bosons. To describe the gregariousness—or otherwise—of particles, we must properly account for high particle densities. We saw an inkling of such high densities previously in Section 7.5, with copper’s conduction electrons—although, of course, electrons are not massless particles and also turn out not to be bosons.

To apply some of the above ideas—particularly occupation number—to massive bosons and other particles that are not bosons, we return to the ideas of Sections 2.3 and 2.4. There, we studied the higher-dimensional phase space in which an entire classical system of particles occupies a single point: for example, the gas of N distinguishable point particles can be represented at each moment by a single point moving in a $6N$ -dimensional phase space. The Heisenberg uncertainty principle then partitioned that phase space into cells whose “volume” was set by Planck’s constant h . To investigate the extent to which a system must be treated quantum mechanically, we now examine the density of its particles in the “everyday” 6-dimensional position–momentum

space, where this space has 3 position and 3 momentum coordinates. In principle, each cell of that position–momentum space contains some number of particles; in practice, any cell can be empty for a moment, because the occupancy of each cell is a function of time.⁴

In particular, we find the occupation number $\bar{n}(E, T)$ of this 6-dimensional position–momentum space by counting the numbers of particles in its cells. In Section 3.3, we likened energy to the money in a bank account: no matter what form the funds originally took, after being deposited in a bank account, each quantum of currency (such as a dollar) lost its individuality. We cannot visit the bank and ask to withdraw “the third dollar” from the \$100 that we deposited yesterday, because the dollars in the account are completely indistinguishable. This loss of individuality also applies to identical particles, and it has consequences when counting particles to determine an occupation number.

For the sake of argument, consider two *distinguishable* particles, and revisit the discussion around Figures 2.11 and 2.12. In those figures, each particle occupies a cell of 2-dimensional position–momentum space (one position and one momentum coordinate). In Figure 2.11, we argued that when the particles are distinguishable and occupy different cells, the number of ways of occupying the various cells (the number of microstates) is $2!$ times as large as the number of ways for when the particles are identical. In Figure 2.12, we argued that when the particles are distinguishable and occupy the *same* cell, the number of ways of occupying the various cells is *equal* to the case for when the particles are identical. In what follows, we will always be referring to microstates, but will shorten the word to “states” for conciseness.

These two figures are combined in Figure 7.9, but now with more detail. In the upper left in the figure, we see that the case of particle 1 occupying the upper-left cell in the plot and particle 2 occupying the lower-right cell defines a separate state from the case of the particles being swapped: there are two distinct states here. Furthermore, both particles might occupy the same cell (bottom-left in figure); this constitutes a third state of the two-particle system. There are $\Omega_{\text{dist}} = 3$ states of distinguishable particles here.

Recall, from Figure 2.11, that if the two particles are identical classical, then when they occupy separate cells, we divide the number of states for the distinguishable case by $2!$ to get the correct number for the identical-classical case. That procedure works for the upper-left two states in Figure 7.9. But (recollect Figure 2.12) it fails when the particles occupy the same cell. In Figure 7.9, we would naïvely divide $\Omega_{\text{dist}} = 3$ states by $2!$ to infer that the number of states for identical-classical particles is $\Omega_{\text{ic}} = 1.5$. Clearly, some-

⁴ We are dealing with a single 6-dimensional position–momentum space above. But recall the distinction made in Section 2.3 between position–momentum space and phase space. If each of the N particles is given its own personal 6-dimensional *position–momentum* space, then joining N of those spaces together results in the $6N$ -dimensional *phase* space. In that phase space, the entire set of N particles occupies a single cell.

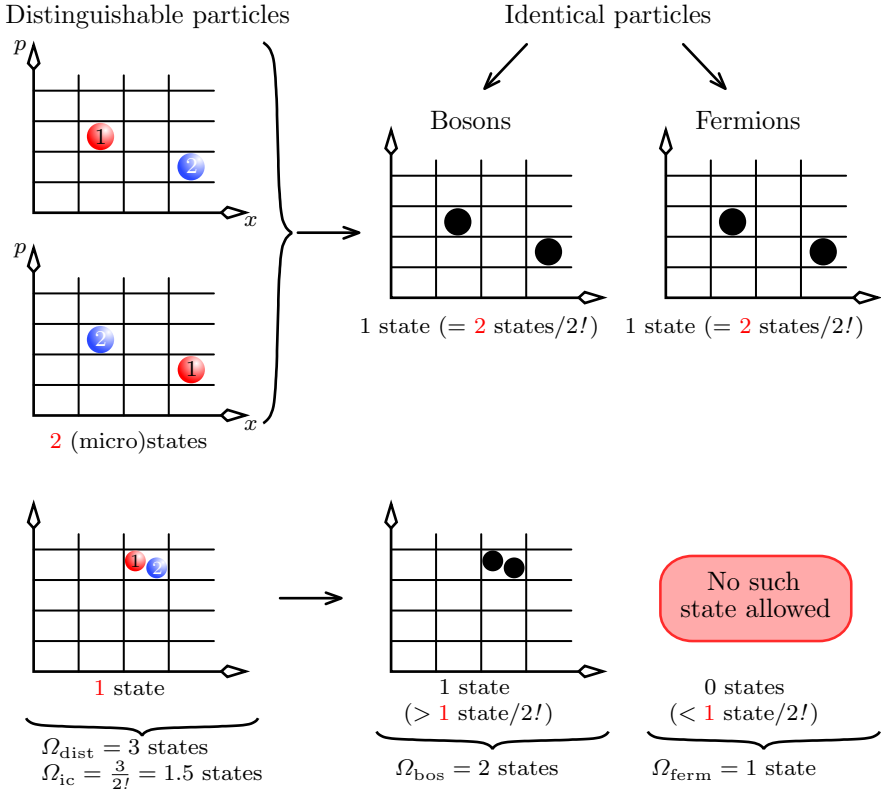


Fig. 7.9 The (micro)states accessible to distinguishable particles (subscript “dist”), identical quantum particles (subscripts “bos” and “ferm”, for bosons and fermions), and identical-classical particles (subscript “ic”). Each two-dimensional position–momentum plot defines one state. At upper left, two distinguishable particles each occupy separate cells, and thus have two states available for these cell occupations (these are the two plots shown). When these particles are identical (upper right), irrespective of whether they are bosons or fermions, the number of states occupied by these particles obeys the “divide by N !” rule (where $N = 2$ particles here). The red digits are the numbers (of states of distinguishable particles) that are being divided by $2!$ to predict the numbers of states that identical particles might occupy. Two distinguishable particles can also occupy the same cell (bottom left), in which case the total number of states (plots in left column) is $\Omega_{\text{dist}} = 3$; but this number cannot be divided by $2!$ to predict the number of states for identical-classical particles. Also, when such particles are fully quantum mechanical, whether or not they can occupy the same cell depends on whether they are bosons or fermions

thing isn’t right. The problem is that, our prescription of dividing the number of states that was produced by assuming the particles are distinguishable, is based on those particles occupying different cells in the position–momentum plot. But those particles *don’t* occupy different cells at the bottom left in Figure 7.9.

Apart from identical-classical particles, of real interest is the case of truly identical *quantum* particles, which (unlike the identical-classical case) really do exist in Nature. We have already encountered bosons in our descriptions of Einstein's and Debye's models of heat capacity. As far as is known, Nature admits only one other type of identical particle: the *fermion*. Unlike the gregarious bosons, fermions are solitary: at most, *one* fermion can occupy a given cell in the plots of Figure 7.9.

- Any number of *bosons* can occupy a given cell in Figure 7.9;
- At most, only one *fermion* can occupy a given cell in Figure 7.9.

When bosons or fermions are not crammed together in their quantum-mechanical state space, we can model them as identical classical: we count the number of states available to each type by numbering the particles as in the distinguishable case (upper left in Figure 7.9) and dividing the result by $2!$. Hence, the two upper-left states in Figure 7.9 count as a single state when the particles are bosons or fermions.

When a large number of particles is present, or when the number of available cells is *not* much larger than the number of particles, some particles can be forced to cram into the same cell. Already, we see a problem with the identical-classical particles, for which simply dividing the total number of states for distinguishable particles ($\Omega_{\text{dist}} = 3$ in Figure 7.9) by $2!$ has the unintended side effect of counting as “half a state” the bottom-left state in the figure. This state would be unchanged if the particles were bosons (middle bottom in the figure)—but it would be a whole state, not half a state. Our rule of dividing the number of states occupied by distinguishable particles by $2!$ fails when bosons are crammed together, because doing so produces $1/(2!)$ or half a state being occupied by the bosons, rather than the one state in Figure 7.9.

And in the case of fermions, no two fermions can occupy the same cell in the figure, and so no such state is defined for them. So, our rule of dividing the number of states for distinguishable particles by $2!$ has failed again, since $1/(2!)$ is not zero. The fact is that bosons can occupy *more* states than our “divide by $2!$ ” rule would suggest (1 instead of $1/2$), and fermions will occupy *fewer* states than the rule would suggest (0 instead of $1/2$).

Our aim is to calculate occupation numbers $\bar{n}(E, T)$ for bosons and fermions. We have already done this for the case of what we might call Einstein's “proto-phonons” in (7.18)–(7.23): with n such particles present in each state, we calculated the mean of n with a weighted sum of all n using the probability p_n that the state is occupied. This probability is given by the Boltzmann distribution. The calculation for more general bosons, and also fermions, follows the same lines; but now we allow for a more general state energy than the “ $n\hbar f + \varepsilon$ ” of (7.18). Begin again with (7.18):

$$\bar{n} = \sum_n n p_n . \quad (7.91)$$

Recall the Boltzmann distribution in (5.5),

$$p(E_s, V_s, N_s) \propto \Omega_s \exp \frac{-E_s - PV_s + \mu N_s}{kT}. \quad (7.92)$$

Set $\Omega_s = 1$, because we are considering a single state. A state has n particles, so set $N_s = n$. A little after (7.25), we switched viewpoints from a state as having an energy equal to the total energy of the particles occupying it, to a *single* energy, which it “bestowed” on each particle occupying it. The value of E_s is still the total energy of the particles in the state—a change in viewpoint is fine, but the Boltzmann distribution stills requires the *total* energy of the system under analysis. Hence, set

$$E \equiv \left[\begin{array}{l} \text{energy of a state, which that state “bestows”} \\ \text{on each particle occupying it} \end{array} \right]. \quad (7.93)$$

Each of the particles in the state has this energy E , rather than $nhf + \varepsilon$, and so the system’s total energy is $E_s = nE$. Also, these quantum states—such as the normal modes of Debye’s oscillating crystal—can be considered to have a fixed volume,⁵ and hence the volume term in (7.92) cancels out in the normalisation of the probability (that is, in the partition function). The $p(E_s, V_s, N_s)$ in (7.92) is now the probability of n quantum particles being present in a state. Equation (7.92) becomes

$$p_n \propto \exp \frac{-nE + \mu n}{kT} = \exp \frac{-n(E - \mu)}{kT}. \quad (7.94)$$

For shorthand, write

$$\alpha \equiv \frac{-(E - \mu)}{kT}. \quad (7.95)$$

The probability of n particles being present in a state is then $p_n \propto e^{n\alpha}$, or

$$p_n = \frac{e^{n\alpha}}{Z}, \quad (7.96)$$

where the partition function is

$$Z = \sum_n e^{n\alpha}. \quad (7.97)$$

We have seen this calculation of $\bar{n}$ before in (7.18)–(7.21), but can repeat the details here. Begin by noticing that

$$\frac{\partial Z}{\partial \alpha} = \sum_n n e^{n\alpha}. \quad (7.98)$$

⁵ If indeed a volume should even be defined for them. Recall that the pressure/volume term in the Boltzmann distribution is representative of mechanical interactions with the bath.

It follows that the occupation number $\bar{n}$ is

$$\bar{n} = \sum_n n p_n \stackrel{(7.96)}{=} \frac{1}{Z} \sum_n n e^{n\alpha} = \frac{1}{Z} \frac{\partial Z}{\partial \alpha}. \quad (7.99)$$

We must now work with fermions and bosons separately, since each type of particle takes a different set of values for n in the partition function sum (7.97).

- **Fermions:** The calculation of $\bar{n}$ is very simple, because for fermions, the number n of particles per quantum state can be only 0 or 1. The partition function (7.97) becomes

$$Z_{\text{ferm}} = \sum_{n=0}^1 e^{n\alpha} = 1 + e^{\alpha}. \quad (7.100)$$

Applying (7.99) yields

$$\bar{n}_{\text{ferm}} = \frac{1}{Z_{\text{ferm}}} \frac{\partial Z_{\text{ferm}}}{\partial \alpha} = \frac{1}{1 + e^{\alpha}} \frac{\partial}{\partial \alpha} (1 + e^{\alpha}) = \frac{1}{e^{-\alpha} + 1}. \quad (7.101)$$

Of course, we could just as well have calculated $\bar{n}_{\text{ferm}}$ directly here from first principles:

$$\bar{n}_{\text{ferm}} = 0 \times p_0 + 1 \times p_1 = p_1; \quad (7.102)$$

and then, since $p_n \propto e^{n\alpha}$ and n equals only 0 or 1, the correct normalisation is trivial to write down:

$$\bar{n}_{\text{ferm}} = p_1 = \frac{e^{1\alpha}}{e^{0\alpha} + e^{1\alpha}} = \frac{e^{\alpha}}{1 + e^{\alpha}} = \frac{1}{e^{-\alpha} + 1}. \quad (7.103)$$

- **Bosons:** Any number of bosons can be present, and we are familiar with this case from the calculation for Einstein's model around (7.22). That is, the sum in (7.97) runs over all whole-number values of n and is simply an infinite geometric series:

$$Z_{\text{bos}} = \sum_{n=0}^{\infty} e^{n\alpha} = \frac{1}{1 - e^{\alpha}}. \quad (7.104)$$

The occupation number for bosons is thus

$$\bar{n}_{\text{bos}} = \frac{1}{Z_{\text{bos}}} \frac{\partial Z_{\text{bos}}}{\partial \alpha} = (1 - e^{\alpha}) \frac{\partial}{\partial \alpha} \frac{1}{1 - e^{\alpha}} = \frac{1}{e^{-\alpha} - 1}. \quad (7.105)$$

Placing (7.101) and (7.105) together, we note how remarkable it is that, while the rules differ dramatically for the numbers of fermions and bosons that can occupy a state, their occupation numbers differ only by a sign:

$$\bar{n}(E, T) = \frac{1}{\exp \frac{E - \mu(T)}{kT} \pm 1} \quad \begin{cases} \text{fermions} \\ \text{bosons.} \end{cases} \quad (7.106)$$

(Although the chemical potential μ is a function of temperature, we generally will not indicate its T dependence explicitly.) Compare (7.106) with Debye's occupation number of phonons in (7.30): it's apparent that phonons have a chemical potential of $\mu = 0$. This makes sense; after all, the existence of a chemical potential indicates a propensity for diffusion of particles, but because phonons can crowd together with an arbitrarily high number density, they offer no resistance to “incoming” phonons, and so they don't diffuse.

The application of statistical mechanics to fermions—with the relevant occupation number in (7.106)—is *fermion statistics*, often called *Fermi–Dirac statistics*. For bosons, again with the relevant occupation number in (7.106), we have *boson statistics*, often called *Bose–Einstein statistics*. Equation (7.106) gives the *Fermi–Dirac* and *Bose–Einstein distributions*.

We have already seen an example of boson statistics in Debye's model of heat capacity: phonons are bosons. In Chapter 8, we'll extend Debye's model by including the fermion statistics of valence electrons; and in Chapter 9, we'll use a bosonic treatment of photons to derive the theory of blackbody radiation.

Quantum mechanically, fermions turn out to have odd half-integral spin ($1/2, 3/2, \dots$), and examples are electrons, positrons, protons, neutrons, neutrinos, and muons. Bosons turn out to have whole-number spin ($0, 1, 2, \dots$), and include α particles, pions, photons, and deuterons. Every fundamental particle known is either a fermion or a boson: no other choice of spin seems to be allowed by Nature. Just why a particle's spin should determine its occupation number is a subject of relativistic quantum mechanics, but the fundamental reason is not yet understood.⁶

It's important to remember that both occupation numbers in (7.106) are a direct application of the Boltzmann distribution. In the limit of $(E - \mu)/(kT) \gg 1$ (that is, high energy/low temperature/low particle density), (7.106) becomes

$$\bar{n}(E, T) \simeq \exp \frac{-E + \mu}{kT}. \quad (7.107)$$

⁶ I defer here to Feynman's opinion of our understanding of spin. Probably, many quantum field theorists will disagree with him on this point. But I think the Great Hall of Quantum Field Theory should have a sign posted on its gates that reads “Abandon all hope of using your hard-won logic, ye who enter here”. Although all disciplines must, and do, pass through a “young” phase, the standard accounts of the subject seem to glorify an absence of mathematical rigor and logical flow, in a series of disjoint and somewhat garbled topics, full of epicycles and band-aids. But quantum field theory has had predictive successes in particle physics; and so the hope is that, perhaps in decades to come, it will transform into something coherent.

This occupation number defines *Maxwell–Boltzmann statistics*, which models both distinguishable and identical-classical particles. But there is certainly no suggestion here that the Boltzmann distribution itself is a limiting case of Fermi–Dirac and Bose–Einstein statistics. The Boltzmann distribution follows quite universally from the concept of entropy, as seen in Section 5.1: essentially, the probability of a system having energy E is inversely proportional to the number of states at E ; this number of states equals $e^{S/k}$ by the definition of entropy, and the First Law of Thermodynamics says that $S/k = \beta(E + PV - \mu N)$, resulting in the Boltzmann expression “probability $\propto \exp -\beta(E + PV - \mu N)$ ”. This applies throughout statistical mechanics, both classically and quantum mechanically. It is built into the quantum statistics of (7.106).

But we might ask how (7.107) relates to the hydrogen energy-level discussion in Sections 5.4 and 5.7, where we calculated the numbers of hydrogen atoms in various energy levels. If we consider the energy of a system as apportioned to different levels “ i ”, as in the case of the hydrogen atom, then the mean number of particles at level i in this high-energy regime is [with $\beta \equiv 1/(kT)$, as usual]

$$\bar{n}_i \stackrel{(7.107)}{=} \exp(-E_i\beta + \mu\beta). \quad (7.108)$$

The total number of particles, N , is the sum of the mean numbers of particles in each of the energy levels:

$$N = \sum_i \bar{n}_i \stackrel{(7.108)}{=} e^{\mu\beta} \sum_i e^{-E_i\beta}. \quad (7.109)$$

Hence, $e^{\mu\beta} = N / \sum_i e^{-E_i\beta}$, and (7.108) becomes

$$\bar{n}_i = \frac{e^{-E_i\beta} N}{\sum_i e^{-E_i\beta}} = p_i N, \quad \text{where } p_i \propto e^{-E_i\beta}. \quad (7.110)$$

This is just what we might expect for the mean number of particles per energy level, based on our discussion of the hydrogen atom in Section 5.4.

We reiterate that no Ω_i appeared in the discussion of this section, unlike, say, in (5.45). We discussed this in Section 5.3. The Ω_i in (5.45) merely accounts for degeneracy: the fact that several quantum states might all have the same energy. The probability for the system to occupy any *particular* one of those states is found by setting $\Omega_i = 1$.

7.7 Occupation Numbers of Fermion and Boson Gases

Insight into the behaviour of fermions and bosons can be gained from plotting the occupation number $\bar{n}(E, T)$ of a gas of fermions, and that of a gas of bosons, versus particle energy E in (7.106) for various values of temperature. But constructing such plots is complicated by the fact that it involves $\mu(T)$, whose temperature dependence might be unknown. Luckily, this temperature dependence can be extracted from what is essentially a *normalisation* equation: (7.34), which expresses the assumed-known number of particles N in terms of the occupation number $\bar{n}(E, T)$ and density of states $g(E)$:

$$N = \int_0^\infty \bar{n}(E, T) g(E) dE. \quad (7.111)$$

Since we (presumably) know the value of N , we can extract the temperature dependence of $\mu(T)$ from this equation.

To see how this is done, start with the much simpler case of Maxwell–Boltzmann statistics (7.107): the simpler form of $\bar{n}(E, T)$ there allows for an exact calculation of μ for a system of N non-interacting massive distinguishable or massive identical-classical particles.

The form of $\bar{n}(E, T)$ is a simple exponential in (7.107). Next, what is the density of states $g(E)$? Its precise shape varies for different materials and is generally very difficult to calculate in detail; but we can, at least, approximate it for a gas of non-interacting particles. Remember that we are treating the gas as having a set of states that can be occupied by various numbers of its particles. Those states are just the cells of the everyday position–momentum space that the gas can occupy: three spatial dimensions and three momentum dimensions. Counting those states then equates to counting the number of states available to *one* free particle moving in three spatial dimensions, given the constraint of some total energy E .

We saw that counting procedure in Section 2.4: there, the “cell size” in phase space was so small relative to the entire set of cells to be counted, that the states could be counted by evaluating an integral. In that section, we treated a single free massive point particle with energy E , but we calculated $\Omega_{\text{tot}}(E)$ rather than $\Omega(E)$, and so considered all energies in the range 0 to E . The particle moved in three spatial dimensions and stored its energy in three momentum variables; this situation was drawn in Figure 2.6. Each six-dimensional cell of the particle’s phase space then had volume h^3 , with each factor of h coming from a pair of space and momentum variables. Equation (2.32) gave the relevant integrals, with the resulting total number of states given by (2.33):

$$\Omega_{\text{tot}}(E) = \frac{V 4\pi (2mE)^{3/2}}{3h^3}. \quad (7.112)$$

But the particles of the quantum gas will usually have spin. In general, what is called a “spin- s massive particle” has $2s+1$ possible spins,⁷ and so the number of states available to a general massive particle of spin s is $2s+1$ times the value in (7.112), or

$$\Omega_{\text{tot}} = \frac{(2s+1) 4\pi V (2mE)^{3/2}}{3h^3}. \quad (7.113)$$

It follows that for a gas of massive distinguishable or massive identical-classical particles, *whose temperature is high enough that counting discrete cells can be approximated by doing an integral*,

$$g(E) = \Omega'_{\text{tot}}(E) = \underbrace{\frac{(2s+1) 2\pi V (2m)^{3/2}}{h^3}}_{\equiv C} \sqrt{E}, \quad (7.114)$$

where the constant C is defined for shorthand here and later.

Thus, $g(E)$ is proportional to $\sqrt{E}$: a simple function to deal with. Now apply (7.111), taking $\bar{n}(E, T)$ from (7.107), and $g(E)$ from (7.114):

$$N = \int_0^\infty e^{\beta(-E+\mu)} C \sqrt{E} dE = e^{\beta\mu} C \int_0^\infty e^{-\beta E} \sqrt{E} dE. \quad (7.115)$$

Apply a change of variables $x \equiv \sqrt{E}$, to obtain $\sqrt{E} dE = 2x^2 dx$. Then,

$$\int_0^\infty e^{-\beta E} \sqrt{E} dE = 2 \int_0^\infty e^{-\beta x^2} x^2 dx \stackrel{(1.98)}{=} \frac{2}{\beta} \sqrt{\frac{\pi}{\beta}}. \quad (7.116)$$

Equation (7.115) becomes (now with all parameters written explicitly)

$$\frac{N}{V} = e^{\beta\mu} (2s+1) \left(\frac{2\pi m k T}{h^2} \right)^{3/2}. \quad (7.117)$$

Hence,

$$e^{\beta\mu} = \frac{N/V}{2s+1} \left(\frac{h^2}{2\pi m k T} \right)^{3/2}. \quad (7.118)$$

(As a quick check: note that the particle density N/V and temperature T are both intensive variables, and so μ as calculated here is also intensive, as expected.) With this value of μ , (7.107) becomes

⁷ That is, a spin- s massive particle has $2s+1$ possible z components to its spin. For example, a spin- $1/2$ particle has possible z components of $-1/2, 1/2$ (times $\hbar$). A massive spin-1 particle has possible z components of $-1, 0, 1$ (times $\hbar$). For massless particles, the situation is a little different: a spin- s massless particle has just $2s$ possible z components to its spin. A photon has spin 1, giving it the two possible z components of just $-1, 1$ (times $\hbar$). These correspond to its two possible polarisations.

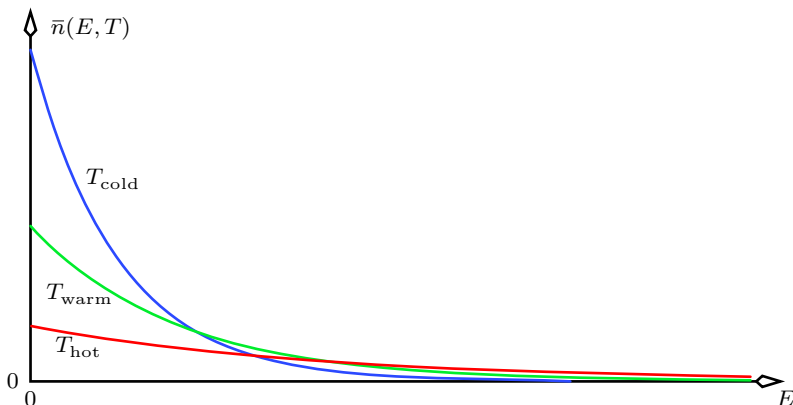


Fig. 7.10 $\bar{n}(E, T)$ versus E for the Maxwell–Boltzmann distribution, from (7.119)

$$\bar{n}(E, T) = \frac{N/V}{2s+1} \left(\frac{h^2}{2\pi m k T} \right)^{3/2} \exp \frac{-E}{kT}. \quad (7.119)$$

Figure 7.10 shows a plot of $\bar{n}(E, T)$ versus E . It has an exponential fall-off with energy that is depressed even further by the chemical potential, which supplies an overall factor of $T^{-3/2}$. Let’s remind ourselves what this plot depicts, by referring to (7.33). The occupation number $\bar{n}(E, T)$ is the mean number of particles per state, where that state has energy E . That is, each of these particles has energy E . The number of states per unit energy interval is given by $g(E)$ in (7.114).

This calculation of $\bar{n}(E, T)$ for Maxwell–Boltzmann statistics (7.107) was easy enough, because $e^{\beta\mu}$ factored out of (7.107), enabling it to factor out of (7.115), and thus to be expressed in terms of known parameters in (7.118). As explained in the next section, the same procedure gives $\mu(T)$ and $\bar{n}(E, T)$ for fermions and bosons—but there, no factoring out occurs, and so the calculations must be done numerically.

On a final side note: recalling the expression for half the mean de Broglie wavelength $^{1/2} \langle \lambda \rangle$ in (7.89), we can write (7.119) as

$$\bar{n}(E, T) = \frac{\langle \lambda \rangle^3}{8(2s+1) V/N} \exp \frac{-E}{kT}. \quad (7.120)$$

Compare this with (7.72): we again see the connection of the de Broglie wavelength to the classical volume occupied by the particles.

7.7.1 Calculating $\mu(T)$ and $\bar{n}(E, T)$ for Fermions

With the simpler Maxwell–Boltzmann case above as a guide, return to (7.111), but now with the fermion occupation number $\bar{n}(E, T)$ from (7.106):

$$\bar{n}(E, T) = \frac{1}{\exp \frac{E - \mu(T)}{kT} + 1}. \quad (7.121)$$

Again, we use (7.114) for the density of states: $g(E) = C\sqrt{E}$. For a numerical calculation, we examine a system containing $N = 1000$ fermions. The constant C is not relevant here, so suppose the fermions' masses are such that $C = 1$ unit of whatever system of units we prefer to use. (As long as the two energies E and μ are measured in the corresponding units—such as joules for SI—the following argument is unchanged.) The “normalisation equation” (7.111) then becomes

$$1000 = \int_0^\infty \frac{\sqrt{E} \, dE}{\exp \frac{E - \mu(T)}{kT} + 1}. \quad (7.122)$$

Given some value of T , we must solve (7.122) for $\mu(T)$.

Simplest is the limit $T \rightarrow 0$: here, the integrand of (7.122) is $\sqrt{E}$ for $E < \mu(0)$ and zero for $E > \mu(0)$. This greatly simplifies the equation to

$$1000 = \int_0^{\mu(0)} \sqrt{E} \, dE = \frac{2}{3} \mu(0)^{3/2}. \quad (7.123)$$

It follows that

$$\mu(0) = 1500^{2/3} \simeq 131.04. \quad (7.124)$$

Similarly, calculate μ for $kT = 1$ by starting with

$$1000 = \int_0^\infty \frac{\sqrt{E} \, dE}{\exp \frac{E - \mu}{1} + 1}. \quad (7.125)$$

This can be solved numerically, to obtain

$$\mu(kT = 1) \simeq 131.03. \quad (7.126)$$

Next, for $kT = 5$:

$$1000 = \int_0^\infty \frac{\sqrt{E} \, dE}{\exp \frac{E - \mu}{5} + 1}, \quad \text{producing } \mu(kT = 5) \simeq 130.88. \quad (7.127)$$

These and two more values are listed in Table 7.2. The value of $\mu(kT)$ remains very close to $\mu(0) \simeq 131.04$ only when $kT \ll \mu(0)$. At much higher values of kT , the value of μ drops markedly.

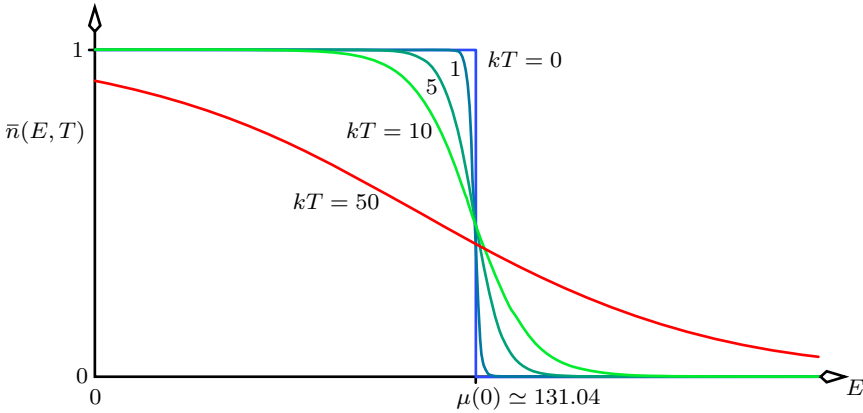


Fig. 7.11 $\bar{n}(E, T)$ versus E for the Fermi–Dirac distribution, from (7.121). The values of temperature and μ are taken from Table 7.2. The graph’s shape at $T = 0$ is that of a step function. Its corners become increasingly rounded as T increases. Note that the blue–green curves all intersect each other at about $\bar{n} = 1/2$, but the red curve does not: only for “low” temperatures [$kT \ll \mu(0)$] is there an approximate symmetry about $E = \mu(0)$

The above procedure now helps us plot $\bar{n}(E, T)$ as a function of energy E per state for various values of T . For each chosen value of T (really kT), we calculate the relevant $\mu(kT)$ as we did immediately above, and then simply plot $\bar{n}(E, T)$ versus E using (7.121). A set of such plots is shown in Figure 7.11. The main character of fermion statistics is a steep drop-off in $\bar{n}(E, T)$ occurring around $E = \mu(0)$. At $T = 0$, the graph becomes a step function, with its step precisely at $E = \mu(0)$. This value $\mu(0)$ appears in various calculations frequently enough to deserve its own name: it is the system’s *Fermi energy* E_F :

$$E_F \equiv \mu(0). \quad (7.128)$$

You will sometimes find $\mu(T)$ called the Fermi energy. But $\mu(T)$ already has a name: it is the chemical potential. In contrast, $\mu(0)$ is special, because it relates to the simple step-function shape taken on by the number density $\bar{n}(E, T)$ at $T = 0$, and this is why it has been given a name of its own.

Figure 7.12 shows the important product $\bar{n}(E, T)g(E)$, whose integral over particle energy E is the total number N of fermions (7.111). In its rightmost plot, we see what is called the *Fermi sea* of fermions: very few

Table 7.2 Values of μ calculated for various values of kT , from the calculations in (7.122)–(7.127)

kT :	0	1	5	10	50
$\mu(kT)$:	131.04	131.03	130.88	130.40	112.53

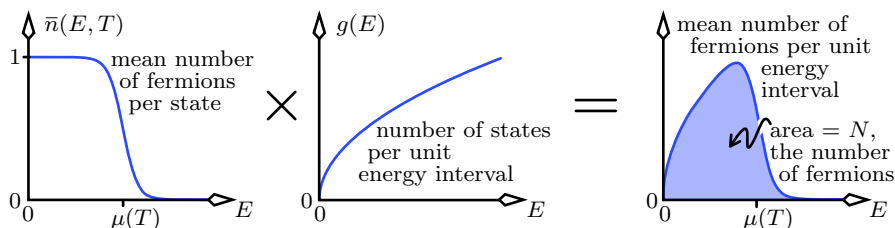


Fig. 7.12 Recall (7.32): $\bar{n}(E, T)$ from (7.121) multiplies $g(E)$ from (7.114), to give the number of fermions per unit energy interval. The rightmost plot shows the (shaded) *Fermi sea* of fermions. To visualise the sea analogy, rotate the plot 90° counter-clockwise, so that the energy axis becomes height above the sea floor—and imagine that the sea bed is sloped

fermions exist with low energy (because there are few states to occupy at low energies), more fermions with higher energies, and suddenly a sharp fall-off occurs at $E = \mu(T)$ —like the surface of an ocean. Perturbing this value of $\mu(T)$ numerically (that is, without changing T) mainly just shifts the fall-off sideways. As the system’s temperature increases, the *slope* of $\bar{n}(E, T)$ ’s fall-off in Figure 7.11 becomes shallower, and, at least for “low” temperatures [$kT \ll E_F$], what it loses from its head is approximately balanced by a gain at its foot. But the area under $\bar{n}(E, T)g(E)$ must remain constant (equal to N). We conclude that at low temperatures, $\bar{n}(E, T)g(E)$ ’s fall-off cannot shift much sideways with a change in temperature. In other words, the value of $\mu(T)$ can have only very little temperature dependence when $kT \ll E_F$.

Thus, provided we work in this “low-temperature” regime of $kT \ll E_F$, it will be valid to write $\mu(T) \simeq E_F$. We’ll find, in Chapter 8, that the Fermi energy of the conduction electrons in copper is $E_F \approx k \times 81,000$ K. It follows that the assumption “ $kT \ll E_F$ ” equates to working only with temperatures below 81,000 K. This, of course, is sufficient for *any* study of the quantum mechanics of copper’s conduction electrons. Hence, we can always replace $\mu(T)$ with the constant E_F for these electrons. This will simplify the conduction calculations of Chapter 8 tremendously.

With these thoughts in mind, replace (7.121) with the excellent approximation

$$\bar{n}(E, T) \simeq \frac{1}{\exp \frac{E - E_F}{kT} + 1}. \quad (7.129)$$

Let’s repeat (7.124)’s calculation of the Fermi energy $E_F = \mu(0)$, but now for a general N and C . Recall that $\bar{n}(E, T)$ becomes a step function at $T = 0$:

$$\bar{n}(E, 0) = \begin{cases} 1 & E \leq E_F, \\ 0 & E > E_F. \end{cases} \quad (7.130)$$

This very simple form of $\bar{n}(E, 0)$, along with $g(E) = C\sqrt{E}$ in (7.114), allows us to write (7.111) as

$$N = \int_0^\infty \bar{n}(E, 0) g(E) dE = \int_0^{E_F} C E^{1/2} dE = \frac{2C}{3} E_F^{3/2}. \quad (7.131)$$

It follows that the Fermi energy is

$$E_F = \left(\frac{3N}{2C} \right)^{2/3} \stackrel{(7.114)}{=} \left(\frac{3Nh^3}{2(2s+1)2\pi V(2m)^{3/2}} \right)^{2/3} = \frac{h^2}{2m} \left(\frac{3N}{4\pi(2s+1)V} \right)^{2/3}. \quad (7.132)$$

Of course, this returns the result in (7.124) for the choice we made there of $N = 1000$ and $C = 1$. We'll use (7.132) in Chapter 8's study of conduction electrons.

7.7.2 Calculating $\mu(T)$ and $\bar{n}(E, T)$ for Bosons

Producing a plot of $\bar{n}(E, T)$ versus E for a gas of massive bosons is a similar exercise to that which we have just done for fermions. We start with $N = 1000$ bosons in (7.111), insert the boson occupation number $\bar{n}(E, T)$ from (7.106) and the density of states $g(E) = C\sqrt{E}$ from (7.114), and solve for μ at various values of kT .

But, unlike the fermion case, attempting to simplify the analysis by setting C to 1 in (7.114) is simply inconsistent with the existence of 1000 bosons in (7.111). That is a statement about boson properties, in that the resulting parameters such as N/V that are present in C are inconsistent with an assemblage of 1000 bosons for this value of C . Instead, set $C = 1000$: this value is quite consistent with the existence of 1000 bosons. Equation (7.111) becomes

$$1000 = \int_0^\infty \frac{1000\sqrt{E} dE}{\exp \frac{E-\mu(T)}{kT} - 1}. \quad (7.133)$$

For now, we avoid the problematic case of $T = 0$: we'll see why, when studying liquid helium in Section 7.8. Begin with $kT = 1$ (in a consistent set of units, as for fermions above) and solve

$$1000 = \int_0^\infty \frac{1000\sqrt{E} dE}{\exp \frac{E-\mu}{1} - 1}. \quad (7.134)$$

A numerical solution is

$$\mu(kT = 1) \simeq -0.285. \quad (7.135)$$

Next, for $kT = 2$:

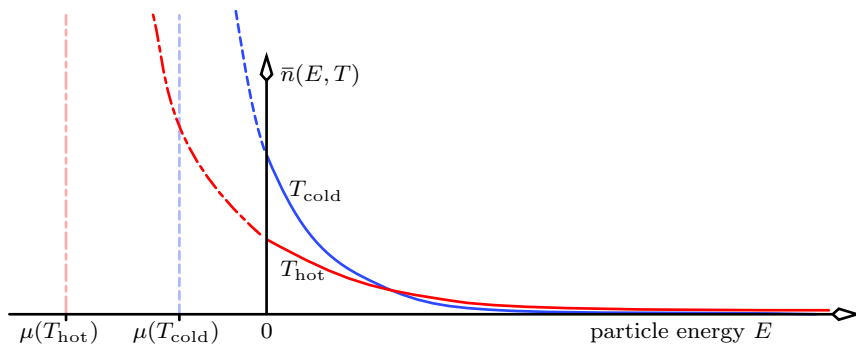


Fig. 7.13 $\bar{n}(E, T)$ versus E for the Bose–Einstein distribution for two values of temperature. Only the solid curves (where E is positive) are relevant, but their analytical expressions are also drawn for negative values of E to aid in picturing how each curve relates to its value of μ . Each curve asymptotes to one of the dashed vertical lines drawn at energy values of $\mu(T_{\text{hot}})$ and $\mu(T_{\text{cold}})$. Note that $\bar{n}(E, T)$ in (7.106) is negative for $E < \mu$ (not drawn here, but the two branches of $\bar{n}(E, T)$ resemble the two branches of the hyperbola “ $y = 1/x$ ”). It follows that if μ were positive, we would obtain negative values of $\bar{n}(E, T)$ for some positive values of energy E , which would make no sense. We conclude that μ must always be negative

$$1000 = \int_0^{\infty} \frac{1000\sqrt{E} dE}{\exp \frac{E-\mu}{2} - 1}, \quad \text{producing } \mu(kT = 2) \simeq -2.122. \quad (7.136)$$

Similarly, $\mu(kT = 3) \simeq -4.812$.

Analogously to the fermion case in Figure 7.11, the above values of μ enable $\bar{n}(E, T)$ to be plotted as a function of E in Figure 7.13 for, say, two representative temperatures T_{cold} and T_{hot} . Each plot of $\bar{n}(E, T)$ resembles the two-armed hyperbola “ $y = 1/x$ ”, but we have drawn only the part where $\bar{n}(E, T)$ is positive. $\bar{n}(E, T)$ asymptotes to positive infinity as $E \rightarrow \mu$ from the right, and to negative infinity as $E \rightarrow \mu$ from the left. Clearly, if μ were positive, negative values of the occupation number $\bar{n}$ would result for energies less than μ ; but these negative values have no meaning. It follows that μ must always be negative for a system with a fixed number of bosons. And when that system is heated, its chemical potential μ must decrease (move toward $-\infty$), since this rearranges the populations of states according to the Boltzmann distribution, which demands that higher temperatures drive up the populations of higher-energy states at the expense of a depletion in populations of lower-energy states. We see this occupation of higher energy levels occurring in Figure 7.13, by comparing the “ T_{hot} ” curve with the “ T_{cold} ” curve (for $E \geq 0$).

Mimicking Figure 7.12, we might attempt to plot $\bar{n}(E, T) g(E)$ for bosons in Figure 7.14. The result is almost correct; but in the next section, we’ll see that to be more correct, we must include what turns out to be a well-populated ground state separately. This state is lost in the continuum ap-

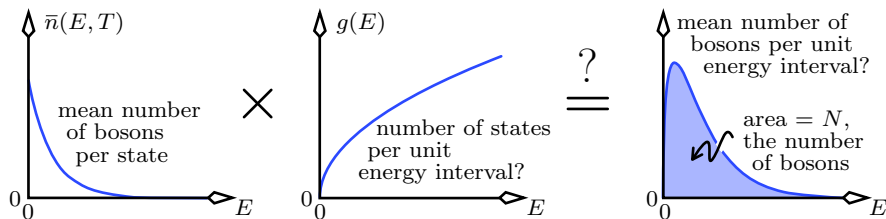


Fig. 7.14 An attempt at forming the product $\bar{n}(E, T) g(E)$ for bosons, mimicking Figure 7.12. In fact, this is not quite right: Section 7.8 will show that our formalism has led to the all-important ground state being omitted here

proximation of $g(E)$ in (7.114). Other states are lost too; but the ground state is the important one, as bosons tend to pile into it as the temperature goes to zero.

7.8 Low-Temperature Bosons and Liquid Helium

The approach to quantum statistics that centres on the occupation number $\bar{n}(E, T)$, from (7.106), and the density of states, $g(E)$, is complicated by the fact that to render the maths tractable, we have assumed the density of states $g(E)$ to be continuous over energy, whereas the quantised nature of the states says that $g(E)$ is *not* continuous. This assumption doesn't lead to any problems for fermions. That's because at most one fermion can occupy a given state, and hence we make negligible error when a continuous $g(E)$ leads to a real, physical state being omitted from an integral over energy.

On the other hand, things are different for low-temperature bosons. Any number of bosons can occupy a state, and thus if a state is accidentally omitted from consideration, a large number of bosons might remain unaccounted for. We can demonstrate this breakdown of the continuous- $g(E)$ assumption for an ideal gas of a fixed number of bosons in the following way.

Begin by revisiting the calculation in Section 7.7, to analyse the dependence of μ on T for a boson gas with fixed N . Do this by applying (7.111) to massive bosons, using $g(E)$ from (7.114):

$$N = C \int_0^\infty \frac{\sqrt{E} dE}{e^{\beta(E-\mu(T))} - 1}. \quad (7.137)$$

Now focus on the integrand of (7.137):

$$f(E) \equiv \frac{\sqrt{E}}{e^{\beta(E-\mu)} - 1}. \quad (7.138)$$

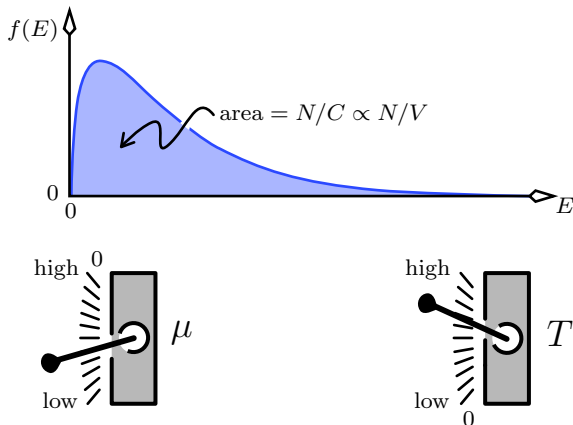


Fig. 7.15 Two levers that control the amount of μ and T present in $f(E)$ via (7.138). The area under the curve is calculated in (7.144). Note the differing positions of the zero values on the two sliders

Suppose that μ and T are independent variables. We will first show that increasing one of either μ or T while holding the other fixed has the effect of increasing $f(E)$ for all energies E . Imagine a device that plots $f(E)$ versus E for any positions of two levers, as shown in Figure 7.15. The movement of a lever is really a series of infinitesimal nudges, meaning that it's sufficient to ask what happens to $f(E)$ when one of μ or T is increased infinitesimally. Consider increasing μ by $d\mu$ at fixed T : the infinitesimal increase in $f(E)$ is

$$df = \frac{\partial f}{\partial \mu} d\mu = \frac{\sqrt{E} e^{\beta(E-\mu)} \beta d\mu}{(e^{\beta(E-\mu)} - 1)^2} = \text{positive number} \times d\mu. \quad (7.139)$$

The signs of df and $d\mu$ are the same. So, shifting the μ lever toward higher values of μ (thus $d\mu > 0$) results in an increase $df > 0$. This means the function $f(E)$ increases for *all* E . Similarly, consider shifting the temperature lever: increasing T by dT at fixed μ . Now the infinitesimal increase in $f(E)$ is

$$df = \frac{\partial f}{\partial T} dT = \frac{\sqrt{E} e^{\beta(E-\mu)} \frac{E-\mu}{kT^2}}{(e^{\beta(E-\mu)} - 1)^2} dT. \quad (7.140)$$

Is this positive or negative? The only unknown quantity here is $E - \mu$. But consider that the occupation number $\bar{n}$ must be greater than or equal to zero:

$$\bar{n} = \frac{1}{e^{\beta(E-\mu)} - 1} \geq 0, \quad \text{in which case } e^{\beta(E-\mu)} - 1 \geq 0. \quad (7.141)$$

Thus,

$$e^{\beta(E-\mu)} \geq 1, \quad \text{and so } \beta(E-\mu) \geq 0, \quad \text{and therefore } E - \mu \geq 0. \quad (7.142)$$

(Note that this implies $\mu \leq 0$ for an ideal gas of bosons, because μ is less than or equal to all possible values of energy $E \geq 0$. We saw the same thing in the analysis around Figure 7.13.) It's now evident that (7.140) has the form

$$df = \text{positive number} \times dT. \quad (7.143)$$

Hence, the effect of the temperature lever is similar to the effect of the μ lever: pushing either lever toward a higher value increases the value of $f(E)$ for all E .

Next, recall, from (7.137) and (7.138), that

$$\int_0^\infty f(E) dE = \frac{N}{C} \frac{(7.114)}{(2s+1) 2\pi V (2m)^{3/2}}. \quad (7.144)$$

From this, it's clear that *if*—for whatever reason—the particle density N/V of the massive bosons is constant, the area under $f(E)$ -versus- E will also be constant, irrespective of the lever positions in Figure 7.15.

Suppose now that we cool the system. Decreasing T (shifting the T lever down in Figure 7.15) makes $f(E)$ decrease for all E ; so, to keep the area N/C under the curve fixed, μ must increase: we must shift the μ lever up in the figure. This shows that μ is indeed a function of temperature T .

But, as T decreases, the increasing μ (which we know is negative) can never increase so much as to become positive. We will show that there is a critical temperature $T_c > 0$ at which μ reaches zero: the μ lever hits the top of its travel—but the T lever still has some downward movement available. For temperatures below T_c , our model must break down in some way, because μ cannot increase further. Naturally, this critical temperature T_c is found by setting $\mu = 0$ in (7.137), with $\beta_c \equiv 1/(kT_c)$:

$$N = C \int_0^\infty \frac{\sqrt{E} dE}{e^{\beta_c E} - 1}. \quad (7.145)$$

Solve this for T_c by changing variables to $x \equiv \sqrt{\beta_c E}$. With $dx/dE = \beta_c/(2x)$, we have

$$\sqrt{E} dE = 2x^2 dx (kT_c)^{3/2}. \quad (7.146)$$

This converts (7.145) to

$$N = (kT_c)^{3/2} C \underbrace{\int_0^\infty \frac{2x^2 dx}{e^{x^2} - 1}}_{\simeq 2.31516}, \quad (7.147)$$

where the last integral has been evaluated numerically.⁸ Hence,

⁸ It is sometimes pointed out that the value of this integral is $1/2\sqrt{\pi}\zeta(3/2)$, where ζ is the illustrious Riemann zeta function. Although that is true, it would be misleading to state that this means the zeta function has any relation to low-temperature bosons. By

$$(kT_c)^{3/2} = \frac{N}{2.31516 C} \stackrel{(7.144)}{=} \frac{Nh^3}{2.31516 (2s+1) 2\pi V (2m)^{3/2}}. \quad (7.148)$$

Convert the N/V in this expression to quantities that are more easily determined, via

$$N = \frac{\text{total mass of gas}}{\text{mass per boson}} = \frac{\rho V}{m}, \quad (7.149)$$

where ρ is the mass density of the gas. The critical temperature is then

$$T_c = \frac{1}{k} \left[\frac{\rho}{2.31516 m (2s+1) 2\pi} \right]^{2/3} \frac{h^2}{2m}. \quad (7.150)$$

This critical temperature will be high—and thus easy to observe in the lab—when the bosons have a high mass density ρ and a low mass m .

At this critical temperature, μ hits its ceiling value of zero, and our model must break down for colder temperatures. What has gone wrong? Remember that at very low temperatures, Figure 7.13 shows that the number of particles in a state can be very high (the asymptote of the curve shifts toward $E = 0$). We must start to be more careful to account for the occupation of what is really a *discrete* set of states by gregarious bosons. But we used a continuous approximation of $g(E)$ in (7.137); specifically, $g(E) = C\sqrt{E}$ from (7.114). This continuum form of $g(E)$ is roughly correct at high temperatures, where classical mechanics applies; but it is simply incorrect at low temperatures, where the set of bosons becomes fully quantum mechanical. Contrast this with the fermion case: since, at most, only one fermion can occupy a state, we make almost no error by using $g(E) = C\sqrt{E}$ at low temperatures for fermions—and we'll do just that in the next chapter.

A way around this problem of modelling a discrete set of states as a continuum is to deal with the number of particles in the ground state separately from the numbers in excited states, since these excited states can be adequately approximated by (7.137). Hence, we write

$$\underbrace{N}_{\substack{\text{total number of} \\ \text{bosons (fixed)}}} = \underbrace{N_0(T)}_{\substack{\text{number in} \\ \text{ground state}}} + \underbrace{N_{\text{ex}}(T)}_{\substack{\text{number in} \\ \text{excited states}}}, \quad (7.151)$$

where

$$N_{\text{ex}}(T) = C \int_0^\infty \frac{\sqrt{E} dE}{e^{\beta(E-\mu)} - 1}. \quad (7.152)$$

The analysis that was depicted in Figure 7.15 says that $N_{\text{ex}}(T)$ attains a maximum for $\mu = 0$. We have already worked out this maximum: it is N

the same token, $\zeta(2) = \pi^2/6$, which means that the circumference of a circle divided by its diameter equals $\sqrt{6}\zeta(2)$; but that does not imply that the zeta function must be important to any discussion of circles. Three or four isolated values of the zeta function appear in physics, but the function itself plays no part in any current formalism.

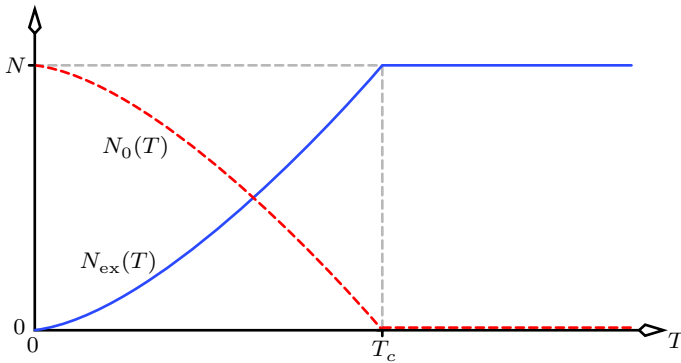


Fig. 7.16 The predicted population N_0 of massive bosons in the ground state, and the predicted population N_{ex} in all excited states, as functions of temperature T from (7.155). Above the critical temperature T_c , essentially all the bosons occupy excited states. As the temperature drops below T_c , the bosons begin to drop into the ground state: this is Bose–Einstein condensation

in (7.147):

$$N_{\text{ex}}(T_c) = N = 2.31516 C(kT_c)^{3/2} \equiv \alpha T_c^{3/2}. \quad (7.153)$$

At higher temperatures, no more bosons are available to be excited, and so $N_{\text{ex}}(T \geq T_c) = N$. What about lower temperatures? The “3/2” power in (7.153) suggests that the same power law might apply at temperatures below T_c , and this is indeed what we postulate:

$$N_{\text{ex}}(T \leq T_c) = \alpha T^{3/2} \stackrel{(7.153)}{=} \frac{N}{T_c^{3/2}} T^{3/2} = N \left(\frac{T}{T_c} \right)^{3/2}. \quad (7.154)$$

The number in the ground state is then

$$N_0(T) = \begin{cases} N - N_{\text{ex}}(T \leq T_c) = N \left[1 - \left(\frac{T}{T_c} \right)^{3/2} \right] & (T \leq T_c), \\ 0 & (T > T_c). \end{cases} \quad (7.155)$$

These populations are shown as functions of temperature in Figure 7.16. There, we see the bosons beginning to “condense” into the ground state as the temperature drops below T_c . This is called *Bose–Einstein condensation*. Realise that the relevant theory (such as the use of $g(E) = C\sqrt{E}$) was derived for an *ideal gas*. But does Bose–Einstein condensation occur with the non-ideal gases that exist in Nature?

Recalling (7.150)’s result that high density and low mass produce a high T_c (which is easier to observe experimentally), the classic candidate for showing Bose–Einstein condensation in the lab is helium-4. Its spin-zero bosonic atoms display exceptionally weak interatomic forces, which allow it to resist liquefying as it cools. As a result, it can be treated as an ideal gas down to

very low temperatures. It liquefies at the uncommonly low temperature of 4.2 K, at which point the newly condensed liquid is called *helium-I*. Helium-I displays no remarkable properties, and can essentially be treated as an ideal gas.

As helium-I is cooled below a transition temperature of 2.17 K, it begins to behave in a remarkably strange way, and thus acquires the name *helium-II*. Helium-II *behaves* as though it were a mixture of two interpenetrating fluids: a “normal fluid” with properties just like any other fluid, and a “superfluid”, which displays some remarkable properties indeed. The superfluid appears to have zero entropy, presents no resistance to flow, has zero viscosity, and will not support any turbulence. Helium-II will leak through the smallest hole, as well as climbing the walls of its container and trickling down the outside.

In what is experimentally a highly successful model of helium-II, when the temperature drops below 2.17 K, the proportion of the superfluid component rises from 0% at 2.17 K to 100% at 0 K. This two-fluid model of helium-II is *phenomenological* only: that is, there is no implication that helium-II really *is* composed of two such fluids.

How does this superfluid behaviour relate to Bose–Einstein condensation? We can get a clue to a match with the above bosonic theory by calculating the critical temperature for helium from (7.150). Remember the assumption of a constant N/V that was stated just after (7.144): this assumption led to the prediction of Bose–Einstein condensation. The mass density is $\varrho = mN/V$, and it turns out that as helium-I is cooled to 2.17 K, ϱ increases to about 140 kg/m³. At still lower temperatures, ϱ stays fairly constant at about 120 kg/m³. So, N/V is approximately constant. The relevant parameters in (7.150) are

$$s = 0, \quad \varrho \simeq 120 \text{ kg/m}^3, \quad m \simeq \frac{0.004 \text{ kg}}{6.022 \times 10^{23}}, \quad (7.156)$$

in which case

$$T_c \simeq \frac{1}{1.381 \times 10^{-23}} \left[\frac{120 \times 6.022 \times 10^{23}}{2.31516 \times 0.004 \times 1 \times 2\pi} \right]^{2/3} \times \frac{(6.626 \times 10^{-34})^2 \times 6.022 \times 10^{23}}{2 \times 0.004} \text{ K} \\ \simeq 2.8 \text{ K}. \quad (7.157)$$

This prediction of the onset of Bose–Einstein condensation at $T_c \simeq 2.8$ K matches well the observed onset of non-classical behaviour at 2.17 K in liquid helium.

But does this agreement imply that the superfluid is the sought-after Bose–Einstein condensate? Unfortunately, the situation is not that straightforward. Neutron-scattering experiments reveal that a set of zero-momentum atoms appears to be present in helium-II. This set begins to form around T_c , and rises to about 14% of the helium-II at $T = 0$. This contrasts with the super-

fluid’s 100% proportion of helium-II at zero temperature. There is simply no straightforward connection between the superfluid present in the real non-ideal-gas helium and the Bose–Einstein condensate predicted for an ideal gas. The zero-momentum atoms in helium-II do not appear to be a separate substance; instead, the state of having zero momentum is presumably being “passed around” different atoms. But even to speak of separate atoms here is dubious, in the light of the de Broglie discussion of helium in Section 7.5. There, we showed that the de Broglie wavelength of its atoms is longer than its interatomic spacing, meaning that the atoms have lost their individuality at these low temperatures.

Decades after the discovery of helium-II, this uneasy relationship between the superfluid and the condensate remains the subject’s status quo.

7.9 Excursus: Particle Statistics from Counting Configurations

Given the very peaked nature of the probability densities that we routinely encounter in statistical mechanics, it’s reasonable to assume that the *mean* number of particles in a quantum state of energy E in (7.106) or (7.107) is approximated extremely well by the *most likely* number of particles in that state.⁹ It turns out that this most likely number can be found from a pure counting argument that requires neither a concept of entropy nor the Boltzmann distribution. We do that in this section, using the analogy of a set of balls being placed into pots that sit on the shelves of a tall book case. Our ultimate aim is to reproduce (7.106) for fermions and bosons, and (7.107) for distinguishable particles.

The book case represents a single quantum system, shown in Figure 7.17. Each of its shelves corresponds to an energy level of that system: you can think of the gravitational potential energy on shelf i as the energy E_i per particle at level i of the system. So, the higher the shelf, the larger the value of E_i . Shelf i contains Ω_i coloured pots, with each pot corresponding to one of the system’s Ω_i quantum states at energy E_i . The pots on any particular shelf are coloured differently amongst themselves; this reminds us that even if the particles in the system are indistinguishable, the quantum states that they occupy are always *distinguishable*, because each quantum state is specified by a unique set of quantum numbers.

Imagine that we have a fixed number N of balls that we must place in the pots. We are free to place them wherever we like, on any shelf and in any pot, subject to two constraints:

- we place all N of the balls, leaving none out and adding none, and

⁹ Remember that “most likely” corresponds to the peak of the probability distribution.

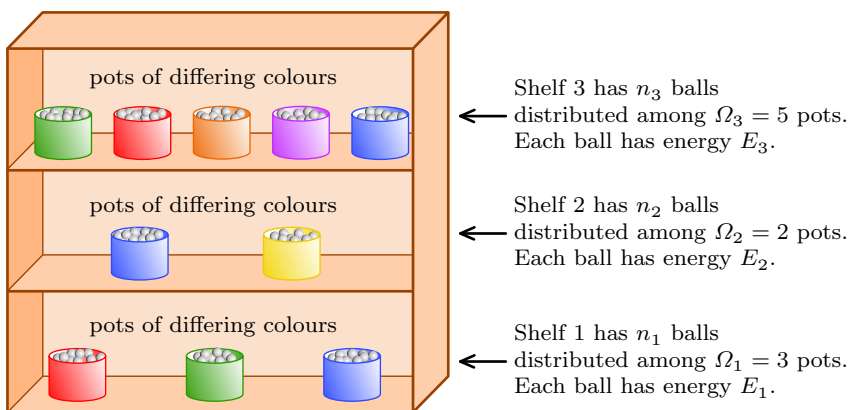


Fig. 7.17 Envisage a generic quantum system as a book case, where each shelf represents an energy level of the system: higher shelves mean higher energies. On each shelf sits one or more pots, each pot with its own colour for that shelf, and with each pot representing a quantum state at that energy. Any one pot on shelf 2 can be painted, say, blue: this doesn't conflict with the blue pot on shelf 1, because it's clear that the pots are on different shelves

- we have a set amount E of energy available for the task, which we must spend exactly. The final arrangement of the balls will have this total energy E . (This will all be gravitational potential energy if we use that gravitational analogy.)

Suppose, initially, that no pots are present on the shelves, and we must simply place balls on the shelves. For this particular argument, let the balls be as good as identical: we bought them in a sports shop, and they all look alike. (Or, if they don't all look alike, we take no notice of their individual markings.) When we have only a few balls to place and few shelves available, we have little choice as to how to place the balls. Consider exactly three shelves, with energies $E_1 = 1$ joule, $E_2 = 2$ joules, $E_3 = 3$ joules. We are required to place $N = 3$ balls and must spend a total of $E = 6$ joules of energy. This can be done in only two ways, as shown in Figure 7.18. How do we know that? If shelf i is to have n_i balls, the above two constraints produce two equations to be solved for whole numbers n_1, n_2, n_3 :

$$n_1 + n_2 + n_3 = 3 \text{ (balls)}, \quad n_1 + 2n_2 + 3n_3 = 6 \text{ (joules)}. \quad (7.158)$$

These simultaneous equations have two whole-number solutions:

$$(n_1, n_2, n_3) = (1, 1, 1) \text{ and } (0, 3, 0). \quad (7.159)$$

The corresponding ball placements are shown in Figure 7.18. The greater the number of balls and the higher the available energy, the more configurations are possible. For example, suppose that on the same three shelves, we must

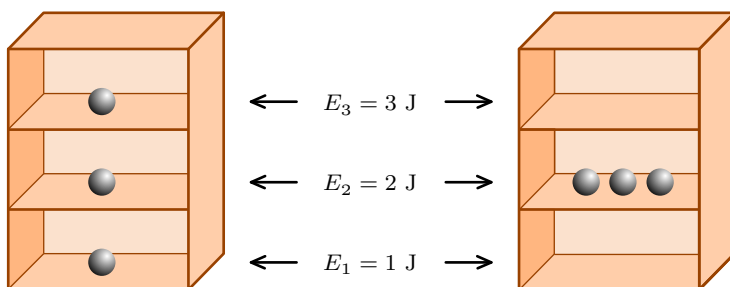


Fig. 7.18 There are only two possible ways to place 3 balls onto the three available shelves (without pots), so as to give a total energy of 6 joules

place $N = 12$ balls with a total of $E = 22$ joules of energy. This can be done in *six* ways, shown in Figure 7.19. A ball represents a particle, of course, and when $N = 10^{23}$ balls and there are a great many shelves, the total number of allowed configurations rises dramatically.

Next, include the pots on the shelves, which gives us an additional choice of which pot to place each ball in. Whether a given ball is placed in one pot or in its neighbour on the same shelf has no bearing on the two constraints of total ball number and total energy, but these different choices of pot are certainly different configurations of the ball layout. Our task is then the following. We have N balls to place, and a total energy E to spend, and we must count the number of configurations in which we can arrange the balls. We take a snapshot of each configuration, and when we have exhausted all possible

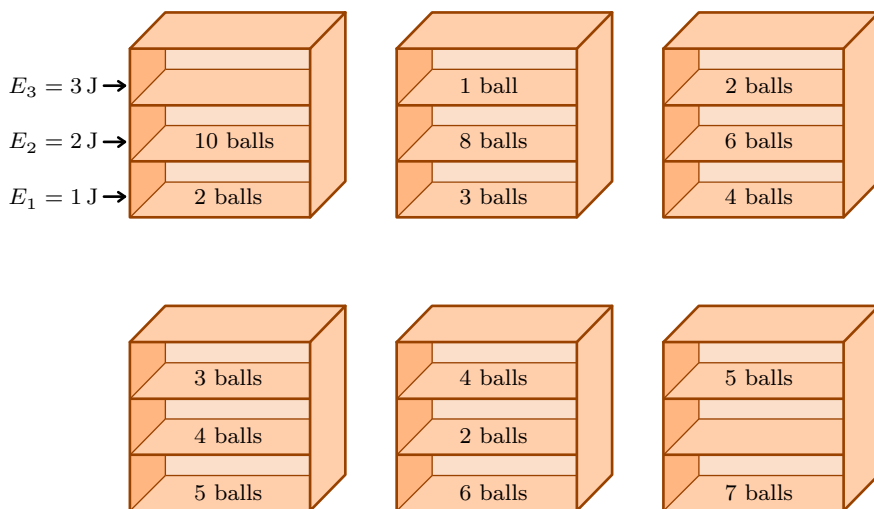


Fig. 7.19 There are six ways to place 12 balls onto the three available shelves without pots, to give a total energy of 22 joules

configurations, we analyse the resulting set of photographs to find which set of $n_1, n_2, n_3, \dots$ is the most numerous. If we now take the viewpoint that the distribution of particles in a real system follows the fundamental postulate of statistical mechanics (which says that any configuration is just as likely as the next: see Section 2.1), then this set of numbers $n_1, n_2, n_3, \dots$ will be the most likely distribution of particles across a set of energy levels.

In the above examples in Figures 7.18 and 7.19, we applied the constraints of “ N balls, with total energy E ”, as in (7.158). This expressed, say, n_1 and n_2 as functions of n_3 , and then we listed the resulting triplets of whole numbers n_1, n_2, n_3 , and depicted these in the two figures. But, when the numbers of particles are large, making such a list of allowed possibilities becomes completely impractical. Instead, we’ll count the number of allowed configurations for a *given set* of $n_1, n_2, n_3, \dots$, and then we’ll use calculus to vary $n_1, n_2, n_3, \dots$ according to the “ N balls, with total energy E ” constraint. This will produce the set with the maximal number of configurations, which is really ultimately what we want.

Our plan is to calculate the distribution of particle numbers in energy for distinguishable classical particles, then fermions and bosons. We expect to obtain the exponential dependence on energy in (7.107) for classical particles, and (7.106) for fermions and bosons. We cannot hope to obtain those two equations *exactly*, because our current picture of balls on shelves allows no concept of temperature.

Counting Configurations of Distinguishable Particles

First, focus on placing *distinguishable* balls onto shelf i . The number of ways that n_i such balls can be placed into Ω_i pots (remember that the pots are *always* distinguishable, even for fermions and bosons) equals the number of ways to paint Ω_i different colours onto n_i distinguishable balls. (To prove this, imagine that each pot is filled with paint of that pot’s colour: when a ball goes into that pot, it is coated in that pot’s colour.) Distinguishable balls can be numbered, so we could colour ball 1 with any of Ω_i colours, and the same for ball 2, and so on. Hence, the total number of ways that the n_i distinguishable balls can be placed into the Ω_i pots must be

$$\underbrace{\Omega_i}_{\text{number of ways of painting ball 1}} \times \underbrace{\Omega_i}_{\text{number of ways of painting ball 2}} \times \underbrace{\Omega_i}_{\text{number of ways of painting ball 3}} \times \cdots = \Omega_i^{n_i}. \quad (7.160)$$

Next, consider *all* of the shelves, with the following set of numbers of distinguishable balls to place on each: $n_1 = 3$ balls on shelf 1, $n_2 = 2$ balls on shelf 2, $n_3 = 2$ balls on shelf 3, and so on. (These numbers are immaterial, but serve as a concrete example for illustration.) The number of ways to put, say, balls 1, 2, 4 into the Ω_1 pots on shelf 1, AND balls 5, 6 into the Ω_2 pots

on shelf 2, AND balls 7, 9 into the Ω_3 pots on shelf 3, etc., is then

$$\Omega_1^3 \Omega_2^2 \Omega_3^2 \dots \quad (7.161)$$

This is just one set of configurations that involve the numbered balls in the way described (balls 1, 2, 4 on shelf 1, etc.). But, similarly, the number of ways to put balls 1, 3, 52 into the Ω_1 pots, AND balls 7, 13 into the Ω_2 pots, AND balls 102, 400 into the Ω_3 pots, etc., is also given by (7.161). So, the total number of ways to place the distinguishable balls must be $\Omega_1^3 \Omega_2^2 \Omega_3^2$ times the number of orderings of ball numbers that we can write. Here are two such orderings:

$$\begin{array}{llll} \text{ordering 1:} & \underbrace{1, 2, 4}_{\text{shelf 1}} & \underbrace{5, 6}_{\text{shelf 2}} & \underbrace{7, 9}_{\text{shelf 3}} \dots \\ \text{ordering 2:} & \underbrace{1, 3, 52}_{\text{shelf 1}} & \underbrace{7, 13}_{\text{shelf 2}} & \underbrace{102, 400}_{\text{shelf 3}} \dots \end{array} \quad (7.162)$$

Remember that the ordering of the numbers on each shelf in (7.162) has no relevance; for example, “1, 2, 4” on shelf 1 of ordering 1 is the same as “4, 1, 2”.

To count the number of orderings that we could list here is to consider the multinomial distribution of Section 1.2. That is, we will “over-count” the number of orderings by writing all $N!$ permutations of the N numbered balls, and then remember that writing “1, 2, 4” in (7.162) is no different from writing “4, 1, 2”: we wish to list *combinations* there, not *permutations*. Hence, we correct for the over-counting by dividing the total number of permutations, $N!$, by (in this case) $3!2!2!$. The total number of ways to place the distinguishable balls must then be

$$\Omega_1^3 \Omega_2^2 \Omega_3^2 \dots \times \frac{\left[\begin{array}{l} \text{total number of orderings, treating} \\ \text{e.g. “1, 2, 4” as distinct from “4, 1, 2”} \end{array} \right]}{\text{correction factor for the over-counting}}. \quad (7.163)$$

In the above example, this is

$$\Omega_1^3 \Omega_2^2 \Omega_3^2 \dots \times \frac{N!}{3!2!2!}. \quad (7.164)$$

More generally, replace the above values $(n_1, n_2, n_3, \dots) = (3, 2, 2, \dots)$ with generic numbers of balls $n_1, n_2, n_3, \dots$ per shelf. Finally, we can write that the total number of ways to place N distinguishable balls onto shelves, such that n_i balls end up in Ω_i pots on shelf i , is

$$\Omega_{\text{dist}}(n_1, n_2, \dots, \Omega_1, \Omega_2, \dots) = \frac{N!}{n_1! n_2! n_3! \dots} \Omega_1^{n_1} \Omega_2^{n_2} \Omega_3^{n_3} \dots \quad (7.165)$$

Our next task is to find the values of $n_1, n_2, \dots$ that maximise this number of ways Ω_{dist} of allocating the distinguishable balls to the pots. These values won't necessarily be the *mean* numbers of balls on each shelf, but they will certainly be the *most likely* numbers of balls on each shelf.

As per the discussion just prior to and including (7.158), Ω_{dist} must be maximised subject to the constraints of fixed particle number N and fixed total energy E :

$$\sum_i n_i = N, \quad \sum_i n_i E_i = E. \quad (7.166)$$

The large number of products in (7.165) means that it will be easier (while still equivalent) to maximise $\ln \Omega_{\text{dist}}$ instead of Ω_{dist} . We saw this type of extremisation subject to constraints when analysing the Brandeis dice in Section 5.11. This is the method of Lagrange multipliers, which determines the set $n_1, n_2, \dots$ by maximising (actually extremising) (7.165) subject to the constraints (7.166). Just as we saw in (5.168), the Lagrange approach solves the following equation for each k :

$$\frac{\partial}{\partial n_k} \left[\begin{array}{c} \text{expression} \\ \text{to extremise} \end{array} \right] = \sum_M (\text{multiplier } M) \times \frac{\partial}{\partial n_k} \text{constraint } M. \quad (7.167)$$

The expression to extremise is $\ln \Omega_{\text{dist}}$ from (7.165). Writing the multipliers as α, β , (7.167) becomes

$$\begin{aligned} \frac{\partial}{\partial n_k} \ln \Omega_{\text{dist}} &= \alpha \frac{\partial}{\partial n_k} \sum_i n_i + \beta \frac{\partial}{\partial n_k} \sum_i n_i E_i \\ &= \alpha + \beta E_k. \end{aligned} \quad (7.168)$$

Now write the logarithm of (7.165):

$$\ln \Omega_{\text{dist}} = \ln N! + \sum_i n_i \ln \Omega_i - \ln(n_i!). \quad (7.169)$$

We must place this into the left-hand side of (7.168). To make it tractable, approximate the factorials $n_i!$ using Stirling's rule (as usual!). The calculation is a little tidier if we use the simpler version (1.27) of the rule: $\ln x! \simeq x \ln x - x$. [It must be said, though, that the following results turn out to be unchanged if we use the more correct version (1.23).] Write (7.169) as

$$\ln \Omega_{\text{dist}} \simeq \ln N! + \sum_i n_i \ln \Omega_i - n_i \ln n_i + n_i. \quad (7.170)$$

Using this, the Lagrange-multiplier equation (7.168) becomes

$$\frac{\partial}{\partial n_k} \left(\ln N! + \sum_i n_i \ln \Omega_i - n_i \ln n_i + n_i \right) \simeq \alpha + \beta E_k. \quad (7.171)$$

Remember that N is a constant, and so taking the partial derivative yields¹⁰

$$\ln \Omega_k - \ln n_k - \frac{n_k}{n_k} + 1 \simeq \alpha + \beta E_k. \quad (7.172)$$

Solving this for n_k produces

$$n_k \simeq e^{-\alpha} \Omega_k e^{-\beta E_k}. \quad (7.173)$$

This result should be considered as approximate, since it was derived using Stirling's rule. Aside from that, it says that the most likely number of balls to end up on shelf k is proportional to $\Omega_k e^{-\beta E_k}$. In other words, the most likely number of particles to be found in energy level k with the constraints of fixed total particle number and fixed total energy is $\Omega_k e^{-\beta E_k}$, where Ω_k is the number of states per energy level. This is the Boltzmann expression (5.5) with $\beta = 1/(kT)$, and with α related to the other system parameters, such as pressure. Temperature and these other parameters are physical quantities, whereas the above argument was purely an exercise in counting configurations; so, we cannot hope to arrive at " $\beta = 1/(kT)$ " without injecting more physics into this scenario of pots and balls. But injecting more physics would then either return to thermodynamical ground that we have already covered, or else we would have to (re)define temperature such that $\beta = 1/(kT)$. We won't pursue either path; instead, it's sufficient to observe the key conclusion here: that the exponential dependence on energy in the Maxwell–Boltzmann distribution results from the counting argument that we followed above. This is all really just a repackaging of Jaynes' Brandeis Dice in Section 5.11.

Using the above arguments, we found an expression for the number of configurations of distinguishable particles Ω_{dist} in (7.165), and then maximised it with respect to $n_1, n_2, n_3, \dots$ to produce an expression for each n_i . We'll now do the same for fermions and bosons, to generate analogous quantities Ω_{ferm} in (7.175) ahead, and Ω_{bos} in (7.176). We will then maximise these with respect to $n_1, n_2, n_3, \dots$ to produce expressions for each n_i , and will find that Fermi–Dirac and Bose–Einstein statistics emerge.

Counting Configurations of Fermions

The main new feature required for modelling identical quantum particles as balls on shelves is that the balls are now identical: they do not come “printed” with numbers. This anonymity actually simplifies the discussion.

For the case of fermions, at most one ball can appear in any given pot. Consider shelf 1 with, say, $\Omega_1 = 4$ pots. In how many ways can $n_1 = 3$ identical balls be placed in these pots? Remember that the pots are *always* distin-

¹⁰ Perhaps we should have summed the n_i to give N in (7.171) before differentiating? We could have: it would only shift the value of α by 1, which is immaterial here.

guishable, so we can start by numbering them (or denoting them by colours; but numbering is fine). The following table shows the number of balls in each of the four pots on shelf 1. Each row shows one of the possible configurations:

Pot 1	Pot 2	Pot 3	Pot 4
1	1	1	0
1	1	0	1
1	0	1	1
0	1	1	1

Clearly, only four configurations are possible. We can enumerate them usefully in the following way. For each configuration (each row), write down a pot's number for each ball in that pot. In the first configuration, row 1 in the table above, the pots that are occupied are pots 1, 2, and 3. So, let the numbers (1, 2, 3) label the first row—and their order is irrelevant; (2, 1, 3) would do just as well. In the same way, the other rows are labelled consecutively (1, 2, 4), (1, 3, 4), and (2, 3, 4).

It's apparent that the number of configurations (that is, labels) equals the number of ways of sampling the numbers 1, 2, 3, 4 three at a time, with no regard for order, and with no replacement allowed. That is, we seek the number of ways of sampling Ω_1 numbers, choosing n_1 at a time with no order and no replacement. This is, of course, the number of *combinations* $C_{n_1}^{\Omega_1}$:

$$C_3^4 = \frac{4!}{3!1!} = 4. \quad (7.174)$$

Recall that shelf i has Ω_i pots, with a total of n_i balls. The maximum value that n_i can have occurs when one ball is placed in each pot; thus, this maximum value is $n_i = \Omega_i$. When $n_i > \Omega_i$, the number of balls that can be placed is zero. It follows that if we define $C_{n_i}^{\Omega_i} \equiv 0$ when $n_i > \Omega_i$, we can use this combinatorial notation $C_{n_i}^{\Omega_i}$ for *any* choice of n_i and Ω_i .

For the entire system, the number of possible configurations is the product of the number of configurations for each shelf:

$$\Omega_{\text{ferm}}(n_1, n_2, \dots, \Omega_1, \Omega_2, \dots) = \prod_i C_{n_i}^{\Omega_i}. \quad (7.175)$$

For example, if $n_1 > \Omega_1$, no configuration with n_1 balls on shelf 1 is possible, and so $\Omega_{\text{ferm}} = 0$. This is ensured in (7.175) by the fact that $C_{n_1}^{\Omega_1}$ equals zero.

Counting Configurations of Bosons

Whereas, at most one “fermion ball” can be placed in any given pot, no such limit applies to bosons: any number of “boson balls” can be placed in a single

pot. Just as we did for fermions, consider again shelf 1 with $\Omega_1 = 4$ pots. In how many ways can $n_1 = 3$ identical balls be placed in these pots? The following table shows the number of balls in each pot on shelf 1, with each row showing one possible configuration:

Pot 1	Pot 2	Pot 3	Pot 4
3	0	0	0
2	1	0	0
1	2	0	0
0	3	0	0
0	2	1	0
	$\vdots$		
0	0	0	3

Count the number of configurations (rows) using the same scheme applied above to fermions: for each configuration, write down a pot's number for each ball in that pot. In the first row above, pot 1 contains three balls and all other pots are empty, so label this configuration with $(1, 1, 1)$. The second row has two balls in pot 1 and one ball in pot 2, so label it $(1, 1, 2)$. Continue until the last row, which is labelled $(4, 4, 4)$. The order of the numbers in each label is immaterial: $(1, 1, 2)$ is the same as $(1, 2, 1)$. Without loss of generality, we will drop the subscript "1" from Ω_1 and n_1 in what follows.

Examining the labels $(1, 1, 1)$, $(1, 1, 2)$, $\dots$, $(4, 4, 4)$, it's evident that the number of configurations (labels) equals the number of ways of sampling the numbers 1, 2, 3, 4 three at a time, with no regard for order and *with replacement*. That is, this number of configurations equals the number of ways of sampling Ω numbers, choosing n at a time with no order, and with replacement allowed. Here is one approach to calculating this number of ways.

First, tabulate those alternative labels in rows:

1	1	1
1	1	2
1	1	3
1	1	4
1	2	2
	$\vdots$	
4	4	4

Now rewrite each row in the table above as follows. Draw a cross for each 1 that appears, then draw a vertical divider, then draw a cross for each 2, then another divider, then a cross for each 3, another vertical divider, and then a cross for each 4. For example, $(1, 1, 4)$ becomes

$$\times \times ||| \times$$

There are always $n = 3$ crosses present, because we sampled $n = 3$ numbers. There are always $\Omega = 4$ sets of crosses to be written down in between the vertical dividers (one set for 1s, one set for 2s, one set for 3s, and one set for 4s), and so $\Omega - 1 = 3$ dividers must be present. Summarising, there are always n crosses and $\Omega - 1$ dividers present. This gives a total number of symbols (crosses and dividers) of $n + \Omega - 1 = 6$.

Next, number these crosses and dividers from left to right. In the above case for $(1, 1, 4)$, the crosses appear at positions 1, 2, 6 of the cross/divider symbols: that is, the 1st, 2nd, and 6th symbols are crosses. We have transformed the triplet $(1, 1, 4)$ into the triplet $(1, 2, 6)$, with none of the numbers in the latter triplet being repeated, because they denote positions from left to right. And, because these latter numbers denote positions, they can be re-ordered to, say, $(1, 6, 2)$, while still representing the original sample $(1, 1, 4)$.

Now pause to see what has happened: sampling with replacement has had the effect of generating an $n = 3$ -digit set for each sample, with the digits now being drawn $n = 3$ at a time from the set $\{1, 2, 3, \dots, n + \Omega - 1\}$, with *no* replacement and *no* order. In other words, the number of samples that we are trying to count equals the number of *combinations* of $n + \Omega - 1$ objects taken n at a time. Thus, this number of combinations, $C_n^{n+\Omega-1}$, is the original sought-after number of ways in which n boson balls can be placed into Ω pots on a shelf. Over all shelves, then, the total number of ways of arranging the bosonic balls is the product of this number of ways for each shelf:

$$\Omega_{\text{bos}}(n_1, n_2, \dots, \Omega_1, \Omega_2, \dots) = \prod_i C_{n_i}^{n_i + \Omega_i - 1}. \quad (7.176)$$

The Four Primary Combinatorial Results

Here are the four primary combinatorial results that we have drawn from, using more traditional language. In how many ways can we choose x objects from n objects? The answers are:

	Order matters	Order doesn't matter
With replacement:	n^x	C_x^{n+x-1}
No replacement:	$P_x^n \equiv \frac{n!}{(n-x)!}$	$C_x^n \equiv \frac{n!}{x!(n-x)!}$

7.9.1 Deriving the Fermi–Dirac and Bose–Einstein Distributions from Counting Configurations

We have finally arrived at the point of applying the method of Lagrange multipliers to Ω_{ferm} in (7.175) and Ω_{bos} in (7.176). To reiterate, we must find the values of $n_1, n_2, \dots$ that maximise Ω_{ferm} , and likewise the values of $n_1, n_2, \dots$ that maximise Ω_{bos} . These values will be the most likely numbers of balls on each shelf—meaning, the most likely numbers of fermions and bosons at each energy level of the system.

Just as for the distinguishable case (7.166), the constraints are a fixed number of balls N and a fixed total energy E . And again, it's equivalent and easier to maximise $\ln \Omega_{\text{ferm}}$ and $\ln \Omega_{\text{bos}}$. The latter two calculations are similar, so we run them in parallel. Begin with a generic Ω that represents either of (7.175) and (7.176):

$$\Omega = \prod_i C_{n_i}^{a_i} \quad \begin{cases} a_i = \Omega_i & (\text{fermions}), \\ a_i = n_i + \Omega_i - 1 & (\text{bosons}). \end{cases} \quad (7.177)$$

Write

$$\begin{aligned} \ln \Omega &= \sum_i \ln \frac{a_i!}{n_i! (a_i - n_i)!} \\ &= \sum_i \ln(a_i!) - \ln(n_i!) - \ln[(a_i - n_i)!]. \end{aligned} \quad (7.178)$$

Apply the simplified Stirling's rule (1.27):

$$\begin{aligned} \ln \Omega &\simeq \sum_i a_i \ln a_i - \cancel{\mathcal{A}_i} - n_i \ln n_i + \cancel{\mathcal{B}_i} - (a_i - n_i) \ln(a_i - n_i) + \cancel{\mathcal{A}_i} - \cancel{\mathcal{B}_i} \\ &= \sum_i a_i \ln a_i - n_i \ln n_i - (a_i - n_i) \ln(a_i - n_i). \end{aligned} \quad (7.179)$$

The Lagrange-multiplier equation (7.168) is, for our generic Ω ,

$$\frac{\partial}{\partial n_k} \ln \Omega = \alpha + \beta E_k. \quad (7.180)$$

Insert the right-hand side of (7.179) into the left-hand side of (7.180):

$$\begin{aligned} \frac{\partial a_k}{\partial n_k} \ln a_k + \frac{a_k}{\cancel{\mathcal{A}_k}} \frac{\partial \cancel{\mathcal{A}_k}}{\partial n_k} - \ln n_k - \cancel{\mathcal{A}} \\ - \left(\frac{\partial a_k}{\partial n_k} - 1 \right) \ln(a_k - n_k) - \left(\frac{\partial \cancel{\mathcal{A}_k}}{\partial n_k} - \cancel{\mathcal{A}} \right) \simeq \alpha + \beta E_k. \end{aligned} \quad (7.181)$$

We must now consider fermions and bosons separately.

- **Fermions:** Equation (7.177) says $a_k = \Omega_k$. Thus, $\partial a_k / \partial n_k = 0$, and (7.181) becomes

$$-\ln n_k + \ln(\Omega_k - n_k) \simeq \alpha + \beta E_k. \quad (7.182)$$

This solves for n_k to return

$$n_k \simeq \frac{\Omega_k}{\exp(\alpha + \beta E_k) + 1}. \quad (7.183)$$

This value of n_k is the most likely number of balls (that is, fermions) that each have energy E_k . These fermions are spread over the Ω_k pots (quantum states) comprising energy level k . We now approximate the most likely number of fermions per quantum state at energy E_k —meaning, the occupation number—as this number n_k divided by the number of pots (quantum states). Hence, the occupation number is approximated as

$$\frac{n_k}{\Omega_k} \simeq \frac{1}{\exp(\alpha + \beta E_k) + 1}. \quad (7.184)$$

This expression is consistent with (7.106), which resulted from a different argument. We won't further analyse what α and β might be, for the same reason given just after (7.173) for the Boltzmann distribution.

- **Bosons:** Equation (7.177) says that $a_k = n_k + \Omega_k - 1$. Thus, $\partial a_k / \partial n_k = 1$, and (7.181) becomes

$$\ln(n_k + \Omega_k - 1) - \ln n_k \simeq \alpha + \beta E_k. \quad (7.185)$$

This solves for n_k / Ω_k to return the boson occupation number as

$$\frac{n_k}{\Omega_k} \simeq \frac{1 - 1/\Omega_k}{\exp(\alpha + \beta E_k) - 1}. \quad (7.186)$$

This expression is *almost* consistent with (7.106), apart from the “ $-1/\Omega_k$ ” in the numerator. But, given that we used Stirling's approximation to reach this point, perhaps we should not expect too exact an agreement with (7.106).

And What About Identical-Classical Particles?

We “constructed” identical-classical particles back in Section 1.1.1 as a way of introducing pseudo-quantum particles without then knowing anything about real quantum particles. Our treatment was necessarily simplistic: we simply divided the number of states accessible to N classical particles by $N!$, with no reference to any quantum ideas. The same approach here would divide (7.165) by $N!$, to produce (with subscript “ic” for “identical classical”)

$$\Omega_{\text{ic}}(n_1, n_2, \dots, \Omega_1, \Omega_2, \dots) = \frac{\Omega_1^{n_1} \Omega_2^{n_2} \Omega_3^{n_3} \dots}{n_1! n_2! n_3! \dots}. \quad (7.187)$$

This is not necessarily a whole number, and so this simple treatment cannot be fully correct. Even so, if we plough on by applying the method of Lagrange multipliers to $\ln \Omega_{\text{ic}}$, almost the same equations result as those that appeared for $\ln \Omega_{\text{dist}}$; for example, (7.171) appears again, but without the $\ln N!$ term. But the absence of that factorial has no effect on the partial derivative in (7.171). The upshot is that we arrive at (7.173) once more.

Why did the division by $N!$ go awry here? Return to the example of placing three balls on the shelves in Figure 7.18. At left in the figure, the balls have different energies. If these balls are distinguishable, they can be numbered and placed on the three shelves in $3!$ different ways (configurations). If those number labels are erased, only one way of distributing the balls will result (being that shown in the figure): we must indeed divide the $3!$ numbered configurations by $3!$ to arrive at this single one. But what happens when the balls have the *same* energy, at the right in Figure 7.18, but are distinguishable? When they are numbered, only one configuration is possible. If we erase those numbers, it's still the case that only one configuration is possible. We cannot divide the one configuration here by $3!$ to turn “distinguishable” into “identical classical”.

The point here is that if we are to divide a number of classical configurations by $3!$ to turn distinguishable particles into identical-classical particles, they are required all to have different energies. Classically (such as when we introduced identical-classical particles in Chapter 1), we assumed a system's energy to be a continuum; in which case, clearly, its particles *do* have different energies. We were then able to divide the number of available states by $N!$ to resolve, for example, Gibbs' Paradox in Section 7.4. But quantum particles *can* have the same energy, by occupying the same quantum state. This means we cannot generally just divide a number of classical configurations by $N!$ to turn a classical system into a true quantum one. The success of the idea of identical-classical particles rests on our being able to model a classical system's particles as having different energies, even if those energies are only infinitesimally different.



Chapter 8

Fermion Statistics in Metals

In which we apply the statistical mechanics of fermions to derive a model of electrons' contribution to a metal's heat capacity, electrical conductivity, and thermal conductivity. We discuss conductors, semiconductors, and insulators, and finish with a description of light-emitting diodes.

8.1 Conduction Electrons' Contribution to Heat Capacity

Debye's model of heat capacity in the previous chapter matches experimental data closely over the entire temperature range for non-metallic crystals. But, as we noted in Section 7.3, it fails in the case of metals at very low temperature. Whereas Debye predicted the low-temperature result of $C_V^{\text{mol}} \propto T^3$ in (7.52), the experimental result is $C_V^{\text{mol}} \propto T$. A key feature that distinguishes a metal from, say, a salt crystal is that the metal conducts electricity: it appears to have a great number of "conduction electrons" that are free enough to form an electric current when an electric field is applied. These are the *valence electrons*, the outer electrons that are bound only loosely to atoms. Our modern view pictures a metal as a lattice of atoms immersed in a "sea" or "gas" of essentially free electrons; roughly speaking, each atom contributes one electron to this sea.

But, if these electrons form a sort of gas, shouldn't they also contribute to the metal's heat capacity at *high* temperature? This high-temperature limit is given by the Dulong–Petit law: $C_V^{\text{mol}} = 3R$, where R is the gas constant. (This value was also correctly predicted by Einstein's and Debye's models of crystal heat capacity.) Recall Section 4.1 for the explanation: the lattice of N metal atoms has $6N$ quadratic energy terms, because each of a lattice atom's 3 dimensions of motion has 2 vibrational quadratic energy terms, and there are N atoms in total. The lattice's total internal energy is then $E = 6N \times kT/2 = 3NkT$. Equation (4.18) says that

$$C_V^{\text{mol}} = \frac{1}{n} \left(\frac{\partial E}{\partial T} \right)_{V,N} = \frac{N_A}{N} \frac{\partial}{\partial T} 3NkT = 3N_A k = 3R. \quad (8.1)$$

So much for the lattice's contribution to the metal's heat capacity. What about a contribution from the "gas" of valence electrons? Each free electron

(approximately one per atom, meaning N in total) might contribute an additional 3 quadratic energy terms, all translational. These $3N$ extra terms are then expected to increase the metal's total internal energy to

$$E = 3NkT + 3N \times kT/2 = 9/2 NkT. \quad (8.2)$$

We would then expect a molar heat capacity of

$$\begin{aligned} C_V^{\text{mol}} &= \frac{1}{n} \left(\frac{\partial E}{\partial T} \right)_{V,N} = \frac{N_A}{N} \frac{\partial}{\partial T} 9/2 NkT = 9/2 N_A k = 9/2 R \\ &= \underbrace{3R}_{\text{crystal}} + \underbrace{3/2 R}_{\text{valence electrons}}. \end{aligned} \quad (8.3)$$

The electrons are thus predicted to supply an extra $3/2 R$ to the molar heat capacity. But this total value of $9/2 R$ is not observed experimentally; so, either this gas of electrons doesn't exist (although such an idea conflicts with the success of the electric-conduction model), or the gas of electrons simply cannot be acting classically. We will assume the electron gas *does* exist, in which case it must have a strongly quantum nature. And indeed, in Section 7.5, we compared the high spatial density of electrons in copper with their de Broglie wavelength, and concluded that they *must* be treated quantum mechanically. It appears, then, that the electron gas has *no* quadratic energy terms that can each contribute $1/2 kT$ to its energy; it seems to be confined or constricted in some quantum-mechanical way.

This quantum-mechanical behaviour of being somehow constricted arises naturally if we assume electrons to be fermions, so that, at most, only one can occupy any given state. Such a model does indeed predict the experimental ultra-low temperature limit of $C_V^{\text{mol}} \propto T$ for metals. We'll describe how it all works in this section.

The total internal energy of the crystal is considered to be the sum of the energy contributed by the Debye model of lattice vibrations and the energy resulting from the Fermi sea of valence electrons; hence, C_V^{mol} is likewise a sum of those contributions. We found the Debye contribution to C_V^{mol} in (7.52). Now we calculate the electron contribution—call it “ $C_V^{\text{mol}}(\text{electrons})$ ”—to the molar heat capacity using the form of (7.1), where $\bar{E}$ is now the mean energy per *valence electron*, not the mean energy per *lattice oscillator* as it was in (7.1). Start with

$$C_V^{\text{mol}}(\text{electrons}) = N_A \left(\frac{\partial \bar{E}}{\partial T} \right)_{V,N}. \quad (8.4)$$

The mean energy per valence electron $\bar{E}$ is found from (7.31). That equation gave the total energy “ E_{tot} ” of the quantum particles, and so we must divide E_{tot} by the total number of valence electrons N here:

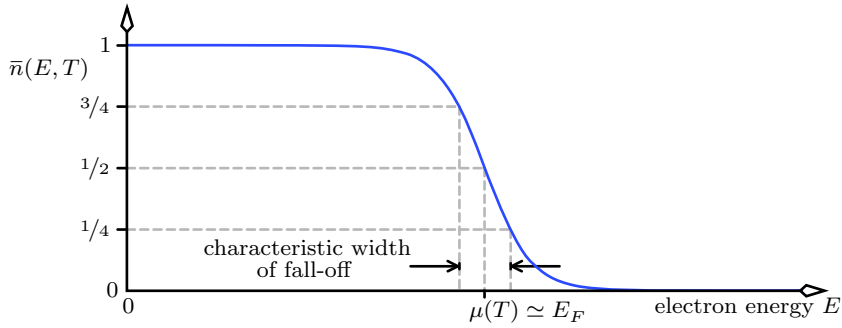


Fig. 8.1 For temperatures that are not “too high” (as determined in the text), $\bar{n}(E, T)$ looks much like the simple step-function form of $\bar{n}(E, 0)$ in Figure 7.11. To calculate the extent to which $\bar{n}(E, T)$ departs from this simple form, we must estimate the width of the fall-off around $\mu(T)$ as defined by two representative values of $\bar{n}(E, T)$ that bracket the middle of the fall-off

$$\bar{E} \stackrel{(7.31)}{=} \frac{1}{N} \int_0^\infty E \bar{n}(E, T) g(E) dE. \quad (8.5)$$

To evaluate this integral, we will use (7.114) for $g(E)$, but it turns out that the approximation (7.129) for $\bar{n}(E, T)$ is not accurate enough to predict the correct low-temperature dependence of $C_V^{\text{mol}}(\text{electrons})$. Instead, use the original expression for $\bar{n}(E, T)$ from (7.106):

$$\bar{n}(E, T) = \frac{1}{\exp \frac{E - \mu(T)}{kT} + 1}, \quad g(E) = C\sqrt{E}. \quad (8.6)$$

We must calculate the chemical potential $\mu(T)$ by using the “normalisation” (7.111). We will explore the shape of $\bar{n}(E, T)$ versus E for copper’s valence electrons.

Recall that Figure 7.11 showed the general form of $\bar{n}(E, T)$ versus E . In particular, for temperatures that are not “too high” (we’ll determine what that means shortly), $\bar{n}(E, T)$ versus E looks like the plot in Figure 8.1. For the purpose of an integration involving $\bar{n}(E, T)$, this plot can be approximated by the one in Figure 8.2. To use Figure 8.2 later in (8.20) and the integration (8.21), we need only determine the characteristic width of the fall-off in Figure 8.1. What is this width? The symmetry of the plot around its fall-off in Figure 8.1 says that the middle of the fall-off occurs at $E = \mu(T)$, where the occupation number falls to value 1/2:

$$\bar{n}(\mu(T), T) \stackrel{(8.6)}{=} \frac{1}{\exp \frac{\mu(T) - \mu(T)}{kT} + 1} = \frac{1}{2}. \quad (8.7)$$

Define a characteristic width of the fall-off as, say, bracketed by values of E that correspond to occupation numbers of 3/4 and 1/4. So, set $\bar{n}$ to both 3/4

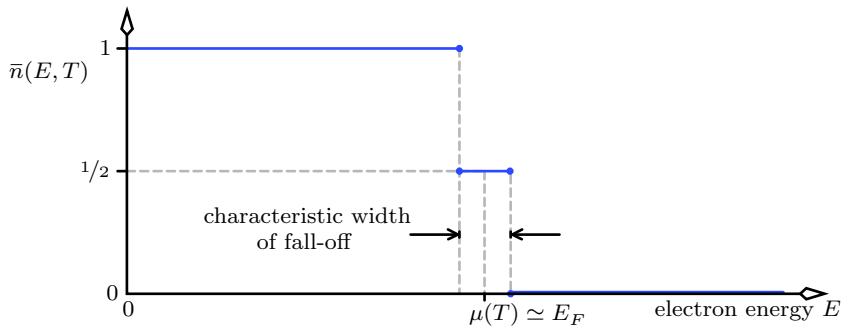


Fig. 8.2 The plot in Figure 8.1 can be approximated by the piece-wise function shown here

and $1/4$:

$$\frac{1}{\exp \frac{E - \mu(T)}{kT} + 1} = \begin{cases} 3/4 \\ 1/4. \end{cases} \quad (8.8)$$

Solving for E gives us (where the upper and lower signs correspond to solving for $3/4$ and $1/4$, respectively)

$$E = \mu(T) \mp kT \ln 3 \simeq \mu(T) \mp kT. \quad (8.9)$$

The characteristic width of the fall-off is thus $2kT$ for this choice of how the width is defined. Aside from the factor of 2, the main point here is that the width is around kT —which might not be surprising in hindsight, since kT characterises so many instances of energy in statistical mechanics.

Before evaluating the integral (8.5), it's useful to analyse the form of $\bar{n}(E, T)$ a little more deeply. In the discussion just after (7.128), we stated that when $kT \ll E_F$, the chemical potential is $\mu(T) \simeq E_F$. What is the Fermi energy of copper's valence electrons? Equation (7.132) says

$$E_F = \frac{h^2}{2m} \left(\frac{3N}{4\pi(2s+1)V} \right)^{2/3}. \quad (8.10)$$

Electrons have two spins, so $s = 1/2$ and $2s + 1 = 2$. Equation (8.10) becomes

$$E_F = \frac{h^2}{8m} \left(\frac{3N}{\pi V} \right)^{2/3}. \quad (8.11)$$

We calculated the number density of copper's valence electrons in (7.57) to be $N/V = 8.47 \times 10^{28}$ electrons/m³, assuming one free electron per atom. The electron's mass is $m = 9.11 \times 10^{-31}$ kg. Equation (8.11) then becomes (in SI units with a final conversion to electron volts)

$$E_F = \frac{(6.626 \frac{-34}{-31})^2}{8 \times 9.11 \frac{-31}{-31}} \left(\frac{3 \times 8.47 \frac{28}{-31}}{\pi} \right)^{2/3} \frac{1}{1.602 \frac{-19}{-31}} \text{ eV} \simeq 7.0 \text{ eV}. \quad (8.12)$$

We now define the *Fermi temperature* T_F via

$$kT_F \equiv E_F. \quad (8.13)$$

The equipartition theorem tells us that a temperature of at least approximately T_F is needed for an external classical influence (such as the motion of lattice atoms) to excite many electrons into states with energies higher than E_F . The Fermi temperature of copper is

$$T_F = \frac{7.0 \text{ eV}}{k} = \frac{7.0 \times 1.602 \frac{-19}{-23}}{1.381 \frac{-23}{-23}} \text{ K} \simeq 81,000 \text{ K}. \quad (8.14)$$

This very high value implies that the behaviour of electrons in copper metal is extremely insensitive to temperature changes. The physical explanation for this is that if two colliding valence electrons were to interact, then for one electron to be pushed into an unoccupied level of higher energy, the other electron would have to drop to a lower unoccupied level; but almost all lower levels are occupied, even at temperatures as high as T_F . Most electrons are not near the “surface of the Fermi sea” at $E = 7 \text{ eV}$, and so interactions with the copper atoms at room temperature (whose energies are around $kT \simeq 1/40 \text{ eV}$) cannot excite the vast majority of electrons to energies above 7 eV . Only electrons with energy within about kT of E_F can be excited into unoccupied states. It follows that even at very high temperatures, the occupation number of the electrons varies little from its distribution at zero temperature. That means we can treat the valence electrons as non-interacting, even though, classically, one would expect them to interact electromagnetically.

Figure 7.12's depiction of the Fermi sea also shows that most electrons have an energy not very far below E_F . It follows that the mean energy $\bar{E}$ of each electron is not far below E_F . With $g(E) = C\sqrt{E}$, equation (8.5) yields

$$\begin{aligned} \bar{E} &= \frac{1}{N} \int_0^\infty E \bar{n}(E, T) g(E) dE \\ &\simeq \frac{1}{N} \int_0^\infty E \bar{n}(E, 0) g(E) dE \stackrel{(7.130)}{=} \frac{1}{N} \int_0^{E_F} E \times 1 \times CE^{1/2} dE \\ &\stackrel{(7.131)}{=} \frac{2/5 CE_F^{5/2}}{2/3 CE_F^{3/2}} = 3/5 E_F. \end{aligned} \quad (8.15)$$

The Fermi energy E_F serves to define the *Fermi speed* v_F of the valence electrons, via

$$\frac{1}{2}mv_F^2 \equiv E_F. \quad (8.16)$$

(Note that some authors define the Fermi speed via the *mean* electron energy: $\frac{1}{2}mv_F^2 \equiv \bar{E} = \frac{3}{5}E_F$.) The Fermi speed of copper's valence electrons is

$$v_F = \sqrt{\frac{2E_F}{m}} \simeq \sqrt{\frac{2 \times 7.0 \times 1.602 \frac{-19}{-31}}{9.11 \frac{-31}{-31}}} \text{ m/s} \simeq 1570 \text{ km/s}. \quad (8.17)$$

This value is typical of most metals. Because the occupation number $\bar{n}(E, T)$ is very insensitive to temperature changes below $T_F \simeq 81,000 \text{ K}$ (meaning *all* temperatures of interest for metallic copper), the Fermi speed is also very insensitive to such temperature changes. Compare the value of 1570 km/s to the Maxwell mean speed $\bar{v} \propto T^{1/2}$ in (6.53), which the electrons would have if they were a classical cloud of particles. At room temperature, the Maxwell expression (6.53) says that this mean speed is

$$\bar{v} = \sqrt{\frac{8kT}{\pi m}} = \sqrt{\frac{8 \times 1.381 \frac{-23}{-31} \times 298}{\pi \times 9.11 \frac{-31}{-31}}} \text{ m/s} \simeq 107 \text{ km/s}. \quad (8.18)$$

Because no more than one electron can occupy each state, the electrons can only pile up into states of ever higher energy, which forces them to move with a much larger range of speeds (around 1570 km/s) than classical physics would predict (around 107 km/s). We'll encounter this difference between classical and quantum-mechanical speeds of valence electrons when studying electrical conduction in Section 8.2.

Let's now return to our calculation of C_V^{mol} (electrons) in (8.4). This is a derivative of the average energy per electron, $\bar{E}$, for which we need $\bar{n}(E, T)$ and $g(E)$ in (8.6). We will approximate $\bar{n}(E, T)$ by the piece-wise function shown in Figure 8.2, whose characteristic drop-off width is about kT . But we require $\mu(T)$ (which almost equals E_F). We can find $\mu(T)$ from the “normalisation condition” (7.111).

The characteristic width of roughly $2kT$ shown in (8.9) suggests that we focus on the value of $\bar{n}(E, T)$ to within a few kT of $E = \mu(T)$. We can investigate the sensitivity of the molar heat capacity to our choice of “a few kT ” by defining a parameter α to be around 1 or 2:

$$\alpha \equiv \text{a number in the region of 1 or 2}. \quad (8.19)$$

Use this to write the approximation of $\bar{n}(E, T)$ in Figure 8.2 as (while dropping explicit mention of the dependence of μ on T)

$$\bar{n}(E, T) = \frac{1}{\exp \frac{E-\mu}{kT} + 1} \simeq \begin{cases} 1 & E < \mu - \alpha kT, \\ 1/2 & \mu - \alpha kT < E < \mu + \alpha kT, \\ 0 & \mu + \alpha kT < E. \end{cases} \quad (8.20)$$

Equation (7.111) becomes

$$\begin{aligned}
N &= \int_0^\infty \bar{n}(E, T) C E^{1/2} dE \\
&\simeq \int_0^{\mu - \alpha kT} C E^{1/2} dE + \int_{\mu - \alpha kT}^{\mu + \alpha kT} \frac{C E^{1/2}}{2} dE + \int_{\mu + \alpha kT}^\infty 0.
\end{aligned} \tag{8.21}$$

The fall-off zone around $E = \mu(T)$ is extremely narrow, so use the approximation $E = \mu$ within it. Equation (8.21) becomes

$$\begin{aligned}
N &= \frac{2C}{3} (\mu - \alpha kT)^{3/2} + \frac{C\mu^{1/2}}{2} 2\alpha kT \\
&= \frac{2C\mu^{3/2}}{3} \left(1 - \frac{\alpha kT}{\mu}\right)^{3/2} + C\mu^{1/2} \alpha kT.
\end{aligned} \tag{8.22}$$

(In the various calculations that follow, we use the binomial theorem freely and retain only leading-order terms.) But note that $\alpha kT \ll \mu \simeq E_F$, and so we will expand the parenthesis in the last line of (8.22) to order T^2 . (Because this T^2 turns out not to cancel anywhere, we needn't expand to higher orders.) Setting $\zeta \equiv \alpha kT/\mu$, write the last line of (8.22) as

$$\begin{aligned}
\frac{N}{C\mu^{3/2}} &= \frac{2}{3}(1 - \zeta)^{3/2} + \zeta \simeq \frac{2}{3} \left(1 - \frac{3}{2}\zeta + \frac{3/2 \cdot 1/2}{2!} \zeta^2\right) + \zeta \\
&= \frac{2}{3} + \frac{\zeta^2}{4}.
\end{aligned} \tag{8.23}$$

In other words,

$$N = C\mu^{3/2} \left[\frac{2}{3} + \frac{1}{4} \left(\frac{\alpha kT}{\mu} \right)^2 \right]. \tag{8.24}$$

Now solve this for μ . Recall that μ differs only marginally from E_F for temperatures below T_F , so write $\mu = E_F + O(T)$ (that is, E_F plus terms of order T that are much smaller than E_F). It follows that

$$E_F^{3/2} \stackrel{(7.132)}{=} \frac{3N}{2C} \stackrel{(8.24)}{=} \mu^{3/2} \left[1 + \frac{3}{8} \left(\frac{\alpha kT}{E_F + O(T)} \right)^2 \right]. \tag{8.25}$$

Clearly then,

$$E_F \simeq \mu \left[1 + \frac{3}{8} \left(\frac{\alpha T}{T_F} \right)^2 + O(T^3) \right]^{2/3} \simeq \mu \left[1 + \frac{1}{4} \left(\frac{\alpha T}{T_F} \right)^2 \right]. \tag{8.26}$$

Hence,

$$\mu \simeq E_F \left[1 - \frac{1}{4} \left(\frac{\alpha T}{T_F} \right)^2 \right]. \tag{8.27}$$

(As a check, we see that $\mu = E_F$ at zero temperature, as expected.) Now that $\mu(T)$ is known for $T \ll T_F$ (in other words, for all realistic temperatures), we can calculate $\bar{E}$ from (8.5). The calculation proceeds in exactly the same way as that of (8.21) and (8.22), except that $g(E) = CE^{1/2}$ is replaced by $Eg(E) = CE^{3/2}$:

$$\begin{aligned}
 N\bar{E} &= \int_0^\infty E \bar{n}(E, T) g(E) dE \\
 &\stackrel{(8.20)}{=} \int_0^{\mu - \alpha kT} CE^{3/2} dE + \int_{\mu - \alpha kT}^{\mu + \alpha kT} \frac{C\mu^{3/2}}{2} dE + \int_{\mu + \alpha kT}^\infty 0 \\
 &= \frac{2C}{5}(\mu - \alpha kT)^{5/2} + \frac{C\mu^{3/2}}{2}2\alpha kT \\
 &= \frac{2C\mu^{5/2}}{5}(1 - \zeta)^{5/2} + C\mu^{3/2}\alpha kT \\
 &= C\mu^{5/2} \left[\frac{2}{5}(1 - \zeta)^{5/2} + \zeta \right] \simeq C\mu^{5/2} \left[\frac{2}{5} \left(1 - \frac{5}{2}\zeta + \frac{5/2 \cdot 3/2}{2!} \zeta^2 \right) + \zeta \right] \\
 &= C\mu^{5/2} \left[\frac{2}{5} + \frac{3\zeta^2}{4} \right]. \tag{8.28}
 \end{aligned}$$

Thus,

$$\bar{E} = \frac{C\mu^{5/2}}{N} \left[\frac{2}{5} + \frac{3}{4} \left(\frac{\alpha kT}{\mu} \right)^2 \right]. \tag{8.29}$$

Substitute μ from (8.27) into (8.29), to obtain

$$\begin{aligned}
 \bar{E} &= \frac{CE_F^{5/2}}{N} \left[1 - \frac{5}{2} \cdot \frac{1}{4} \left(\frac{\alpha T}{T_F} \right)^2 \right] \left[\frac{2}{5} + \frac{3}{4} \left(\frac{\alpha T}{T_F} \right)^2 \right] \\
 &\simeq \frac{CE_F^{5/2}}{N} \left[\frac{2}{5} + \frac{1}{2} \left(\frac{\alpha T}{T_F} \right)^2 \right]. \tag{8.30}
 \end{aligned}$$

But

$$E_F^{5/2} = E_F \times E_F^{3/2} \stackrel{(7.132)}{=} E_F \frac{3N}{2C}, \tag{8.31}$$

and so (8.30) becomes

$$\bar{E} = E_F \frac{3}{2} \left[\frac{2}{5} + \frac{1}{2} \left(\frac{\alpha T}{T_F} \right)^2 \right] = E_F \left[\frac{3}{5} + \frac{3}{4} \left(\frac{\alpha T}{T_F} \right)^2 \right]. \tag{8.32}$$

Now we apply (8.4) to calculate the electron contribution to the molar heat capacity:

$$\begin{aligned}
C_V^{\text{mol}}(\text{electrons}) &= N_A \left(\frac{\partial \bar{E}}{\partial T} \right)_{V,N} = N_A E_F \frac{3}{2} \left(\frac{\alpha}{T_F} \right)^2 T \\
&= \frac{3}{2} R \alpha^2 \frac{T}{T_F}, \tag{8.33}
\end{aligned}$$

where R is the gas constant. With the choice $\alpha \approx 1$, $C_V^{\text{mol}}(\text{electrons})$ has the Dulong–Petit value of $3R/2$ [recall (8.3)] in the region of $T = T_F$. But the value of $C_V^{\text{mol}}(\text{electrons})$ in (8.33) is *far* less than $3R/2$ at all realistic temperatures $T \ll T_F$. We see that electrons really contribute very little to the metal's total heat capacity. We also see that the contribution from the valence electrons is proportional to T , which suggests that we're on the right track to agree with experiment. We discuss that in detail next.

Comparison with Measurement of Heat Capacity

The total internal energy of the crystal is the sum of the contributions from the Debye model of lattice vibrations and the Fermi model of the valence-electron sea. It follows that the *total* molar heat capacity C_V^{mol} of a metal crystal is just the sum of the Debye term ($\propto T^3$) and the Fermi term ($\propto T$):

$$C_V^{\text{mol}} = \underbrace{\frac{12\pi^4 R}{5} \left(\frac{T}{T_D} \right)^3}_{\text{Debye (7.52)}} + \underbrace{\frac{3}{2} R \alpha^2 \frac{T}{T_F}}_{\text{Fermi (8.33)}}. \tag{8.34}$$

This predicts that a plot of measured values of C_V^{mol}/T versus T^2 should yield a straight line whose slope and intercept are taken from:

$$\frac{C_V^{\text{mol}}}{T} = \underbrace{\frac{12\pi^4 R}{5T_D^3}}_{\text{slope}} T^2 + \underbrace{\frac{3R\alpha^2}{2T_F}}_{\text{intercept}}. \tag{8.35}$$

Laboratory measurements of the molar heat capacity of copper at low temperatures yield typical values of

$$\begin{aligned}
\text{slope} &= 5 \times 10^{-5} \text{ J K}^{-4} \text{ mol}^{-1}, \\
\text{intercept} &= 7 \times 10^{-4} \text{ J K}^{-2} \text{ mol}^{-1}. \tag{8.36}
\end{aligned}$$

What are the theoretical predictions in (8.35)? The slope is fully set by Debye's theory of lattice vibrations, and the intercept fully by the Fermi theory of the electron sea. We found the slope in (7.60) to be $5.5 \times 10^{-5} \text{ J K}^{-4} \text{ mol}^{-1}$, in excellent agreement with (8.36)'s prediction.

To calculate the predicted intercept from (8.35), we need copper's Fermi temperature $T_F \simeq 81,000 \text{ K}$ from (8.14). Hence,

$$\text{intercept} = \frac{3R\alpha^2}{2T_F} \simeq \frac{3 \times 8.314 \alpha^2}{2 \times 81,000} \text{ SI units} = 1.5 \times 10^{-4} \alpha^2 \text{ SI}. \quad (8.37)$$

Recall that α is a measure of the width of the fall-off region of $\bar{n}(E, T)$, as shown in Figure 8.1; its value in (8.19) is somewhat vaguely in the region of “several” (which needn’t be a whole number). Choosing values 1, 2, and 3, produces intercepts of $(1.5, 6.0, 13.5) \times 10^{-4}$ SI, respectively. These intercepts certainly encompass the experimental value in (8.36). This agreement of theory with experiment shows that the idea of treating valence electrons as though they form a Fermi sea in the lattice is a good one.

8.1.1 A More Accurate Approximation of $\bar{n}(E, T)$

It turns out that we can do better in the integration of (8.21) and (8.22), and so avoid invoking the factor of α that was introduced in (8.19). We show how to do that here, and then determine what value of α would have given the same result: this forms a good insight into just how to define the width of the fall-off region of a function like $\bar{n}(E, T)$ that has an exponential form.

Begin with the general integral

$$I = \int_0^\infty \bar{n}(E, T) \phi(E) dE, \quad (8.38)$$

where $\phi(E)$ is a generic function that denotes either of $g(E) = CE^{1/2}$ [which we used in (8.21)] and $Eg(E) = CE^{3/2}$ [used in (8.28)]. The thing to notice is that $\bar{n}'(E, T) \simeq 0$ everywhere outside the fall-off region, where a prime here and in what follows indicates $\partial/\partial E$. Now evaluate (8.38) using an integration by parts. To facilitate that, define

$$\Phi'(E) \equiv \phi(E). \quad (8.39)$$

Equation (8.38) is then

$$I = \bar{n}(E, T) \Phi(E) \Big|_{E=0}^\infty - \int_0^\infty \bar{n}'(E, T) \Phi(E) dE. \quad (8.40)$$

We can take $\bar{n}(\infty, T) \Phi(\infty) = 0$. Also, $\bar{n}(0, T) \Phi(0) = \Phi(0)$, and so

$$I = \Phi(0) - \int_0^\infty \bar{n}'(E, T) \Phi(E) dE. \quad (8.41)$$

What is $\Phi(0)$? As usual for fermions, we are treating $\phi(E)$ as a continuous function, and hence we might set $\Phi(0) \equiv 0$. But $\phi(E)$ is *not* really continuous. After all, if $\phi(E)$ is set to be the density of states $g(E)$, then $\Phi(E)$ is the

number of states $\Omega_{\text{tot}}(E)$; but that number of states is never zero, according to the Third Law of Thermodynamics. Even so, the value of $\Phi(0)$ will be much smaller than I : remember that $I = N$ in (7.111) when $\phi(E) = g(E)$. So, we will ignore $\Phi(0)$: namely, set

$$\Phi(0) \equiv 0. \quad (8.42)$$

We now have

$$I = - \int_0^\infty \bar{n}'(E, T) \Phi(E) dE. \quad (8.43)$$

Now focus on $\Phi(E)$ in the fall-off region, by expanding it as a Taylor series around μ :

$$\Phi(E) = \Phi(\mu) + \phi(\mu)(E - \mu) + \frac{\phi'(\mu)}{2!}(E - \mu)^2 + \dots \quad (8.44)$$

Combine (8.43) and (8.44):

$$\begin{aligned} I &= - \int_0^\infty \bar{n}'(E, T) \left[\Phi(\mu) + \phi(\mu)(E - \mu) + \frac{\phi'(\mu)}{2!}(E - \mu)^2 + \dots \right] dE \\ &= -\Phi(\mu) \int_0^\infty \bar{n}'(E, T) dE - \phi(\mu) \int_0^\infty \bar{n}'(E, T)(E - \mu) dE \\ &\quad - \frac{\phi'(\mu)}{2} \int_0^\infty \bar{n}'(E, T)(E - \mu)^2 dE - \dots \end{aligned} \quad (8.45)$$

Make a change of variables $x \equiv \beta(E - \mu)$, in which case

$$\bar{n}(E, T) \stackrel{(7.106)}{=} \frac{1}{e^x + 1}. \quad (8.46)$$

Then, since we are working at a fixed temperature,

$$\bar{n}'(E, T) dE = d\bar{n}(E, T) = \frac{d}{dx} \bar{n}(E, T) dx = \frac{-e^x dx}{(e^x + 1)^2} = \frac{-dx}{4 \text{ch}^2 \frac{x}{2}}, \quad (8.47)$$

where “ch” is the hyperbolic cosine.¹ Equation (8.45) becomes

$$I = \Phi(\mu) \int_{\frac{-\mu}{kT}}^\infty \frac{dx}{4 \text{ch}^2 \frac{x}{2}} + \phi(\mu) \int_{\frac{-\mu}{kT}}^\infty \frac{kTx dx}{4 \text{ch}^2 \frac{x}{2}} + \frac{\phi'(\mu)}{2} \int_{\frac{-\mu}{kT}}^\infty \frac{(kT)^2 x^2 dx}{4 \text{ch}^2 \frac{x}{2}} + \dots \quad (8.48)$$

But note that $\mu/k \simeq E_F/k = T_F = 81,000 \text{ K}$ for copper, whereas T is just a few kelvins for the low temperatures that we are working with here; hence,

¹ The hyperbolic cosine is, of course, often written as “cosh”, although you will find it as “ch” in some tables of integrals. Replacing “sinh, cosh, tanh” with the lesser-known alternatives “sh, ch, th” is very convenient in lengthy hand-written calculations with the hyperbolic functions.

$-\mu/(kT) \ll 0$. The integrands in (8.48) are immensely suppressed at negative values of x by the huge (positive) values of the hyperbolic cosines in their denominators. This all means that we make no real error by replacing the lower limit of integration $-\mu/(kT)$ with $-\infty$. That produces three standard integrals whose values can be found in standard tables of integrals. We will simply quote the results:²

$$\begin{aligned} I &\simeq \frac{\Phi(\mu)}{4} \underbrace{\int_{-\infty}^{\infty} \frac{dx}{\text{ch}^2 \frac{x}{2}}}_{=4} + \frac{\phi(\mu)kT}{4} \underbrace{\int_{-\infty}^{\infty} \frac{x dx}{\text{ch}^2 \frac{x}{2}}}_{=0} + \frac{\phi'(\mu)(kT)^2}{8} \underbrace{\int_{-\infty}^{\infty} \frac{x^2 dx}{\text{ch}^2 \frac{x}{2}}}_{=4\pi^2/3} \\ &= \Phi(\mu) + \phi'(\mu)(kT)^2 \pi^2/6. \end{aligned} \quad (8.49)$$

We can use this value of I to determine the integrals in (8.21) and (8.28) more accurately. Equation (8.21) was the normalisation:

$$N = \int_0^{\infty} \bar{n}(E, T) C E^{1/2} dE. \quad (8.50)$$

Compare this to the generic form (8.38), to set

$$I = N, \quad \text{and} \quad \phi(E) = C E^{1/2}. \quad (8.51)$$

Refer to (8.39) and (8.42), to write $\Phi(E) = 2/3 C E^{3/2}$. Then, (8.49) yields

$$N = C \mu^{3/2} \left[\frac{2}{3} + \frac{\pi^2}{12} \left(\frac{kT}{\mu} \right)^2 \right]. \quad (8.52)$$

This compares well with (8.24).

Similarly, in (8.28),

$$I = N \bar{E}, \quad \text{and} \quad \phi(E) = E g(E) = C E^{3/2}. \quad (8.53)$$

Hence, $\Phi(E) = 2/5 C E^{5/2}$. Equation (8.49) becomes

$$\bar{E} = \frac{C \mu^{5/2}}{N} \left[\frac{2}{5} + \frac{\pi^2}{4} \left(\frac{kT}{\mu} \right)^2 \right]. \quad (8.54)$$

This compares well with (8.29). In summary,

$$\begin{aligned} &\text{for } N, \text{ the more accurate version of (8.24) is (8.52);} \\ &\text{for } \bar{E}, \text{ the more accurate version of (8.29) is (8.54).} \end{aligned} \quad (8.55)$$

² The second integral's value of zero is trivial, since the integrand is an odd function.

Whereas the old expressions, (8.24) and (8.29), required the unknown factor α , the new expressions, (8.52) and (8.54), do not.

With these new and improved expressions (8.52) and (8.54) for N and $\bar{E}$, we follow the same procedure that was applied just after (8.29): solve (8.52) for μ , and place this into (8.54). A similar analysis to that used just after (8.24) leads to

$$\mu \simeq E_F \left[1 - \frac{\pi^2}{12} \left(\frac{T}{T_F} \right)^2 \right]. \quad (8.56)$$

This compares well with (8.27). Now substitute this μ into (8.54), to obtain the final result for the mean energy:

$$\bar{E} = E_F \left[\frac{3}{5} + \frac{\pi^2}{4} \left(\frac{T}{T_F} \right)^2 \right]. \quad (8.57)$$

This compares well with (8.32).

We can now apply (8.4) to find the electron contribution to the molar heat capacity:

$$C_V^{\text{mol}}(\text{electrons}) = N_A \left(\frac{\partial \bar{E}}{\partial T} \right)_{V,N} \stackrel{(8.57)}{=} \frac{\pi^2 R}{2} \frac{T}{T_F}, \quad (8.58)$$

where R is the gas constant. Compare this to (8.33): these two expressions would be identical if we chose α to be $\pi/\sqrt{3}$, or about 1.8. This is completely consistent with the original idea of (8.19) that α was to be “around 1 or 2”.

For copper metal at room temperature, the contribution of the lattice atoms to the molar heat capacity is approximately the Dulong–Petit value of $3R = 24.9$ SI units. The contribution of the valence electrons to the molar heat capacity is given by (8.58):

$$C_V^{\text{mol}}(\text{electrons}) = \frac{\pi^2 R}{2} \frac{T}{T_F} \simeq \frac{\pi^2 \times 8.314}{2} \times \frac{298}{81,000} \simeq 0.2 \text{ SI units}. \quad (8.59)$$

We see that the contribution of the valence electrons is negligible. Even at copper's melting point of $T = 1084^\circ\text{C} = 1357 \text{ K}$, this contribution has risen only to about 0.7 SI units. This shows that metals have pretty much the same room-temperature heat capacity as crystalline insulators: both about $3R$.

The Fermi term in the molar heat capacity, (8.34), originally came from (8.33). We now replace it with the more accurate expression in (8.58), to write

$$C_V^{\text{mol}} = \frac{12\pi^4 R}{5} \left(\frac{T}{T_D} \right)^3 + \frac{\pi^2 R}{2} \frac{T}{T_F}. \quad (8.60)$$

This replaces (8.34). The more accurate version of (8.35) is now

$$\frac{C_V^{\text{mol}}}{T} = \underbrace{\frac{12\pi^4 R}{5T_D^3}}_{\text{slope}} T^2 + \underbrace{\frac{\pi^2 R}{2T_F}}_{\text{intercept}}. \quad (8.61)$$

The Debye part of this (the slope) is, of course, unchanged from (8.35). The new number is the intercept, which is now

$$\text{intercept} = \frac{\pi^2 R}{2T_F} \simeq \frac{\pi^2 \times 8.314}{2 \times 81,000} \text{ SI units} = 5.0 \times 10^{-4} \text{ SI}. \quad (8.62)$$

This compares very well with the experimental value of 7×10^{-4} SI in (8.36).

Finally, given that a choice of α as $\pi/\sqrt{3} \simeq 1.8$ would have given agreement between our simple integrations of N and $N\bar{E}$ [in (8.21) and (8.28)] and the more exact integrations that followed (8.38), what bracketing values of the fall-off zone would this correspond to, replacing the 1/4 and 3/4 that we chose arbitrarily in Figure 8.1? In other words, what are the values of $\bar{n}(E, T)$ when $E = \mu \pm \alpha kT$, for $\alpha = \pi/\sqrt{3}$?

$$\bar{n}(\mu \pm \alpha kT, T) \stackrel{(7.106)}{=} \frac{1}{e^{\pm\alpha} + 1} = \frac{1}{\exp(\pm \pi/\sqrt{3}) + 1} = \left\{ \begin{array}{l} 0.14 \\ 0.86 \end{array} \right\}. \quad (8.63)$$

It follows that bracketing choices of 0.14 and 0.86 would have been better than our first estimates of 0.25 and 0.75 in Figure 8.1. These “better” bracketing numbers are approximately e^{-2} and $1 - e^{-2}$. This suggests that using these latter two values might be a good rule of thumb when a similar bracketing of an exponential function is called for in other areas of physics.

8.2 Electrical Conductivity of Metals

We have seen that classical physics fails to predict the correct low-temperature dependence of metallic heat capacity, and that the quantum theory of the valence-electron contribution to the heat capacity of metals gives results that are in close agreement with experiment.

That is all very well, but what else does quantum theory have to say? For the purpose of this discussion, we might take the approach that quantum mechanics is a “new” theory. A new theory is never created in isolation; we must always check for any consequences that might disagree with other experiments. In the current case, we must determine what this “new” quantum model predicts for the temperature dependence of the most obvious manifestation of valence electrons: a metal’s electrical conductivity.

Resistivity and Conductivity

Remember that a *resistor* is a piece of hardware in an electrical circuit. A *resistance* is the simplest idealisation that can represent the resistor when we analyse the circuit. A resistor has resistance, possibly accompanied by a small amount of capacitance and inductance that we can ignore here. The amount of resistance depends on the size and shape of the resistor. The material comprising the resistor has *resistivity*, which is an intensive variable: it does not depend on the size or shape of the resistor.

The treatment in this chapter assumes the simplest model: that resistivity and conductivity are each independent of direction in the material, and so are *scalars*. This scalar nature means that I have written either “resistivity” or “conductivity”, according to which of those terms might be used more frequently. Recalling (4.70), this is a trivial choice of language, because resistivity and conductivity scalars (thermal, electrical, etc.) are always defined as reciprocals of each other. So too are resistance and conductance.

In more complicated materials, resistivity and conductivity can depend on direction, and thus are *tensors*. Given a choice of coordinates, each can be written in component form as a matrix, and then these matrices will be inverses of each other. Their individual elements, of course, need not be reciprocals of each other—although they will be reciprocals if the matrices are diagonal. For some background on tensors, see the brief discussion in Section 1.9.3, and the longer discussions in Sections 5.6.1 and 6.10.

Classical physics fails to predict the correct value of a metal’s *thermal* conductivity, because thermal conductivity is related classically to heat capacity in (6.169). It turns out that classical physics also fails to predict the correct value of a metal’s *electrical* conductivity. But we are in luck: the “new” model of quantum mechanics that treats the valence electrons as a sea of non-interacting fermions turns out to agree with experimental measurements of thermal and electrical conductivities. We show how it all fits together in this section and the next.

Consider a resistor wire that is connected to a battery to make a circuit. A basic result of circuit theory says that the wire’s electrical resistance R is proportional to its length ℓ and inversely proportional to its cross-sectional area A . The constant of proportionality is the *electrical resistivity* ϱ of the material comprising the wire:

$$R = \frac{\varrho \ell}{A}. \quad (8.64)$$

[This is identical to the case of thermal resistance in (4.71). We used $\Delta\ell$ for the length in (4.71) as a result of the mathematical notation followed in that section, but we can just as well write the length as ℓ here.]

We will combine (8.64) with Ohm's experimentally determined rule for a resistor, " $V = IR$ ", to predict the temperature dependence of ϱ . We'll then check whether this dependence matches experiment.

Begin with a classical argument: although we model the electrons that carry the electric current as the gas of weakly bound valence electrons that we worked with in the previous section, we treat those electrons as non-interacting particles of a classical gas. We'll need the following parameters:

$$\begin{aligned}
 \nu_a &= \text{number density of lattice atoms}, & m_a &= \text{mass of one lattice atom}, \\
 \nu_e &= \text{number density of electrons}, & m_e &= \text{mass of one electron}, \\
 v_d &= \text{electron drift speed}, & q &= \text{electric charge of one electron}, \\
 \ell &= \text{length of resistor}, & A &= \text{cross-sectional area of resistor}, \\
 V &= \text{voltage drop across resistor}, & I &= \text{electric current in resistor}, \\
 \bar{v} &= \text{electrons' mean thermal speed from Maxwell distribution}, \\
 \lambda &= \text{mean free path of electrons interacting with metallic lattice}, \\
 \sigma &= \text{cross section of lattice atoms}.
 \end{aligned} \tag{8.65}$$

Combine (8.64) with Ohm's rule, to produce

$$R = \frac{\varrho\ell}{A} = \frac{V}{I}, \quad \text{or} \quad \varrho = \frac{AV}{\ell I}. \tag{8.66}$$

To determine ϱ , we require V and I . For the voltage drop V , realise that the electrons within the resistor are immersed in an electric field $\mathcal{E}$ created by the battery. We take this field to be uniform along the resistor. Basic electrostatics says that a uniform electric field $\mathcal{E}$ gives rise to a voltage drop $V = \mathcal{E}\ell$ along the resistor's length ℓ .

Next, we require the current I in the resistor. The drift of the charges is no different from the motion of particles that we studied in Section 6.8. In particular, examine Figure 6.10, replacing that figure's " v " with the drift speed v_d of the electrons. The charge passing through the wire's cross section A in a time Δt sweeps out a volume of $A \times \text{swept distance} = Av_d\Delta t$. With ν_e electrons in a unit volume, each with charge q , this swept volume holds a total charge of $\nu_e q Av_d\Delta t$. The current is then³

³ Recalling the language of current density is useful here. Remember that in Section 4.1.2's discussion of heat flow, we examined the thermal current density $\mathbf{J}$ in three dimensions. Equation (4.61) showed that $\mathbf{J}d\mathbf{A}$ was the heat current or thermal power (energy per unit time) flowing through an area $d\mathbf{A}$ lying perpendicular to $\mathbf{J}$. The same idea extends to electric current, where we define the *electric current density* $\mathbf{J}$: now, $\mathbf{J}d\mathbf{A}$ is the electric current (charge per unit time) flowing through an area $d\mathbf{A}$ lying perpendicular to $\mathbf{J}$. For the above analysis, $\mathbf{J}$ always points along the

$$I = \text{charge passed}/\Delta t = \nu_e q A v_d \Delta t / \Delta t = \nu_e q A v_d. \quad (8.68)$$

Now we require the electrons' drift speed v_d . The electric field $\mathcal{E}$ applies a force to the electrons that causes them to drift. Estimate their drift speed v_d by imagining the electrons to be constantly colliding with lattice atoms in such a way that the electrons are continually being accelerated from zero speed by a force $\mathcal{E}q$ until their next collision. On average, these collisions occur at time intervals of $\lambda/\bar{v}$, where λ is the mean free path of the electrons (Section 6.7) and $\bar{v}$ is their mean thermal speed from the Maxwell speed distribution (Section 6.3). Approximate v_d as the speed that an electron has acquired at the midpoint of this interval:

$$\begin{aligned} v_d &= \text{acceleration} \times \text{half the time interval between collisions} \\ &= \frac{\mathcal{E}q}{m_e} \times \frac{\lambda}{2\bar{v}}. \end{aligned} \quad (8.69)$$

This combines with (8.68) to give the current:

$$I = \nu_e q A \times \frac{\mathcal{E}q\lambda}{2m_e\bar{v}} = \frac{\nu_e \mathcal{E}q^2 A \lambda}{2m_e\bar{v}}. \quad (8.70)$$

Now insert this along with $V = \mathcal{E}\ell$ into (8.66):

$$\varrho = \frac{AV}{\ell I} = \frac{A}{\ell} \times \mathcal{E}\ell \times \frac{2m_e\bar{v}}{\nu_e \mathcal{E}q^2 A \lambda} = \frac{2m_e\bar{v}}{\nu_e q^2 \lambda}. \quad (8.71)$$

This expression for ϱ has two variables that we must inspect for possible temperature dependence: $\bar{v}$ and λ . Equation (6.53) says that the electrons in the classical gas have a mean speed of $\bar{v} \propto \sqrt{T}$. For λ , we might use (6.120) [not (6.124)] to write the electrons' mean free path through the fixed lattice atoms as

$$\lambda = \frac{1}{\nu_a \sigma}. \quad (8.72)$$

The number density ν_a of lattice atoms has no temperature dependence. Their cross section σ would have no temperature dependence *if* the lattice atoms were stationary spheres of radius r . For such a case, (8.72) would become

$$\lambda = \frac{1}{\nu_a \pi r^2} \quad \text{for stationary atoms.} \quad (8.73)$$

wire, and so we need only write $J = I/A$. It's now clear, from (8.68), that $J = \nu_e q v_d$. We also have

$$J = \frac{I}{A} = \frac{V}{RA} \stackrel{(8.66)}{=} \frac{\mathcal{E}\ell}{\varrho\ell} = \frac{\mathcal{E}}{\varrho}. \quad (8.67)$$

See also the discussion around (6.153).

But the lattice atoms are not stationary, because the lattice has a non-zero temperature T . Suppose an electron “sees” a lattice atom oscillating in the three spatial dimensions x, y, z : along each of these dimensions, the atom oscillates with amplitude $\mathcal{A}$ and angular frequency ω . At any given moment, the atom is displaced from its centre of oscillation by a time-dependent distance r_{osc} , which we’ll assume is somewhat greater than its radius. The atom then presents an average cross section to wandering electrons of

$$\sigma = \langle \pi r_{\text{osc}}^2 \rangle = \pi \langle x^2 + y^2 + z^2 \rangle = 3\pi \langle x^2 \rangle. \quad (8.74)$$

Recall the basic ideas of a classical harmonic oscillator, such as a mass attached to the end of a spring whose other end is fixed to a wall. For an appropriate choice of start time, the mass’s displacement x from equilibrium can be written as

$$x = \mathcal{A} \sin \omega t. \quad (8.75)$$

If the mass were a lattice atom, it would then present a cross section to wandering electrons of

$$\sigma = 3\pi \langle x^2 \rangle = 3\pi \langle \mathcal{A}^2 \sin^2 \omega t \rangle = 3/2 \pi \mathcal{A}^2, \quad (8.76)$$

since $\langle \sin^2 \omega t \rangle = 1/2$. What is $\mathcal{A}$? Note that the oscillator’s total energy is all kinetic when it passes through $x = 0$, at which point its speed is a maximum. This maximum speed v_{max} is the amplitude of the mass’s sinusoidal velocity $\dot{x} = \mathcal{A}\omega \cos \omega t$, and so is $\mathcal{A}\omega$. The mass’s total energy is then

$$1/2 m_a v_{\text{max}}^2 = 1/2 m_a \mathcal{A}^2 \omega^2. \quad (8.77)$$

The lattice atom is oscillating in three dimensions, and so its total energy is triple this value, or $3/2 m_a \mathcal{A}^2 \omega^2$. The equipartition theorem then says that

$$3/2 m_a \mathcal{A}^2 \omega^2 = 6/2 kT, \quad \text{or} \quad \mathcal{A}^2 = \frac{2kT}{m_a \omega^2}. \quad (8.78)$$

It follows that the atom presents a cross section to wandering electrons of

$$\sigma = 3/2 \pi \mathcal{A}^2 = \frac{3\pi kT}{m_a \omega^2}. \quad (8.79)$$

This combines with (8.72) to produce

$$\lambda = \frac{m_a \omega^2}{3\pi \nu_a kT} \quad \text{for oscillating atoms.} \quad (8.80)$$

To summarise the cases of both non-oscillating and oscillating lattice atoms, the classical mean free paths from (8.73) and (8.80) are

$$\lambda = \frac{1}{\nu_a \sigma} = \begin{cases} \frac{1}{\nu_a \pi r^2} & \text{lattice atoms are stationary spheres} \\ & \text{of radius } r, \\ \frac{m_a \omega^2}{3\pi \nu_a kT} & \text{lattice atoms are oscillators.} \end{cases} \quad (8.81)$$

The predictions of the resistivity ϱ are, from (8.71), (6.53), and (8.81),

$$\begin{aligned} \varrho &= \frac{2m_e \bar{v}}{\nu_e q^2 \lambda} \\ &= \frac{2m_e}{\nu_e q^2} \sqrt{\frac{8kT}{\pi m_e}} \times \begin{cases} \nu_a \pi r^2 & \text{lattice atoms are stationary spheres} \\ & \text{of radius } r, \\ \frac{3\pi \nu_a kT}{m_a \omega^2} & \text{lattice atoms are oscillators.} \end{cases} \end{aligned} \quad (8.82)$$

In particular, the electrical resistivity's temperature dependence is predicted to be (with the above references to lattice atoms abbreviated to “stationary spheres” and “oscillators”)

$$\varrho \propto \begin{cases} T^{1/2} & \text{stationary spheres,} \\ T^{3/2} & \text{oscillators.} \end{cases} \quad (8.83)$$

But the experimental result is $\varrho \propto T$, and neither stationary spheres nor oscillators predict this.

The resolution to the problem comes about by assuming electrons are fermions. It's then appropriate to replace the electrons' average Maxwell thermal speed $\bar{v}$ with a value similar to their Fermi speed v_F : say, αv_F where $\alpha \leq 1$. Equation (8.71) is then replaced by

$$\varrho = \frac{2m_e \alpha v_F}{\nu_e q^2 \lambda}. \quad (8.84)$$

Recall that $\bar{v}$ follows the Maxwell distribution (6.53), and so is proportional to $\sqrt{T}$; in contrast, the Fermi speed v_F given in (8.17) is a good measure of the electrons' average speed for all temperatures below 81,000 K for copper [recall (8.14)]. This makes v_F effectively independent of T .

If we treat the lattice atoms as oscillators instead of stationary spheres (in keeping with quantum mechanics), we obtain

$$\varrho = \frac{2m_e \alpha v_F}{\nu_e q^2} \times \frac{1}{\lambda} \stackrel{(8.80)}{=} \frac{2m_e \alpha v_F}{\nu_e q^2} \times \frac{3\pi \nu_a kT}{m_a \omega^2}. \quad (8.85)$$

With this choice of λ , we find that, indeed, $\varrho \propto T$, in accordance with experiment. This wouldn't work if we modelled the lattice atoms as stationary spheres, since λ [in (8.73)] would then have no T dependence.

What does (8.85) predict for copper's resistivity at 298 K? To simplify the calculation slightly, suppose that each lattice atom “releases” β valence electrons, so that $\nu_e = \beta\nu_a$. Equation (8.85) then becomes

$$\varrho = \frac{6\pi m_e \alpha v_F kT}{\beta q^2 m_a \omega^2}. \quad (8.86)$$

We will set $\alpha = 1$: slightly different choices of this correspond to different definitions of E_F to be found in the literature.⁴ The Fermi speed v_F was given in (8.17) as 1570 km/s. Each copper atom produces about one valence electron, so set $\beta = 1$. Copper's molar mass is 63.5 g; and we saw, in (7.14) from Einstein's model of heat capacity, that the copper lattice atoms vibrate with an angular frequency of about $\omega = 2\pi \times 6 \times 10^{12}$ Hz. Equation (8.86) becomes

$$\begin{aligned} \varrho &\simeq \frac{6\pi \times 9.11 \overset{-31}{=} \times 1 \times 1570 \overset{3}{=} \times 1.381 \overset{-23}{=} \times 298}{1 \times (-1.602 \overset{-19}{=})^2 \times 63.5 \overset{-3}{=}/6.022 \overset{23}{=} \times (2\pi \times 6 \overset{12}{=})^2} \Omega \text{ m} \\ &\simeq 29 \text{ n}\Omega \text{ m}. \end{aligned} \quad (8.87)$$

Experiments give a value of 17 n Ω m. Fine-tuning this calculation would call for a more accurate value of ω (since it appears squared here), and also for some adjustment in the choice of α . But we should not forget that the Fermi speed v_F was produced from the Fermi energy E_F , whose value was calculated in (7.132), which used the constant C in (7.114)—and this constant ultimately came from partitioning phase space into cells whose volume was set by Planck's constant h , as suggested by the Heisenberg Uncertainty principle. We have come such a long way from the subject's humble beginnings in Chapter 2 that we can only be impressed by the accuracy that the above calculation of resistivity has achieved.

It's useful to compare the typical speed $v_F = 1570$ km/s of the electrons with their drift speed v_d . We can calculate v_d by rearranging (8.68) to obtain

$$v_d = \frac{I}{\nu_e q A} = \frac{I}{\beta \nu_a q A}. \quad (8.88)$$

Consider a current $I = 1$ amp flowing through a copper wire of diameter 1 mm. As usual, set $\beta = 1$. Equation (7.57) gave the number density of copper atoms as $\nu_a = 8.47 \times 10^{28} \text{ m}^{-3}$. The electrons' drift speed is then

$$v_d = \frac{1}{1 \times 8.47 \overset{28}{=} \times 1.602 \overset{-19}{=} \times \pi \times (0.5 \overset{-3}{=})^2} \text{ m/s} \simeq 0.1 \text{ mm/s}. \quad (8.89)$$

Contrast this tiny *drift* speed of 0.1 millimetres per second with an electron's typical *individual* speed of 1570 *kilometres* per second: the electrons

⁴ See, for example, the comment just after (8.16). For that case, we might choose $\alpha = \sqrt{3/5} \simeq 0.77$.

are bouncing wildly back and forth within the copper lattice, but making very little forward progress in response to the applied voltage. We could also have calculated their drift speed directly from a given applied voltage V along a given length of wire ℓ . Recall, from (8.66), that

$$\frac{I}{A} = \frac{V}{\varrho \ell}. \quad (8.90)$$

This converts (8.88) into an expression involving voltage and wire length; but (8.90) does need the resistivity ϱ —either a measured value, or from (8.86).

8.3 Thermal Conductivity of Metals

In the previous section, we saw that by assuming electrons to be fermions, we can predict the correct temperature dependence of a metal's electrical conductivity. It turns out that the same can be said for a metal's *thermal* conductivity.

Recall the discussion in Section 6.9 of thermal conductivity and mean free path. There, we produced a classical expression (6.169) for a gas's thermal conductivity κ (or alternatively, its thermal resistivity $1/\kappa$) that contained a characteristic speed V . For the case of thermal conductivity of metals, the gas of Section 6.9 is presumably the valence electrons, in which case this characteristic speed might be set equal to the Fermi speed v_F , rather than the electrons' Maxwell mean speed $\bar{v}$. Also, the thermal conductivity in (6.169) involved the heat capacity—but we have learned that when it comes to metals, the molar heat capacity's classical (Dulong–Petit) value of $3R$ is not the full story.

To reiterate, the classical thermal conductivity κ in (6.169) is, with $V \equiv \bar{v}$,

$$\kappa = \frac{\nu_e \bar{v} \lambda}{3} \frac{C_V^{\text{mol}}}{N_A}. \quad (8.91)$$

We can show that this simply doesn't hold for a conductor, as follows. Assume the electrons follow a Maxwell distribution of speeds, so that their mean speed is (6.53), with that equation's m set equal to the electron mass m_e . The electrons' mean free path of interaction with the lattice atoms is given by (8.81). Hence, the thermal conductivity is, from (8.91),

$$\kappa = \frac{\nu_e}{3} \sqrt{\frac{8kT}{\pi m_e}} \frac{C_V^{\text{mol}}}{N_A} \times \begin{cases} \frac{1}{\nu_a \pi r^2} & \text{stationary spheres,} \\ \frac{m_a \omega^2}{3\pi \nu_a kT} & \text{oscillators.} \end{cases} \quad (8.92)$$

The classical prediction sets C_V^{mol} to be a constant in (8.92): either $3R$, or perhaps $3R + 3R/2$ if the valence electrons are treated using equipartition. This yields

$$\kappa \propto \begin{cases} T^{1/2} & \text{stationary spheres,} \\ T^{-1/2} & \text{oscillators.} \end{cases} \quad (8.93)$$

But experimentally, κ is found to be *independent* of temperature. Just as we did with electrical resistivity, the fix is to replace $\bar{v}$ in (8.91) with αv_F ; only we must now also give some thought to the precise value of C_V^{mol} . This quantity derives from Section 6.9's analysis of the flow of energy through a surface. Here, we consider this flow to arise from the movement of valence electrons, and so we replace C_V^{mol} with its electron contribution. That contribution is given, ostensibly, by (8.33)—whose more finely tuned version is (8.58), where R is the gas constant:

$$\text{“}C_V^{\text{mol}}\text{” becomes } C_V^{\text{mol}}(\text{electrons}) = \frac{\pi^2 R}{2} \frac{T}{T_F}. \quad (8.94)$$

With these choices, (8.91) becomes

$$\begin{aligned} \kappa &= \frac{\nu_e \alpha v_F \lambda}{3} \frac{C_V^{\text{mol}}(\text{electrons})}{N_A} = \frac{\nu_e \alpha v_F \lambda}{3} \frac{\pi^2 R}{2 N_A} \frac{T}{T_F} \\ &\stackrel{(8.81)}{=} \frac{\nu_e \alpha v_F}{3} \frac{\pi^2 k T}{2 T_F} \times \begin{cases} \frac{1}{\nu_a \pi r^2} & \text{stationary spheres,} \\ \frac{m_a \omega^2}{3 \pi \nu_a k T} & \text{oscillators.} \end{cases} \end{aligned} \quad (8.95)$$

For κ to be independent of temperature (as found experimentally), it's apparent that once again, we must choose the “oscillating lattice atoms” model. Set $\nu_e = \beta \nu_a$ as we did just before (8.86), and remember that

$$T_F = \frac{E_F}{k} = \frac{m_e v_F^2}{2k}. \quad (8.96)$$

Now (8.95) becomes

$$\begin{aligned} \kappa &= \frac{\beta \nu_a \alpha v_F}{3} \frac{\pi^2 k T}{2} \frac{2k}{m_e v_F^2} \frac{m_a \omega^2}{3 \pi \nu_a k T} \\ &= \frac{\beta \alpha \pi k m_a \omega^2}{9 m_e v_F}. \end{aligned} \quad (8.97)$$

For copper, similarly to (8.87), we obtain

$$\kappa \simeq \frac{1 \times 1 \times \pi \times 1.381 \stackrel{-23}{=} \times 63.5 \stackrel{-3}{=} / 6.022 \stackrel{23}{=} \times (2\pi \times 6 \stackrel{12}{=})^2}{9 \times 9.11 \stackrel{-31}{=} \times 1570 \stackrel{3}{=}} \text{ SI units}$$

$$\simeq 505 \text{ J K}^{-1} \text{ m}^{-1} \text{ s}^{-1}. \quad (8.98)$$

Compare this with the experimental value of around $401 \text{ J K}^{-1} \text{ m}^{-1} \text{ s}^{-1}$. The agreement with theory is impressive, with a comment applying here similar to the paragraph just after (8.87).

8.3.1 The Lorenz Number

Equations (8.71) and (8.91) together say that the electrical resistivity ϱ and the classical thermal conductivity κ satisfy

$$\varrho \propto 1/\lambda, \quad \kappa \propto \lambda. \quad (8.99)$$

Also, experiments indicate that $\varrho \propto T$ and κ is independent of T . This suggests that $\varrho\kappa/T$ might be a convenient quantity to work with, since it should be independent of λ and T . And, indeed, $\varrho\kappa/T$ is called the *Lorenz number*. (The fact that it is independent of temperature is called the *Wiedemann–Franz–Lorenz law*.) We have enough information now to calculate its value, both with a classical model and with a quantum-mechanical model. Equations (8.71) and (8.84) produce

$$\varrho = \begin{cases} \frac{2m_e \bar{v}}{\nu_e q^2 \lambda} & \text{classical,} \\ \frac{2m_e \alpha v_F}{\nu_e q^2 \lambda} & \text{quantum.} \end{cases} \quad (8.100)$$

Likewise, (8.91) and the first part of (8.95) yield

$$\kappa = \begin{cases} \frac{\nu_e \bar{v} \lambda}{3} \frac{C_V^{\text{mol}}}{N_A} & \text{classical,} \\ \frac{\nu_e \alpha v_F \lambda}{3} \frac{C_V^{\text{mol}}(\text{electrons})}{N_A} & \text{quantum.} \end{cases} \quad (8.101)$$

The product of (8.100) and (8.101) is

$$\varrho\kappa = \frac{2m_e}{3N_A q^2} \times \begin{cases} \bar{v}^2 C_V^{\text{mol}} & \text{classical,} \\ \alpha^2 v_F^2 C_V^{\text{mol}}(\text{electrons}) & \text{quantum.} \end{cases} \quad (8.102)$$

Now recall that

$$\bar{v}^2 \stackrel{(6.53)}{=} \frac{8kT}{\pi m_e}, \quad v_F^2 \stackrel{(8.17)}{=} \frac{2kT_F}{m_e}, \quad (8.103)$$

and

$$C_V^{\text{mol}} = 3R \text{ (Dulong-Petit)}, \quad C_V^{\text{mol}}(\text{electrons}) \stackrel{(8.58)}{=} \frac{\pi^2 RT}{2T_F}. \quad (8.104)$$

With these, (8.102) returns a Lorenz number of

$$\frac{\varrho\kappa}{T} = \frac{2k^2}{3q^2} \times \begin{cases} 24/\pi & \text{classical,} \\ \pi^2\alpha^2 & \text{quantum.} \end{cases} \quad (8.105)$$

In the classical case, the distinguishing factor in (8.105) is $24/\pi \simeq 7.6$. In the quantum case with the choice $\alpha = 1$, the distinguishing factor is $\pi^2 \simeq 9.9$; and if, instead, we set $\alpha = \sqrt{3/5}$ [as suggested in the footnote just after (8.86)], the quantum distinguishing factor is $\pi^2 \times 3/5 \simeq 5.9$. In other words, the classical and quantum predictions of the Lorenz number are roughly the same. Before the advent of quantum mechanics, physicists were unsettled by the fact that whereas the classical prediction of the Lorenz number was verified experimentally, predictions of its individual factors ϱ and κ differed from experiment. We see now that the agreement of the *classical* Lorenz number with experiment was quite fortuitous.

8.4 Insulators and Semiconductors

The above treatment of the cloud of valence electrons as a gas of non-interacting fermions took us some way toward incorporating quantum mechanics into the study of electrical conduction. But the quantised values of atomic energy that are so necessary to the discussion of the states available to this gas of electrons are only a first approximation to the actual energy levels of a metal. To dig deeper, we must consider the effect of the entire crystal lattice on the values of the electron energy levels.

Solving Schrödinger's equation for a single electron in a lone atom yields a discrete set of energy levels; but, when we bring two atoms close together, the energy of each level changes due to the influence of the other atom. Figure 8.3 shows how, when many atoms are brought closer and closer together to form a crystal lattice, the simple set of energy levels $E_1, E_2, E_3, \dots$ available to all the atoms in isolation splits into multiple closely spaced levels. ("Read" the figure from right to left, beginning with a large interatomic spacing, and then reducing that spacing.) This process is usually explored by solving Schrödinger's equation for electrons moving in a periodic potential. Bringing N atoms together to form a lattice will cause each energy level to split into N levels, forming a closely spaced set of levels known as a *band*. The energy width of this band of N levels is independent of N , being determined

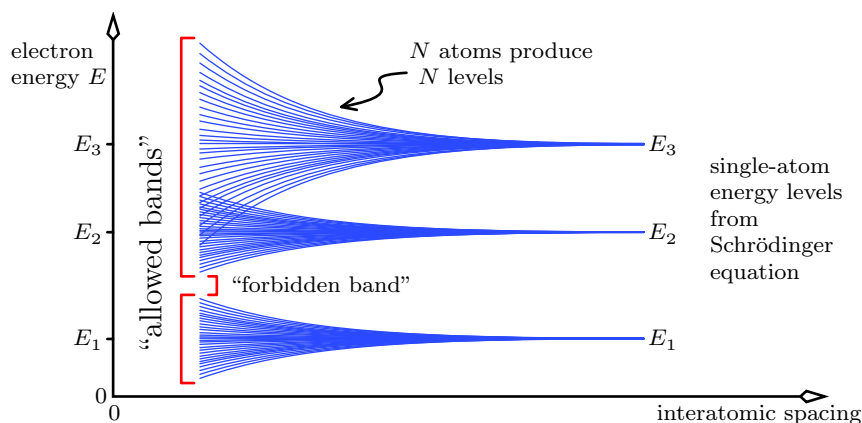


Fig. 8.3 The simple energy levels $E_1, E_2, E_3, \dots$ predicted by Schrödinger's equation for a lone atom split into a band of N levels when N atoms are present. The width of this band is determined by the atomic spacing, and not by N ; so, for large N , a band is an almost continuous spread of energy levels

instead by the atomic spacing. It follows that for large N , the energy levels in a band crowd together to become an almost continuous spread.

Bands are typically a few electron volts wide, and may overlap. The word “band” is used to refer not only to a region of energy populated by energy levels (which is called an *allowed band*), but also to a region that has *no* energy levels (called a *forbidden band*).

One piece of direct evidence for the existence of these bands of energy levels in solids comes from X-ray spectra. For example, the spectrum of X rays emitted by gaseous sodium shows the expected sharp peaks due to energy-level quantisation, but the same peaks produced from solid sodium are broadened. We interpret this broadening to result from a band structure of the energy levels in the solid sodium.

Figure 8.4 shows various ways in which bands can be populated by electrons. Each bar is a one-dimensional axis of energy increasing to the right, and is given a non-zero height only for readability.

1. The top bar in the figure shows a conductor of electricity. Here, the highest-energy electrons occupy only part of an allowed band, shown in red. This means the remaining (neighbouring) empty energy levels of that band (coloured blue) are available for electrons to move into easily. Those electrons can then carry energy through the lattice, and this manifests as electric current. The band containing these highest-energy electrons is called the *conduction band*.
2. The middle bar in the figure shows an insulator. Its highest-energy band is full of electrons (meaning all of that band's energy levels are occupied), and is called the *valence band*. The width of the forbidden band immedi-

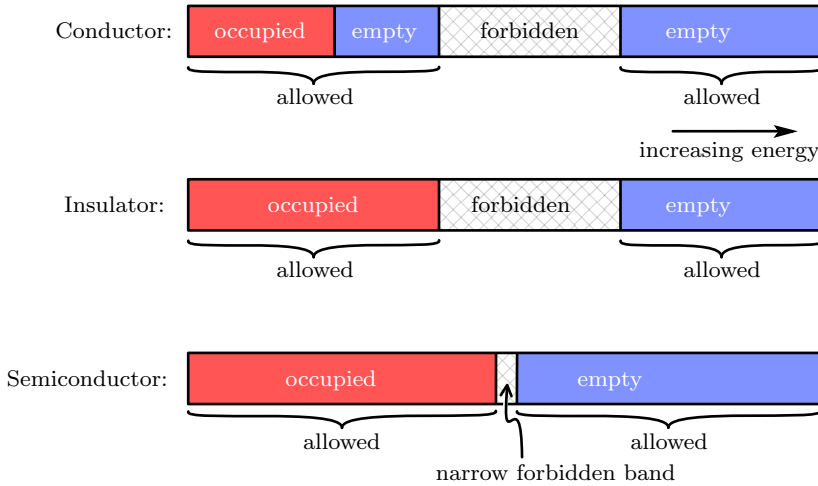


Fig. 8.4 A comparison of band placements and widths for different materials. The height of each bar has no significance. Electrical conductors have a highest-energy band that is only partially occupied by electrons, giving those electrons plenty of scope to move into higher energy levels and carry energy through the lattice. In contrast, the highest-energy band of an insulator is completely occupied, with a large forbidden band that suppresses electrons from being “bumped” (via their interactions) into the empty levels of the next allowed band. If the width of this forbidden band is sufficiently small, some electrons can jump the gap, and the material is a semiconductor

ately above the valence band is comparatively large (greater than about 2 eV), which prevents electrons from jumping across it into the next allowed band. Their motion is thus restricted, and no electric current flows. In diamond, the width of this forbidden band is 7 eV, which makes diamond a strong insulator.

3. The bottom bar in the figure shows a *semiconductor*. Here, the forbidden band immediately above the valence band has a width of less than about 2 eV, allowing a “few” electrons to jump across into the next allowed band. A semiconductor will thus support a small amount of electric current. Well-known semiconductors are silicon (forbidden-band width 1.1 eV) and germanium (0.7 eV).

How does the proportion of electrons found in the conduction band differ for insulators versus semiconductors? Refer to Figure 8.5, which shows a simple band structure for an insulator or semiconductor, on which is superimposed the occupation number $\bar{n}(E, T)$. We assume that the Fermi energy E_F lies halfway between the valence and conduction bands. Write E_c for the energy at the base of the conduction band, and call the width of the forbidden band the “gap width” E_g . We will calculate the number of electrons N_e found in the conduction band as a fraction of total number of electrons, $N = \frac{2}{3} C E_F^{3/2}$ in (7.131). The number of electrons in the conduction band is

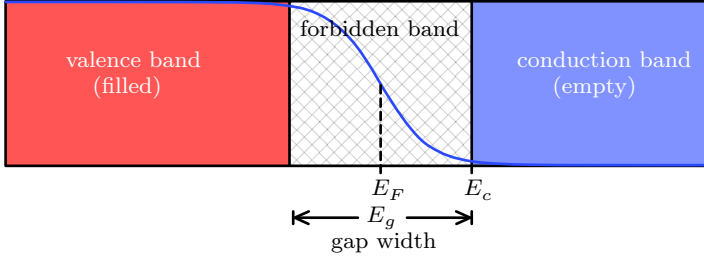


Fig. 8.5 The idealised band structure of an insulator or semiconductor, over which is superimposed a plot of the electron occupation number $\bar{n}(E, T)$ at room temperature. The gap width E_g of an insulator's forbidden band is large, whereas E_g of a semiconductor is small. The energy at the base of the conduction band is labelled E_c . The Fermi energy E_F is assumed to lie halfway between the valence and conduction bands, in which case $E_c - E_F = E_g/2$

$$N_e \stackrel{(7.32)}{=} \int_{E_c}^{\infty} \bar{n}(E, T) g(E) dE = \int_{E_c}^{\infty} \frac{g(E) dE}{e^{\beta(E-E_F)} + 1}. \quad (8.106)$$

The value of $E_c - E_F = E_g/2$ tends to be about one electron volt, whereas at room temperature, $kT \simeq 1/40$ eV. For energies E in the conduction band, it follows that $(E - E_F)/(kT) \gg 1$, and hence $e^{\beta(E-E_F)} \gg 1$. So, we ignore the “+ 1” in the denominator in (8.106), and thus replace the integral with a more manageable version:

$$N_e \simeq e^{\beta E_F} \int_{E_c}^{\infty} e^{-\beta E} g(E) dE. \quad (8.107)$$

As an idealised example, set $g(E) = C\sqrt{E}$ from (7.114):

$$N_e \simeq C e^{\beta E_F} \int_{E_c}^{\infty} e^{-\beta E} \sqrt{E} dE. \quad (8.108)$$

Apply a change of variables $x \equiv \sqrt{E}$ here, giving $\sqrt{E} dE = 2x^2 dx$. This converts (8.108) to

$$N_e \simeq 2C e^{\beta E_F} \int_{\sqrt{E_c}}^{\infty} x^2 e^{-\beta x^2} dx. \quad (8.109)$$

Refer to (1.97) to obtain

$$N_e \simeq \frac{C e^{\beta E_F}}{\beta^{3/2}} \left[\sqrt{\beta E_c} e^{-\beta E_c} + \frac{\sqrt{\pi}}{2} \operatorname{erfc} \sqrt{\beta E_c} \right]. \quad (8.110)$$

Recall that $E_c \gg kT$ (i.e., $\beta E_c \gg 1$). This lets us call on the large- x approximation of $\operatorname{erfc} x$ in (1.113), to write (8.110) as

$$\begin{aligned}
N_e &\simeq C e^{-\beta(E_c - E_F)} kT \sqrt{E_c} \left[1 + \frac{1}{2\beta E_c} \right] \\
&\simeq C e^{-\beta(E_c - E_F)} kT \sqrt{E_c}.
\end{aligned} \tag{8.111}$$

Now use $N = 2/3 C E_F^{3/2}$, to write (while remembering that $E_c - E_F = E_g/2$)

$$\frac{N_e}{N} \simeq \frac{3kT}{2E_F} \sqrt{\frac{E_c}{E_F}} \exp \frac{-E_g}{2kT}. \tag{8.112}$$

To gain a feel for the individual terms here, note that $kT/E_F \ll 1$, and $E_c/E_F \approx 1$. We see that the population N_e in the conduction band decreases exponentially with the gap width E_g .

In practice, the above idealised state density $g(E) = C\sqrt{E}$ overestimates the density of states in the conduction band, and what is more usual is to use $g(E) = C\sqrt{E - E_c}$. Equation (8.107) then becomes

$$N_e \simeq C e^{\beta E_F} \int_{E_c}^{\infty} e^{-\beta E} \sqrt{E - E_c} dE. \tag{8.113}$$

A change of variables $x \equiv \sqrt{E - E_c}$ renders this expression as

$$N_e \simeq 2C e^{\beta(E_F - E_c)} \int_0^{\infty} x^2 e^{-\beta x^2} dx. \tag{8.114}$$

Refer to Figure 8.5 to see that $E_F - E_c = -E_g/2$, and so

$$N_e \simeq 2C e^{-\beta E_g/2} \int_0^{\infty} x^2 e^{-\beta x^2} dx \stackrel{(1.98)}{=} \frac{C}{2\beta} \sqrt{\frac{\pi}{\beta}} e^{-\beta E_g/2}. \tag{8.115}$$

Now introduce $N = 2/3 C E_F^{3/2}$, obtaining

$$\frac{N_e}{N} \simeq \frac{3\sqrt{\pi}}{4} \left(\frac{kT}{E_F} \right)^{3/2} \exp \frac{-E_g}{2kT}. \tag{8.116}$$

This differs from (8.112) mainly by having an extra factor of $\sqrt{kT/E_F}$.

Conduction Electrons in Insulators and Semiconductors

Calculate N_e/N at room temperature for two materials, each of which has the Fermi energy of copper. One material's forbidden band has width 5 eV (an insulator) and the other's has width 1 eV (a semiconductor).

We use $kT = 1/40$ eV and $E_F = 7$ eV:

Insulator: $E_g = 5$ eV, so (8.116) is

$$\frac{N_e}{N} \approx \frac{3\sqrt{\pi}}{4} \left(\frac{1/40}{7} \right)^{3/2} \exp \frac{-5}{1/20} \approx 10^{-47}. \quad (8.117)$$

Semiconductor: $E_g = 1 \text{ eV}$, so

$$\frac{N_e}{N} \approx \frac{3\sqrt{\pi}}{4} \left(\frac{1/40}{7} \right)^{3/2} \exp \frac{-1}{1/20} \approx 10^{-12}. \quad (8.118)$$

The semiconductor has a comparatively *much* larger number of conduction electrons than the insulator. This number also increases with temperature, as is evident from (8.116); materials that are insulators at low temperatures can become semiconductors as their temperature rises.

8.5 Diodes

The dependence of electrical conductivity on temperature that we saw in (8.116) makes semiconductors useful in technology. But modern technology has gone further: it has been able to change the conduction properties of semiconductors drastically by adding minuscule amounts of elements known as *impurities*. These impurities create extra energy levels, which can either accept or donate excited electrons. Thus, the impurities cause semiconductors to conduct more strongly than if the conduction were due to thermal excitation alone. Adding such impurities is called “doping” the semiconductor.

When a valence electron inside a semiconductor is excited by some influence that makes it jump across the small gap from the valence band into the next allowed band, it leaves behind an empty state, or “hole”, in the valence band. This hole can be treated as though it were a positively charged particle, and it contributes to the overall electric current in the semiconductor. Doping the semiconductor with an impurity will cause more or less free electrons *or* holes to move throughout the lattice, where they effectively orbit an atom of the impurity at a large distance.

Doped semiconductors come in two basic flavours, shown in Figure 8.6.

- Shown at left in the figure, a *p-type* semiconductor is one that has been doped with an element such as gallium or indium (typically “Group 3” of the Periodic Table), which accepts electrons or, equivalently, donates holes (holes have positive charge—hence the name p-type). The doping element creates new energy levels at the *bottom* of the forbidden band. These new energy levels are then populated by the donated holes. Thermal excitation or an external electric field can now nudge electrons from the top of the valence band into these holes, resulting in a dramatic increase in conductivity.

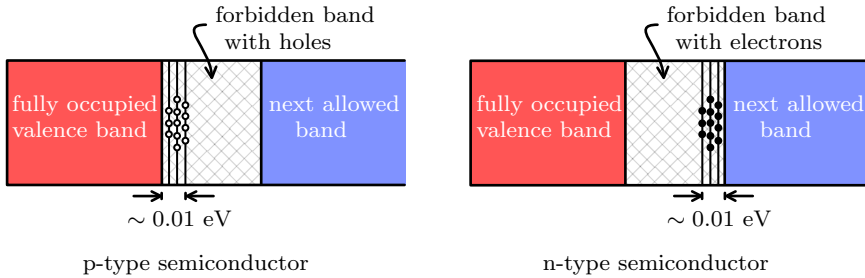


Fig. 8.6 **Left:** A p-type semiconductor is a semiconductor that is doped with an impurity that *accepts* electrons. This is equivalent to “donating holes” into new energy levels at the *bottom* of the forbidden band. Valence electrons can then easily jump from the valence band to fill these holes in the new energy levels, and this increases the conductivity of the semiconductor. **Right:** An n-type semiconductor is a semiconductor that is doped with an impurity that *donates* electrons, which are placed into new energy levels at the *top* of the forbidden band. These electrons can then easily jump into the next allowed band, which increases the conductivity of the semiconductor

- At right in Figure 8.6, an *n-type* semiconductor is one that has been doped with an element such as arsenic or antimony (typically a “Group 5” element of the Periodic Table), which donates electrons (electrons have negative charge—hence the name n-type). The doping element creates new energy levels at the *top* of the forbidden band, which are then populated by the donated electrons. Thermal excitation or an external electric field can nudge these electrons into the next allowed band, causing a dramatic increase in conductivity.

By applying an external electric field, we can effectively switch on and off the ability of a doped semiconductor to conduct. This makes it very useful in electronic circuits, since there, we do have this ability to alter the external electric field. For example, suppose we take two doped semiconductors: the p-type has wandering donated holes, and the n-type has wandering donated electrons. As shown in Figure 8.7, we now bond them together, creating a *pn-semiconductor*. This is one example of a *diode*: here, some of the donated wandering electrons in the n-type will cross the junction of the faces to fill donated wandering holes in the p-type semiconductor. The result is a slight excess of negative charge on the p-type side, and a slight excess of positive charge on the n-type side. This produces a permanent internal electric field across the junction, pointing from n-type to p-type.

Two processes now take place continuously across the junction:

1. *Thermal* fluctuations constantly nudge mobile donated electrons from the n-type side to join the excess negative charge on the p-type side. This acts to increase the strength of the electric field across the junction. The

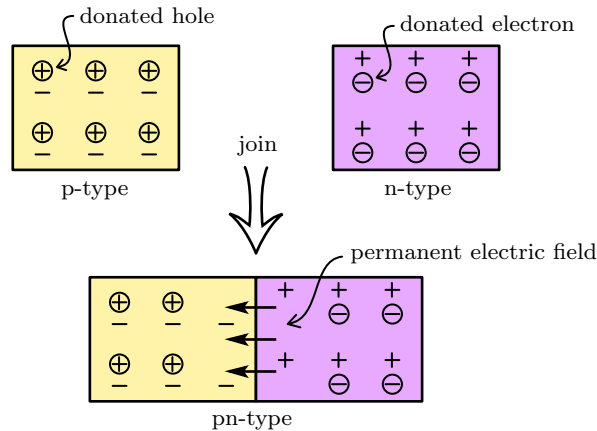


Fig. 8.7 When a p-type semiconductor is physically joined to an n-type semiconductor, some of the mobile donated electrons in the n-type move to fill some of the mobile donated holes in the p-type. This creates a permanent electric field across the junction

Boltzmann distribution can be applied to determine how many mobile electrons are thermally nudged in this way.

- Those electrons that were thermally nudged into the p-type now experience *electrostatic* repulsion from the negatively charged p-type, and attraction to the positively charged n-type. This makes them return home across the junction. This process acts to reduce the field's strength.

These two processes are in equilibrium. The thermal fluctuations do work on the electrons: they push the electrons away from the slightly positive n-type side of the junction and into the slightly negative p-type side of the junction, and so push the electrons against forces from both sides that are trying to return them to their start point. This work done by the fluctuations on the electrons increases their potential energy. The energy term to be used in the Boltzmann distribution is this increase in potential energy that the n-type's free electrons “see” from where they sit on the n-side of the junction, before they are thermally nudged across.

We might ask: should the Boltzmann energy term be the potential energy “seen” on the far side of the junction by the n-type's free electrons, or should it be the potential energy *increase* that they see across the junction? In fact, either choice will give the same result. Potential energy is only ever defined up to an additive constant, and so the electrons can only “know” of the increase in potential energy. We are free to include this additive constant in the Boltzmann exponential in (5.5), but it will only get absorbed into the constant of proportionality in that equation. So, we might as well set the additive constant to be that which causes the potential energy of a free

electron on the n-type side of the junction to be zero. That electron then sees a potential energy on the p-type side of some $U_0 > 0$.

We can picture the continuous thermal and electromagnetic flows of electrons to and fro across the junction as follows. First, it's worth making the point that electrons' negative charge can be confusing when we wish to picture the "conventional current" of circuit theory, since conventional current is the flow of *positive* charge. To ameliorate that, in the following paragraphs, we'll treat the flow of electrons as a particle current K , which is the number of electrons passing some point per unit time.⁵ This flow of charge produces a conventional electric current I of *charge* flow per unit time: $I = -eK$, where $e \simeq 1.602 \times 10^{-19}$ C. We will set the positive direction of both I and K to be from left to right in Figure 8.7—that is, from the p-type to the n-type. Consider the two flows of electrons:

1. Thermal flow: a *particle* current $K_{\text{thermal}} < 0$ of electrons going from n to p (due to thermal fluctuations) boosts these electrons into the higher potential energy U_0 . The number of electrons forming the current obeys the Boltzmann distribution, which closely approximates the tail of the Fermi–Dirac distribution for the electron occupation number at energies above the forbidden band, meaning greater than the Fermi energy E_F .
2. Electrostatic repulsion: this *particle* current of electrons $K_0 > 0$ flows from p to n across the junction, driven by the permanent internal electric field that was created when the p-type and n-type semiconductors were originally fused together.

In equilibrium, the total particle current is zero:

$$K_{\text{thermal}} + K_0 = 0 \quad (\text{equilibrium}). \quad (8.119)$$

Suppose that we now introduce a *bias voltage*, by connecting the p-type semiconductor to the positive terminal of an electric battery rated at voltage $V_b > 0$, and connecting the n-type to the negative terminal of that battery. This choice of connection is called "forward-biasing" the diode, and is shown in Figure 8.8. This bias voltage introduces an external electric field that opposes the permanent field in Figure 8.7. The external field lowers the barrier to thermal motion of the n-type's free electrons; in essence, they are now being attracted to the positive terminal of the battery. Now, K_{thermal} no longer balances K_0 . Note that this external field doesn't change K_0 , which is a kind of ever-present background current arising from the base conditions existing internally to the junction in Figure 8.7. The potential energy of an electron on the p side of the junction has now decreased from U_0 to $U_0 - eV_b$. Because

⁵ To match " I " for electric current, I could have called this particle current " J " instead of " K ". But J is often used in flow calculations to denote a flow per unit area; see, for example, the footnote around (8.67), and Section 9.8 ahead. There is no need to introduce a cross-sectional area into the present diode discussion, and so I have avoided the possible confusion of writing J for a flow of particles.

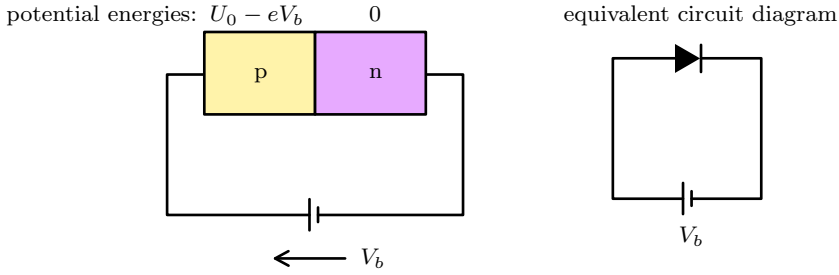


Fig. 8.8 Left: A forward-biased diode. The addition of the battery of voltage $V_b > 0$ means that electrons in the p-type are now attracted to the positive terminal of the battery, which makes them more inclined, so to speak, to want to stay on the p side of the junction. The effect of this is to lower their potential energy from U_0 to $U_0 - eV_b$. **Right:** The equivalent circuit diagram

(8.119) no longer holds, an electric current arises. This is, remembering that conventional current is positive for left-to-right flow of positive charge,

$$\begin{aligned}
 I &= \text{conventional electric current through diode from p-type to n-type} \\
 &= -e \times \text{electron particle current through diode from p to n} \\
 &= -e(K_{\text{thermal}} + K_0).
 \end{aligned} \tag{8.120}$$

But, as discussed earlier, the Boltzmann distribution says that for some normalisation $\mathcal{N}$,

$$\begin{aligned}
 K_{\text{thermal}} &\propto \exp \frac{-\Delta \text{potential energy (n} \rightarrow \text{p)}}{kT} \\
 &= \mathcal{N} \exp \frac{-(U_0 - eV_b)}{kT}.
 \end{aligned} \tag{8.121}$$

What is this normalisation $\mathcal{N}$? With no battery, (8.119) holds; and since then $V_b = 0$, we can write

$$-K_0 \stackrel{(8.119)}{=} K_{\text{thermal}} \stackrel{(8.121)}{=} \mathcal{N} \exp \frac{-U_0}{kT}. \tag{8.122}$$

K_0 is unchanged by whether or not a battery is present, and so (8.121) can be written as

$$K_{\text{thermal}} = -K_0 \exp \frac{eV_b}{kT}. \tag{8.123}$$

The total current I is then, from (8.120) and (8.123),

$$I = -e(K_{\text{thermal}} + K_0) = -e \left(-K_0 \exp \frac{eV_b}{kT} + K_0 \right)$$

$$= eK_0 \left(\exp \frac{eV_b}{kT} - 1 \right). \quad (8.124)$$

Recall that $K_0 > 0$. This current is plotted as a function of the bias voltage in Figure 8.9. When the diode is forward-biased as we have discussed above, the electrons have a decreased potential gap to jump thermally (that is, they are being pulled toward the battery's positive terminal), and the resulting current I through the diode can be very large. I tends to be several amps for $V_b = +1$ volt. (Forwarding-biasing a commercial diode at high voltage is likely to burn it out very quickly.)

But the real marvel occurs when we decide to *reverse* the battery polarity: that is, we connect the p-type to the negative battery terminal, and (of course) connect the n-type to the positive battery terminal. This has the effect of making V_b negative in the above discussion. This is called *reverse-biasing* the diode. It's now even more difficult for the electrons to be thermally bumped into the p-type semiconductor than when no bias was present. This effectively shuts off the thermal current K_{thermal} from (8.124), and then only a residual current $-eK_0$ flows. This residual current is typically around 1 or 2 milliamps—even when the diode is reverse-biased to several hundred volts. (Commercial diodes can be quite robust to being reverse-biased.)

Diodes thus pass current almost entirely in one direction only, and this makes them very useful in electronic devices. For example, they are used as *rectifiers* to convert alternating current to direct current.

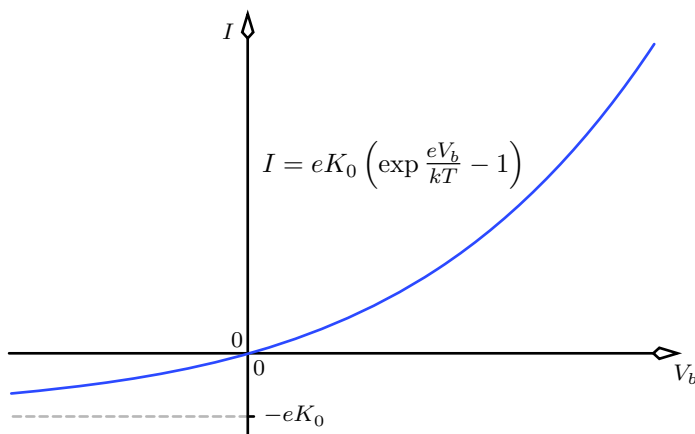


Fig. 8.9 The p-to-n conventional current I in a diode as a function of the forward-bias voltage V_b , from (8.124)

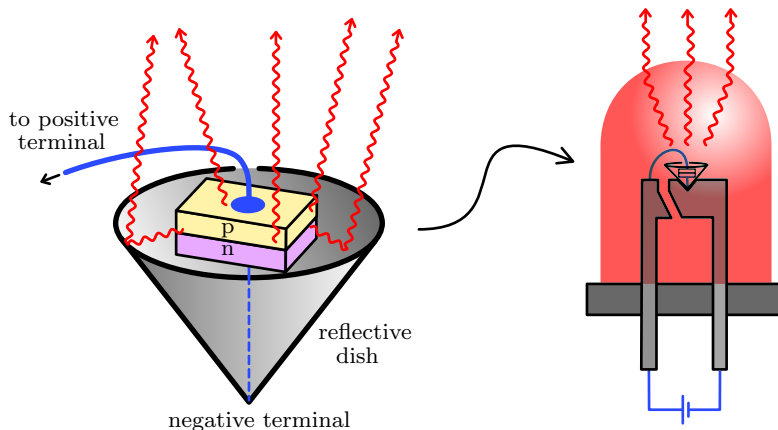


Fig. 8.10 An LED consists of a forward-biased pn-semiconductor placed inside a reflective dish. The whole is encased in a plastic lens that directs the light upward in a well-defined beam

Light-Emitting Diodes

In a forward-biased diode, the positive conventional current flowing in Figure 8.9 corresponds to electrons moving from right to left in Figure 8.8, filling holes in the p-type semiconductor as they move across the junction. Referring to Figure 8.6, we see that these electrons are dropping from states of higher energy to states of lower energy.

This drop to lower-energy states is accompanied by a release of photons. When the diode is encased in a structure that directs these photons more or less into a single direction, it is called a *light-emitting diode*, or LED. A common type of LED is shown in Figure 8.10.⁶

The use of different materials to build semiconductors gives many choices of the width of the forbidden band in Figure 8.6. Standard silicon diodes have a forbidden-band width of 1.1 eV; for these, the drop in energy is relatively small, leading to infra-red light being emitted. This is invisible to the human eye, and such LEDs are used in remote-control devices, such as for changing channels on a television. When the width of the forbidden band is about 2.5 eV, visible light of energy E_γ is produced, whose wavelength is

$$\lambda = \frac{c}{f} = \frac{hc}{hf} = \frac{hc}{E_\gamma} = \frac{6.626 \times 10^{-34} \times 2.998 \times 10^8}{2.5 \times 1.602 \times 10^{-19}} \text{ m} \simeq 496 \text{ nm}. \quad (8.125)$$

⁶ A red LED emits light that is reasonably monochromatic. Why, then, are many—but not all—of the plastic casings for such LEDs coloured red? I don't know. Perhaps the colour of the plastic just serves to identify the LED when more than one colour is present.

LEDs are highly efficient at converting energy to light, and so produce very little wasteful heat. They last much longer, are much smaller, and are more rugged than conventional light bulbs. Not only can modern monochromatic LEDs be produced in all the colours of the spectrum (including infra-red and ultraviolet), but they can also emit intense *white* light. Powered by small batteries and used as long-lasting head-mounted lamps, these LEDs have transformed our ability to work outdoors at night, and to explore dark caves.



Chapter 9

Boson Statistics in Blackbody Radiation

In which we apply the statistical mechanics of bosons to study the thermal noise produced by electrical components, make a brief foray into radar theory and signal processing, and examine the spectra of light produced by ovens and hot bodies. We conclude with studies of the cosmic microwave background, the greenhouse effect, and lasers.

We are all acquainted with the observation that many materials glow when heated. This light is produced when excited atoms drop between energy levels, and when charges accelerate as they bounce among lattice atoms. The electromagnetic theory of just how a flame's hot gas emits light is complex; but, given that a huge number of charges give rise to the emitted light, it can be expected that statistical mechanics can be used to investigate the phenomenon.

But the emission of light from a hot object is not an equilibrium process. Energy is continuously transformed from the object to the light waves, which can then escape. We have only considered equilibrium processes in this book, and the subject of non-equilibrium processes is an advanced branch of statistical mechanics. Nonetheless, the spectrum of light emitted from a hot object can be predicted by considering a related process that *does* occur in equilibrium. The idea that makes this connection is the *principle of detailed balance*. Because accelerating charges emit light, if an object has charges that resonate at some particular frequency, then not only will it readily *emit* light of that frequency, but it will also readily *absorb* light of that frequency. The principle of detailed balance postulates that the object's abilities to emit and absorb are identical:

The Principle of Detailed Balance

When an object is in thermal equilibrium with a bath of electromagnetic waves, then regardless of its colour or makeup, it emits the same spectrum and intensity that it absorbs.

This principle suggests that we can predict the spectrum of light *emitted* by a hot object via examining the light *absorbed* by that hot object. We can begin to do that by considering a closed system, since that enables equilibrium statistical mechanics to be used in the analysis. We can imagine wrapping

the hot object around itself to form a closed cavity—an oven—and analyse how the radiation inside the oven interacts with the oven walls. These walls are assumed to be in equilibrium with the radiation. After using equilibrium statistical mechanics to examine the spectrum inside the oven, we will transfer what we have learned to predict the spectrum of a glowing hot object.

Consider, then, a perfectly absorbing object placed in an oven that is ideal in the sense that this oven is “perfectly emitting” in its interior (an idea that we’ll examine in the next section). In the limit of low temperatures, such a perfectly absorbing object will be black, and so is conventionally called a *black body*. The black body must emit exactly what it absorbs; but, by definition, it absorbs all of the radiation it receives from the oven. Hence, it must emit the same spectrum that the oven produces when everything is in equilibrium. But, presumably, the mechanism for how the body emits doesn’t depend on the oven; and thus, the black body must emit identically when *outside* the oven. We conclude that the spectrum of frequencies produced by a black body equals that found inside an ideal oven.¹

9.1 Spectrum of Radiation Inside an Oven

Our main task is to calculate the *spectral energy density* $\varrho(f)$ of the radiation in an oven:

$$\varrho(f) \equiv \left[\begin{array}{l} \text{amount of electromagnetic energy per unit} \\ \text{frequency } f, \text{ and per unit oven volume} \end{array} \right]. \quad (9.1)$$

(This is also a function of temperature, but—for brevity—we won’t indicate that temperature dependence explicitly throughout this chapter.) Different hot materials emit different amounts of each wavelength, which means we cannot hope to use only general arguments to obtain the spectral energy density of an oven made from any arbitrary material. Also, we cannot expect to discuss the emission of very low frequencies (long wavelengths) from a generic oven; the reason for this is that it’s problematic to analyse electromagnetic radiation whose wavelength is longer than the characteristic size of the body producing it.

So, the central question is: what is the spectrum of electromagnetic frequencies inside the oven? It is sometimes argued that the electric field inside a metal oven will go to zero at the walls, since, otherwise, wall currents would be generated, which would then eliminate the field at the walls. Requiring all

¹ You will often find it written that any light entering a small hole in the side of a large oven will never leave, and thus the oven is a perfectly absorbing body, and hence supposedly must also be a perfectly emitting body. I see no content in this compact line of reasoning; the actual analysis using the principle of detailed balance makes no mention of light entering the hole in the oven wall.

individual waves to have nodes at the walls would force only certain frequencies to be present—although, for all practical purposes in a real oven, these would still approximate a continuum of frequencies. But the field in a ceramic oven need not go to zero at the walls, and so the frequencies present need not be quantised in this way. On the other hand, if the walls inside an oven are reflective enough that a light wave inside bounces back and forth many times, it will be reinforced if a whole number of wavelengths fit into a round trip. Different ovens will have different amounts of internal reflectivity, and different-sized and different-shaped ovens will reinforce some wavelengths but not others.

The task of calculating an oven's spectral energy density is starting to look difficult! To make progress, we consider an idealised oven that holds a continuum of wavelengths. Its wall oscillators produce light of all frequencies. This light bounces about inside the oven, sometimes reflected and sometimes absorbed, ensuring that the spread of frequencies tends quickly toward some equilibrium distribution.

Consider the following argument for why the oven's shape might conceivably have no bearing on its spectrum. Join two differently shaped ovens at the same temperature, allowing radiation to pass between them via a small hole in a wall of each. Suppose the radiation spectra of the two differed around some particular frequency (say, yellow light): one oven produced more yellow light than the other. Then, presumably we could introduce a filter across the hole, that passed only yellow light. That would allow a flow of energy in one direction through the hole, which would presumably act to "unequalise" the temperatures. But it's unreasonable for the system to depart from thermal equilibrium in such a way—it would allow us to "break" the Second Law of Thermodynamics. So, we might conclude that there can be no such one-way flow of energy, and infer that an oven's shape doesn't affect the spectrum of radiation inside.

Actually, this argument is not quite as straightforward as it might appear. Yes, the filter would pass yellow light into the oven whose walls did not naturally emit much yellow light; but the principle of detailed balance says that those walls would not *absorb* much yellow light either, in which case that oven's temperature would presumably not increase. Would the yellow light then build up inside that oven, perhaps interacting with the filter to heat it up until the filter broke down? The hot filter will also radiate, and its emission might be related to the yellow wavelength that it has been designed to allow through. These are difficult questions with no clear answers in sight, and so we will simply postulate that an "idealised" oven's spectral energy density is independent of its shape. Ultimately, we will appeal to experiment for validation. We make the following assumptions of an idealised oven:

Assumptions of an Idealised Oven

- the oven walls are continuously emitting and absorbing radiation,
- the oven’s shape doesn’t affect its spectrum,
- there is no restriction on what frequencies can exist inside the oven,
- the walls contain a huge number of quantised harmonic oscillators, with each wall particle being associated with two such oscillators: one oscillator for each of the two orthogonal directions of the particle’s oscillation along the wall’s surface,
- at thermal equilibrium, the energy of the oven radiation in one “wave state” (defined soon) at frequency f equals the mean energy $\varepsilon(f)$ of a single wall oscillator with frequency f .

The last assumption above is by no means obvious, especially as it depends wholly on how we define a “wave state”. Once we do have such a definition, we can apply the usual statistical mechanical idea of computing a density of such wave states $g(f)$ as a function of, say, frequency. Start by writing

$$\varepsilon(f) = \frac{\text{total energy of radiation in } f \text{ to } f + df}{\text{number of wave states in } f \text{ to } f + df} = \frac{\varrho(f) df V}{g(f) df}. \quad (9.2)$$

The spectral energy density $\varrho(f)$ is then

$$\varrho(f) = \frac{\varepsilon(f) g(f)}{V}. \quad (9.3)$$

We must calculate the mean energy $\varepsilon(f)$ of a wall oscillator of frequency f , and define and calculate the radiation’s density of wave states $g(f)$.

9.1.1 Mean “Extractable” Energy of an Oscillator, $\varepsilon(f)$

At first sight, it’s reasonable to invoke the equipartition theorem to say that each oscillator’s energy will equal $1/2 kT$ times its number of quadratic energy terms. Each oscillator has two such terms (one kinetic and one potential), leading to $\varepsilon(f) = 1/2 kT \times 2 = kT$.

This value of $\varepsilon(f) = kT$ turns out to be completely wrong for a standard hot oven, as we’ll see later in the discussion around (9.44); but it works well for the “one-dimensional oven” model of an electrical resistor that we’ll study in Section 9.2. Historically, it was Planck’s attempt to fix this problem that led to quantum mechanics. Quantum mechanics treats the oscillators’ dependence on frequency as no longer continuous, in which case they no longer

obey this particular requirement for the equipartition theorem to hold.² And because that theorem doesn't hold, $\varepsilon(f)$ need not equal kT . And yet, experimentally, $\varepsilon(f)$ does seem to equal kT for a resistor, but not for a hot oven. Why is that so?

Rather than follow Planck's argument, we will use the modern quantum mechanical picture to calculate $\varepsilon(f)$ without using the equipartition theorem. We are modelling the oven walls as a set of *quantum* oscillators held at a temperature T by their interaction with an environment. The mean energy $\varepsilon(f)$ of a single oscillator of frequency f for a given polarisation results from analysing the relevant Boltzmann distribution. Recall, from (5.74), that energy level n of a quantised oscillator has energy $(n + 1/2)hf$, giving it an energy of nhf over and above the ground level's zero-point energy of $1/2 hf$ (where $n = 0$), which is ever present and cannot be taken away from the oscillator. This zero-point energy is thus not treated as internal energy, and so does not enter our analysis.³

A single oscillator of frequency f then has a mean energy of (excluding the zero-point energy)

$$\varepsilon(f) = \sum_{n=0}^{\infty} E_n p_n, \quad \text{where } E_n \equiv nhf. \quad (9.4)$$

If we can justify ignoring the pressure/volume contribution to the Boltzmann exponential (5.5) for *all* energy levels of the oscillators, the Boltzmann probability becomes $p_n \propto e^{-\beta E_n}$. [Remember that a quantum state has a fixed number of particles, so the chemical potential term in (5.5) is not required here.] It's not at all clear, in fact, that we can ignore the pressure/volume term. That term quantifies the effect that “the system” (in this case, a single oscillator) would have on “the bath” if the system were to inflate in size as it jumped from one energy level to another, hence performing mechanical work on the bath. The pressure/volume term is traditionally ignored in discussions of the mean energy of the oscillators; but, given that we are summing to infinity in (9.4), we can only hope that nothing like the old problem with the hydrogen-atom partition function in Section 5.7 appears here. We'll see shortly that (9.4) does not diverge when the pressure/volume term is omitted.

With the standard shorthand $\beta \equiv 1/(kT)$, equation (9.4) becomes

$$\varepsilon(f) = \frac{\sum_n nhf e^{-\beta nhf}}{\sum_n e^{-\beta nhf}} = \frac{hf \sum_n n e^{n\alpha}}{\sum_n e^{n\alpha}}, \quad \text{with } \alpha \equiv -\beta hf. \quad (9.5)$$

Note that α in (9.5) is negative, and so $e^\alpha < 1$. Hence, the probability normalising factor in (9.5) is a well-behaved geometric series:

² The requirements for the equipartition theorem were listed in Section 3.5.2.

³ We have seen this reasoning before, in Sections 7.1 and 7.2.

$$\sum_{n=0}^{\infty} e^{n\alpha} = \frac{1}{1 - e^{\alpha}}. \quad (9.6)$$

Equation (9.6) can be used to compute the sum $\sum n e^{n\alpha}$ in (9.5), via the partial differentiation that we have used several times already in Sections 7.1, 7.2, and 7.6. That is, write

$$\sum_{n=0}^{\infty} n e^{n\alpha} = \frac{d}{d\alpha} \sum_{n=0}^{\infty} e^{n\alpha} = \frac{d}{d\alpha} \frac{1}{1 - e^{\alpha}} = \frac{e^{\alpha}}{(1 - e^{\alpha})^2}. \quad (9.7)$$

Equation (9.5) now becomes

$$\varepsilon(f) = \frac{hf e^{\alpha}}{(1 - e^{\alpha})^2} (1 - e^{\alpha}) = \frac{hf e^{\alpha}}{1 - e^{\alpha}} = \frac{hf}{e^{-\alpha} - 1}, \quad (9.8)$$

or, finally,

$$\boxed{\varepsilon(f) = \frac{hf}{\exp \frac{hf}{kT} - 1}}. \quad (9.9)$$

Is this result reasonable? Study its behaviour in the regimes of low temperature ($kT \ll hf$) and high temperature ($kT \gg hf$):

- **Low temperature:** It's clear that $\varepsilon(f) \rightarrow 0$ here, and so the mean energy $\varepsilon(f)$ of a single oscillator vanishes, as might be expected. (Don't forget that we have omitted the quantum oscillator's ground-state energy $1/2 hf$; this means the oscillator's energy really tends toward $1/2 hf$ as $T \rightarrow 0$.)
- **High temperature:** In this regime, we can write

$$\exp \frac{hf}{kT} \simeq 1 + \frac{hf}{kT}. \quad (9.10)$$

Substituting this into (9.9) produces $\varepsilon(f) \rightarrow kT$. But this quantum value agrees with the classical value given by the equipartition theorem for a single oscillator: an oscillator has 2 quadratic energy terms (kinetic and potential), and each such term contributes a mean internal energy of $1/2 kT$.

Photons in the Oven

Recall that quantum mechanics views a harmonic oscillator as occupying one of an infinite number of energy levels, with the n^{th} level above the $n = 0$ ground level having energy $n hf$ above that ground level, as in (5.74). In contrast, the Einstein and Debye approaches of quantum statistics in Chapter 7 treat the oscillator as being in some *state*, whose “extractable” energy $n hf$ is due to the state being “occupied” by n bosons called *photons*. Each photon has energy hf and, under further analysis,

turns out to have “spin 1”: this means the z component of its spin can be either of $\pm\hbar$; photon spin cannot have a zero z component.

Suppose that we could look around inside the oven walls, focussing on oscillators with frequency f —say, those emitting yellow light. The mean energy of a single “yellow” oscillator is $\varepsilon(f)$. This is a good time to remind ourselves that a single particle corresponds to *three* oscillators, one for each direction of its vibration; but only the two oscillators whose motion is orthogonal to the direction “into” the oven turn out to be relevant, corresponding to the fact that an oscillating charge doesn’t radiate along its direction of motion. This mean energy $\varepsilon(f)$ is “held” by some mean number $\bar{n}(f, T)$ of photons that each have energy hf , which are associated with the oscillator at temperature T :

$$\bar{n}(f, T) = \frac{\text{mean energy of oscillator}}{\text{energy of a photon}} = \frac{\varepsilon(f)}{hf} \stackrel{(9.9)}{=} \frac{1}{\exp \frac{hf}{kT} - 1}. \quad (9.11)$$

Compare this with (7.106) [since we could just as well write $\bar{n}(f, T)$ as $\bar{n}(E, T)$, because $E = hf$]. Because (9.11) refers to a number of particles, it harks back to (7.33). To see why, realise that the spectral energy density $\varrho(f)$ is the oven energy per unit frequency f and per unit volume. It follows that the frequency density of the number of photons in the oven is

$$\left[\begin{array}{c} \text{number of photons} \\ \text{per unit frequency} \end{array} \right] = \frac{V\varrho(f)}{hf} \stackrel{(9.3)}{=} \frac{\varepsilon(f)g(f)}{hf} \stackrel{(9.11)}{=} \bar{n}(f, T)g(f). \quad (9.12)$$

This is the frequency version of (7.33).

Also, compare (9.11) with the comment after (7.106) to conclude that, like phonons, photons have chemical potential $\mu = 0$. This is consistent with our knowledge of the First and Second Laws of Thermodynamics. Write the First Law $dE = TdS - PdV + \mu dN$, and then note that the total energy E of the interior of the resistor, along with its N photons, is constant. Its volume V is also constant. Because everything is in equilibrium, the Second Law says that the total entropy S is constant. The First Law then reduces to $0 = \mu dN$. But photons are continually being created and destroyed, and so $dN \neq 0$. It must be, then, that $\mu = 0$.

Now that we have an expression for $\varepsilon(f)$ in (9.9), our next task is to calculate $g(f)$. Before attempting the full calculation of $g(f)$ in three dimensions, we will gain much insight by calculating it for a *one*-dimensional oven. Such an oven does indeed exist: it is an electrical resistor. The bane of every electronics engineer is the noisy fluctuation that appears in the voltage drop produced across a resistor when current flows through it. This resistor can

be treated as a one-dimensional oven for the purpose of determining how this electrical noise arises.

9.2 The One-Dimensional Oven: an Electrical Resistor

Every electronics engineer is familiar with the idea of “thermal noise” in electrical resistors. This is electronic noise, usually unwanted, and caused by fluctuations in the voltage across the resistor that go on to affect other elements in the circuit. In this section, we’ll use (9.3) to derive the standard expression for this voltage fluctuation.

A resistor that carries an alternating current can be modelled as a one-dimensional oven filled with electromagnetic waves. This is because the applied alternating voltage makes the valence electrons oscillate, and oscillating charges radiate at the frequency of their oscillation. For a resistor of length L , the one-dimensional analogue of (9.3) is

$$\varrho(f) = \frac{\varepsilon(f) g(f)}{L}, \quad (9.13)$$

where now $\varrho(f)$ is the amount of electromagnetic energy in the resistor per unit frequency f , and per unit resistor *length* (not volume).

Unlike sound waves—which are associated with matter vibrations along their direction of motion—electromagnetic waves are associated only with varying electric and magnetic fields *transverse* to their direction of motion. Also, oscillating charges don’t radiate in their direction of oscillation: they radiate in all other directions, but mostly transverse to their direction of motion. The consequence is that only oscillations transverse to the length of the resistor generate appreciable amounts of electromagnetic waves that can move along the resistor. Hence, just as with a hollow oven, two (not three) “wall oscillators” are associated with each of the valence electrons that reside within the entire length of the resistor: one oscillator for each of the two orthogonal directions of oscillation that are themselves both orthogonal to the length of the resistor. Each of these oscillators has mean energy $\varepsilon(f)$.

The resistor turns out to use the high-temperature limit of $\varepsilon(f)$ in (9.9). The frequencies of alternating currents in electronic circuits are generally in the sub-gigahertz range, in which case $hf \approx 10^{-25}$ J. Contrast this with the value of kT at room temperature: $kT \approx 1/40$ eV $\approx 10^{-21}$ J. It follows that $hf \ll kT$, and hence we are in the high-temperature regime here, where the equipartition value of $\varepsilon(f) \simeq kT$ resulting from (9.10) applies. But, even at much higher frequencies, this equipartition value is still quite accurate. For example, at room temperature and with a frequency of $f = 100$ GHz, the ratio $hf/(kT)$ is only about 0.02.

We'll see later that a hot oven does *not* use the high-temperature limit of (9.9). This could sound paradoxical: it might suggest that resistors are hotter than ovens. What is really happening is that the vibrations of the resistor's electrons are *electrically* induced, whereas the jumps between excited atomic energy levels that produce radiation in, say, a ceramic oven are *thermally* induced. In the absence of any current, the temperature in a resistor would need to be very high to excite the electrons as much as the alternating electric field does. When the current is switched off, the resistor really does revert to being a *cold* "oven", because only its lattice atoms are now being excited, and only to room temperature.

The electromagnetic waves that are coupled to the resistor's oscillators have a density $g(f)$ of wave states. We must define and calculate $g(f)$.

9.2.1 Calculating the Density of Wave States, $g(f)$

Recall that we are calculating $\varrho(f)$ in (9.13). We wrote down $\varepsilon(f)$ in (9.9), and need only use its low-frequency/high-temperature approximation, $\varepsilon(f) \simeq kT$, for the resistor. We must now define and calculate $g(f)$, the radiation's density of wave states.

In Sections 2.4 and 2.5, we found an ideal gas's density of energy states $g(E)$ by defining a state to be a cell in phase space. Equation (2.86) introduced the density of states $g(E) = \Omega'_{\text{tot}}(E)$, which referred to the total number of these states $\Omega_{\text{tot}}(E)$ in the energy range 0 to E . We found $\Omega_{\text{tot}}(E)$ by noting that this total number of cells equals the total phase-space volume available to energies 0 to E , divided by the volume of one cell. Planck's constant h defined a natural size for each cell. In Section 2.6, we extended this calculation of $g(E)$ to massless particles. For these, we also switched to the language of frequency via $E = hf$ using (2.99) (because dependence on frequency is more conventional than dependence on energy in the following discussions), and we calculated the corresponding density⁴ $g(f)$.

We will use a similar approach here to recalculate the density of states $g(f)$ for the resistor's internal radiation, by defining and counting the total number $\Omega_{\text{tot}}(f)$ of such states in the frequency range 0 to f , and then applying the definition of a density, $g(f) = \Omega'_{\text{tot}}(f)$. We will reproduce the result (2.100), but will perform the calculation using the concept of "wave number", which is highly useful for understanding waves in a broader context. There is some latitude in the way in which a "wave state" might be defined; but the following definition has led to predictions in statistical mechanics that stand up to experiment extremely well.

⁴ Recall the discussion in Section 2.6: there and here, the symbol g is being "overloaded" with two meanings. As long as we don't write a specific value such as " $g(4)$ ", there is no problem with giving two separate density functions the same name " g ".

If the fact that we are now referring to $g(f)$ instead of $g(E)$ is confusing, recall that $g(f)$ is *defined* such that $g(f) df = g(E) dE$:

$$\begin{aligned} d\Omega_{\text{tot}} &= g(E) dE = \left[\begin{array}{c} \text{number of wave states in} \\ \text{energy interval } E \text{ to } E + dE \end{array} \right] \\ &= g(f) df = \left[\begin{array}{c} \text{number of wave states in} \\ \text{frequency interval } f \text{ to } f + df \end{array} \right]. \end{aligned} \quad (9.14)$$

It follows that $g(E) = \Omega'_{\text{tot}}(E)$ and $g(f) = \Omega'_{\text{tot}}(f)$.

Because we are positing that a continuous range of frequencies exists inside the resistor, it's apparent that—just as with the phase-space cell approach to discretising the continuous position and momentum in Section 2.4—we should discretise frequencies by grouping them into sets, where each set is called a “wave state”. We require any such binning procedure to generalise to waves in three dimensions, to enable a later calculation of $g(f)$ for waves inside an oven. But, although the frequency of a photon relates simply to energy via “ $E = hf$ ”, it turns out that neither frequency nor wavelength are always the most natural parameters with which to describe a wave. We'll shortly use another quantity that is more natural for defining and constructing a wave state.

The Wave Number and Wave Vector

To see why neither frequency nor wavelength are necessarily the best parameters with which to describe light waves, consider what happens in three dimensions. There, we wish to distinguish between two plane waves of the same frequency that travel in different directions. A *vector* parameter describing the plane wave will encapsulate the directional nature of the “light ray” associated with that plane wave. Let $\mathbf{n}$ be the unit vector pointing in the wave's direction of travel, shown in Figure 9.1. Could we perhaps define a “frequency vector” $\mathbf{f} \equiv f\mathbf{n}$, or a “wavelength vector” $\boldsymbol{\lambda} \equiv \lambda\mathbf{n}$, and use one of these to characterise the wave? The answer is no: a quantity with magnitude and direction is not guaranteed to be a vector, and it's easy to show that $\mathbf{f}$ and $\boldsymbol{\lambda}$ are not reasonable quantities to define: they are not vectors. For, if they *are* to be vectors, their components should have the appropriate meaning: each component should quantify the projection of the wave onto the relevant axis. But the components of $\mathbf{f}$ and $\boldsymbol{\lambda}$ fail to do this.

This can be seen in Figure 9.1, in which a plane wave's direction of travel lies in the xy plane and makes an angle θ with the x axis. Use the idea that the x component of any vector $\mathbf{v}$ that lies in the xy plane is $v_x = v \cos \theta$,

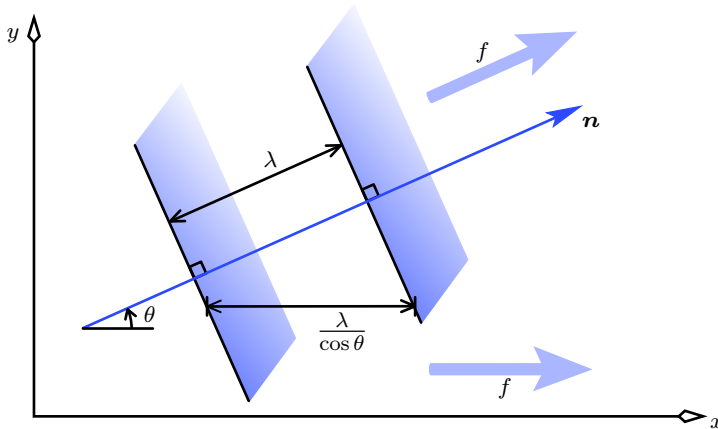


Fig. 9.1 Two wave fronts, usually called “phase fronts” (surfaces of constant phase), of a three-dimensional plane wave that moves along the unit normal $\mathbf{n}$ of those planes. Refer to the text to see how this construction shows why we cannot construct a meaningful wavelength vector as $\boldsymbol{\lambda} \equiv \lambda \mathbf{n}$

where v is the vector’s length. If $\boldsymbol{\lambda}$ were a vector, we would rightfully expect its x component “ $\lambda_x = \lambda \cos \theta$ ” to be “the wavelength along the x axis”, meaning the wavelength of the wave fronts’ intersections with the x axis. This is the distance between the intersection of one crest with the x axis and the intersection of the next crest with the x axis. But the figure shows that this projected wavelength is not $\lambda \cos \theta$, but rather $\lambda / \cos \theta$. So, $\boldsymbol{\lambda}$ doesn’t behave as a vector should.

Likewise, if $\mathbf{f}$ were a vector, we would rightfully expect its x component “ $f_x = f \cos \theta$ ” to be the frequency of the wave fronts’ intersections with the x axis; but that frequency is, in fact, just f ! Thus, neither $\mathbf{f}$ nor $\boldsymbol{\lambda}$ behaves as a vector should, and hence they play no role in physics. The lesson here is that an arbitrary object with length and a direction in space is not guaranteed to be a vector; its components must also have an appropriate physical meaning.

We know that $\cos \theta$ is vital to discussing any possible vector’s x component. It doesn’t appear in the “projected frequency” f above. And it appears in the denominator of the “projected wavelength” $\lambda / \cos \theta$ instead of where we require it to be: the numerator. This suggests that the *reciprocal* of the wavelength might be convertible to a vector. Enter the *wave vector* $\mathbf{k} \equiv k \mathbf{n}$, where $k \equiv 2\pi / \lambda$ is the *wave number*.⁵ The 2π is included for mathematical convenience, but the point is that the factor of $\cos \theta$ now appears in the right place in k_x :

$$k_x \equiv k \cos \theta = \frac{2\pi \cos \theta}{\lambda} = \frac{2\pi}{\lambda / \cos \theta}$$

⁵ Although “wavelength” is written universally without a space, I have stopped short of writing “wavevector”, “wavenumber”, “wavefrequency”, and “wavefunction”.

$$\begin{aligned}
&= 2\pi/(\text{wavelength of wave's projection onto } x \text{ axis}) \\
&\equiv \text{wave number of wave's projection onto } x \text{ axis.}
\end{aligned} \tag{9.15}$$

This is precisely what we require k_x to be. So, $\mathbf{k}$ behaves in the manner of a true vector, unlike $\mathbf{f}$ and $\boldsymbol{\lambda}$. This is why $\mathbf{k}$ is indispensable for characterising waves, which makes it so pervasive in wave theory. In quantum mechanics, $\mathbf{k}$ appears from the outset in de Broglie's celebrated postulate $\mathbf{p} = \hbar\mathbf{k}$, which relates a particle's momentum $\mathbf{p}$ to its wave vector $\mathbf{k}$. We will use $\mathbf{k}$ to define a wave state in what follows.

Using the Wave Number to Define a Wave State

In the single spatial dimension of our resistor, the wave vector $\mathbf{k}$ becomes a real number k whose sign determines the direction in which the wave travels, and whose modulus is $|k| = 2\pi/\lambda = 2\pi f/c$, where c is the speed of the waves in the resistor (which is generally a little slower than the speed of light in a vacuum).

The central importance of the wave number to the theory of waves almost obliges us to formulate a wave state using the language of wave number. Because each wave in the resistor has a particular wave number k , that wave can be allocated a single point in “ k space”. First, we divide this k space into equal-sized cells; then, we define the number of wave states $\Omega_{\text{tot}}(f)$ in the frequency range 0 to f to be

$$\begin{aligned}
\Omega_{\text{tot}}(f) &\equiv \left[\begin{array}{c} \text{number of possible} \\ \text{polarisations} \end{array} \right] \times \left[\begin{array}{c} \text{number of cells} \\ \text{in } k \text{ space} \end{array} \right] \\
&= 2 \times \text{number of cells in } k \text{ space} \\
&= 2 \times \frac{\text{extent of } k \text{ space}}{\text{cell width}},
\end{aligned} \tag{9.16}$$

where the extent of k space corresponds to the frequency range 0 to f .

We will group the waves into cells by comparing the wavelength of each wave with the length of the resistor, because these are the only length scales in the scenario. The idea is shown in Figure 9.2.

1. Locate $k = 0$ in (the one-dimensional) k space, and construct “cell 1” next to it along the k number line as follows. This cell contains all right-moving waves whose wavelengths fit between 0 and 1 times into L . Their wavelengths thus range from infinity down to L , and so their corresponding k values range from “ $2\pi/\infty = 0$ ” to $2\pi/L$. The width of cell 1 is then $2\pi/L$.
2. Next, construct “cell 2” next to cell 1. Cell 2 contains all right-moving waves whose wavelengths fit between 1 and 2 times into L . Their wavelengths thus range from L down to $L/2$, making their corresponding k values range

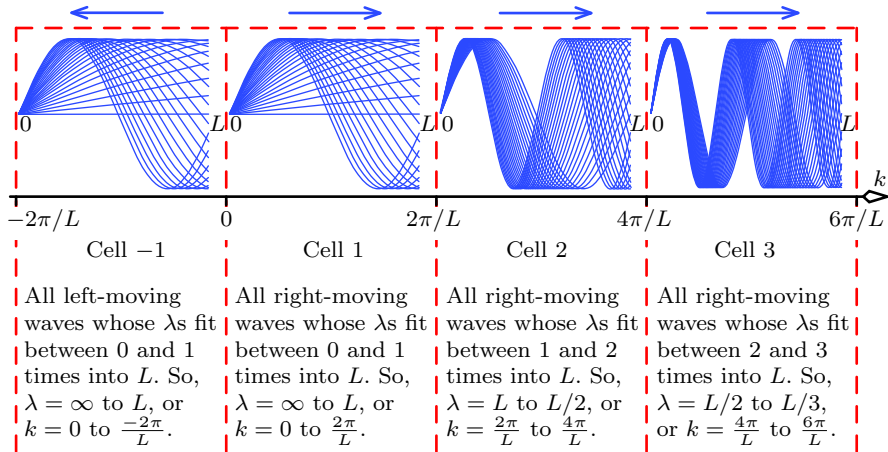


Fig. 9.2 A selection of representative waves that occupy each cell in k space. Each cell holds a *continuum* of wavelengths. The arrow above each cell shows the direction of motion of the waves in x space

from $2\pi/L$ to $2 \times 2\pi/L$. The width of cell 2 is then $2\pi/L$.

3. Similarly, construct “cell 3” next to cell 2. Cell 3 contains all right-moving waves whose wavelengths fit between 2 and 3 times into L . Their wavelengths thus range from $L/2$ down to $L/3$, making their corresponding k values range from $2 \times 2\pi/L$ to $3 \times 2\pi/L$. The width of cell 3 is then $2\pi/L$.
4. In general, construct “cell n ” containing all right-moving waves whose wavelengths fit between $n - 1$ and n times into L . Their wavelengths thus range from $L/(n - 1)$ down to L/n , making their corresponding k values range from $2\pi(n - 1)/L$ to $2\pi n/L$. The width of cell n is then $2\pi/L$.
5. We must also consider left-moving waves: these are grouped in the same way as right-moving waves, but have negative wave number. So, construct “cell -1 ” next to cell 1 as follows. This cell contains all *left*-moving waves whose wavelengths fit between 0 and 1 times into L . Their wavelengths range from infinity down to L , making their corresponding k values (now negative) range from “ $-2\pi/\infty = 0$ ” to $-2\pi/L$. The width of cell -1 is then $2\pi/L$.
6. Cell $-n$ is constructed similarly to cell n , but from left-moving waves.

It’s clear that all of the cells have the same width:

$$\text{width of each cell} = 2\pi/L. \quad (9.17)$$

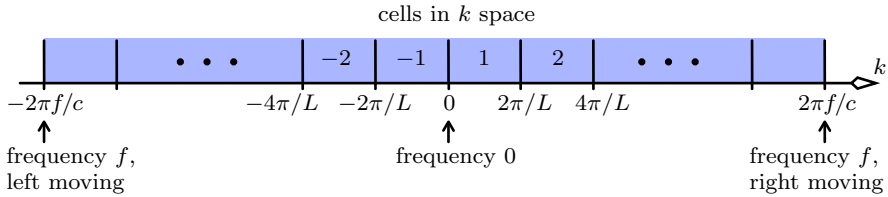


Fig. 9.3 k space in one dimension, for frequencies 0 to f . The space runs from k values of $-2\pi f/c$ to $2\pi f/c$, and each cell has width $2\pi/L$

Next, referring to (9.16), we require the total extent, or width, of k space. This is set by the largest possible k , and is shown in Figure 9.3. Given that $|k| = 2\pi/\lambda = 2\pi f/c$, the frequency range 0 to f maps to $|k|$ values of 0 to $2\pi f/c$. But waves can travel in either direction, and so the allowed values of k are $-2\pi f/c$ to $2\pi f/c$. The extent of the k space filled by waves of frequencies 0 to f is then

$$\text{extent of } k \text{ space} = 4\pi f/c. \quad (9.18)$$

Equation (9.16) now gives the number of wave states as

$$\Omega_{\text{tot}}(f) = 2 \times \frac{\text{extent of } k \text{ space}}{\text{cell width}} = 2 \times \frac{4\pi f/c}{2\pi/L} = 4fL/c. \quad (9.19)$$

The density of states is then

$$g(f) = \Omega'_{\text{tot}}(f) = 4L/c. \quad (9.20)$$

This is precisely what we found in (2.100), where we followed a somewhat different argument.

Here is an alternative mathematical approach to calculating $g(f)$ that you will sometimes encounter. (It differs only in the maths, not the physics.) Instead of calculating the number of states $\Omega_{\text{tot}}(f)$ in the frequency range 0 to f and then differentiating that to find $g(f)$, aim more directly for $g(f)$. Do this by calculating $d\Omega_{\text{tot}}(f) = g(f) df$, the combined number of states in the two intervals of wave number relevant to the frequency f : the first interval is k to $k + dk$, and the second is its partner at the opposite end of k space, $-(k + dk)$ to $-k$:

$$\begin{aligned} g(f) df &= \left[\begin{array}{c} \text{number of possible} \\ \text{polarisations} \end{array} \right] \times \left[\begin{array}{c} \text{number of cells in both} \\ \text{infinitesimal portions of } k \text{ space} \end{array} \right] \\ &= 2 \times (2 \times \text{number of cells in } k \text{ to } k + dk) \\ &= 2 \times \frac{2 dk}{\text{cell width}}. \end{aligned} \quad (9.21)$$

The relationship $k = 2\pi f/c$ leads to $dk = 2\pi df/c$, and so

$$g(f) df = 2 \times \frac{4\pi df/c}{2\pi/L} = 4 df L/c. \quad (9.22)$$

It follows that $g(f) = 4L/c$ yet again.

A Standing-Wave Picture in One Dimension

What if we were to treat the resistor as holding standing waves only, as some texts do? Standing waves relate to whole numbers of *half* wavelengths fitting into L . In that case, only positive values of k would be allowed. This is because a standing wave is a sum of equal-wavelength waves travelling in opposite directions, and thus it needs only a positive wave number to describe it. The previous analysis would now change in two ways. First, the extent of k space would be half the value in (9.18), and so would be $2\pi f/c$. Second, the width of a cell would now be governed by whole numbers of *half* wavelengths fitting into L , and a continuum of waves would no longer be considered:

- Cell 1 would hold only waves with half-wavelength L . Thus, $\lambda = 2L$, and the cell's waves would have $k = 2\pi/\lambda = 2\pi/(2L) = \pi/L$.
- Cell 2 would hold only waves with half-wavelength $L/2$. Thus, $\lambda = 2L/2$, and the cell's waves would have $k = 2\pi/\lambda = 2\pi/L$.
- Cell 3 would hold only waves with half-wavelength $L/3$. Thus, $\lambda = 2L/3$, and the cell's waves would have $k = 2\pi/\lambda = 3\pi/L$.
- Similarly, cell n would hold only waves with half-wavelength L/n . Thus, $\lambda = 2L/n$, and the cell's waves would have $k = 2\pi/\lambda = n\pi/L$.

The cells would all have the same width of π/L : this is half the value that holds for the continuum of waves in (9.17). Equation (9.16) would yield

$$\Omega_{\text{tot}}(f) = 2 \times \frac{\text{extent of } k \text{ space}}{\text{cell width}} = 2 \times \frac{2\pi f/c}{\pi/L} = 4fL/c. \quad (9.23)$$

But this is exactly what we found in (9.19). Then, $g(f) = \Omega'_{\text{tot}}(f) = 4L/c$ again, as in (9.20). This viewpoint gives the same value for $g(f)$ as we found using the “continuum of frequencies” picture. Textbooks often use standing waves in these discussions, because they lead to the same density of states $g(f)$ as the continuum-of-frequencies picture, without requiring any discussion of grouping a continuum of frequencies into cells. Nonetheless, a continuum of frequencies is really no more difficult to analyse.

Let's remind ourselves that we set out to calculate the resistor's spectral energy density $\varrho(f)$ from (9.13). For this, we needed two quantities:

- $\varepsilon(f)$ in (9.9), whose low-frequency limit $\varepsilon(f) \simeq kT$ will be sufficient for most applications of resistors in circuits, and
- $g(f)$, which we have now found is $4L/c$ from multiple approaches: (2.100), (9.20), and the calculations producing (9.22) and (9.23).

The spectral energy density is then

$$\varrho(f) = \frac{\varepsilon(f) g(f)}{L} = \frac{kT \times 4L/c}{L} = 4kT/c. \quad (9.24)$$

This is the amount of electromagnetic wave energy in the resistor per unit frequency per unit resistor length. The quantity of great interest to engineers is the total energy inside the resistor over a frequency range f to $f + df$ for sub-gigahertz frequencies:

$$\begin{aligned} \left[\begin{array}{c} \text{total energy in} \\ \text{resistor in } df \end{array} \right] &= \underbrace{\left[\begin{array}{c} \text{energy in resistor} \\ \text{per unit frequency} \\ \text{per unit length} \end{array} \right] \times \left[\begin{array}{c} \text{frequency} \\ \text{extent } df \end{array} \right] \times \left[\begin{array}{c} \text{resistor} \\ \text{length } L \end{array} \right]}_{= \varrho(f) = 4kT/c} \\ &= 4kTL df/c. \end{aligned} \quad (9.25)$$

This noise in a resistor manifests as a usually unwanted voltage and current. Next, we discuss briefly the consequences of this noise in the field of modern communications.

9.2.2 Excursus: Thermal Noise in a Resistor, and Some Communications Theory

Why might an engineer wish to know the energy (9.25) present in some small frequency range inside a resistor? The rate of flow of this energy out of the resistor is the noise power that he is probably trying to avoid when constructing the circuit, or whose value must be corrected for in a measurement that has used an off-the-shelf piece of electronics. A common situation where we require accurate knowledge of this noise power occurs with the electrical lines that carry the signals used for modern communication.

In particular, consider using electronics to encode the strings of symbols that comprise modern digital communication. Different schemes exist that convert such a string to and from the electrical signal that actually travels along the wires. For the sake of argument, we will suppose that a signal is a sequence of zeroes and ones—for which the term “binary digits” is routinely shortened to “bits”. The presence of a wave pulse of power through the circuit for a set time interval denotes a “1”, and the absence denotes a “0”. The power

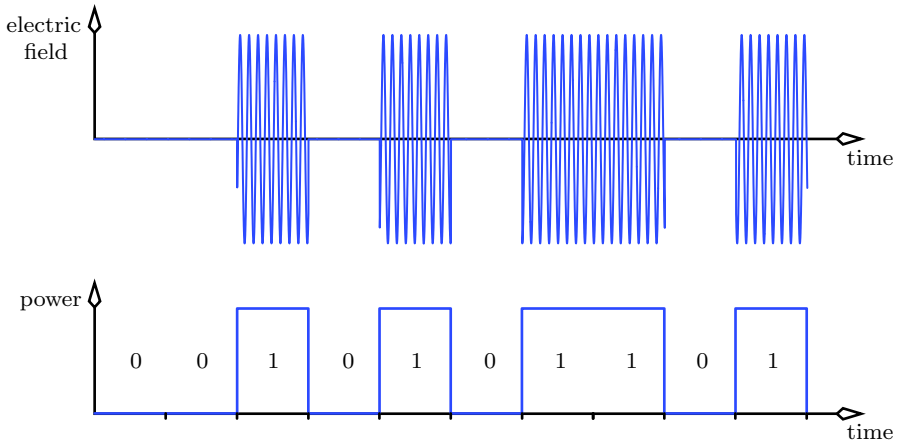


Fig. 9.4 A “square-wave” signal that encodes the sequence 0 0 1 0 1 0 1 1 0 1. The presence of a pulse of oscillating electromagnetic field for a pre-set time interval signals a “1”. **Top:** The oscillating value of the electric field versus time. **Bottom:** The modulation of the field, which determines the power in the circuit. This envelope is the usual way in which a series of pulses is represented

is carried by oscillations of an electromagnetic field in the circuit. Figure 9.4 shows the signal “0 0 1 0 1 0 1 1 0 1”.

This signal has been generated by rapidly switching on and off a carrier wave, which typically has a *far* higher frequency than the occurrence of zeroes and ones in Figure 9.4—the carrier frequency in the figure has been reduced purely to make the oscillations visible. This carrier wave has a single frequency, but when modulated into a series of pulses, the resulting signal is no longer a pure sinusoid. Fourier theory tells us that the signal can be written as the sum of many sinusoids, and these sinusoids tend to be grouped into a range of frequencies known as the signal’s *bandwidth* B .

Let’s investigate the total energy of the electromagnetic waves present in this bandwidth. Equation (9.25) is easily integrated to give the total energy in the frequency band f to $f + B$:

$$\begin{aligned} \text{total energy in resistor of bandwidth } B &= \int_f^{f+B} 4kTL \, df/c \\ &= 4kTLB/c. \end{aligned} \quad (9.26)$$

Envisage this energy as moving along the resistor at the wave speed c , so it emerges in a time of roughly L/c to manifest as noise power. The amount of this noise produced by the resistor is then

$$\text{noise power out} = \frac{\text{energy out}}{\text{time taken}} = \frac{4kTLB/c}{L/c} = 4kTB. \quad (9.27)$$

This is something of an average value, of course, and it's based on the idea of electromagnetic fluctuations occurring inside the resistor. Because the noise in complex circuits arises from many sources interacting with each other in various ways, in practice, the factor of 4 in (9.27) is replaced by the circuit's *noise factor*, F . This result, that the noise power is $FkTB$, has been given the name *Nyquist's theorem for thermal noise in circuits*. It is widely used by electronics engineers.

This thermal noise (also known as *Johnson noise*) manifests as a fluctuating voltage V across the resistance R and a fluctuating current I through it, where $V = IR$. The power dissipated in R can be written as either V^2/R or I^2R . Nyquist's theorem then writes the mean noise power as

$$\langle V^2/R \rangle = \langle I^2R \rangle = FkTB. \quad (9.28)$$

The mean-square voltage and current arising from the noise are then

$$V_{\text{rms}}^2 \equiv \langle V^2 \rangle = FkTBR, \quad I_{\text{rms}}^2 \equiv \langle I^2 \rangle = FkTB/R. \quad (9.29)$$

Central to an engineer's study of how to build a good communications line is the idea of reducing sources of noise-inducing voltage—and, of course, economising on *required* voltage. On the one hand, (9.29) says that to minimise voltage noise, we should minimise the bandwidth B . But we might very well want to design a system that can generate electromagnetic waves over a *large* bandwidth. This is because a high bandwidth means we have a large range of frequencies at our disposal from which to construct signals. Fourier analysis says that this large range can be used to create a signal wave form with a large amount of *structure*.

Why would we require such structure? Modern technology demands an ever-increasing rate of information flow; but we cannot increase the flow rate by increasing the signal speed, because electromagnetic signals travel at a set speed through a given transmission line. Instead, we can only send more bits per second if we shorten the duration of each pulse. But a series of very short pulses sent in, say, ten seconds equates to more structure than a series of only a few pulses sent in that same time interval. That higher structure in the signal with many pulses requires more sinusoids to build it; that is, the many-pulses signal requires more bandwidth than the few-pulses signal.

Figure 9.5 shows an example of each type of signal. At its top left, we see the “0 0 1 0 1 0 1 1 0 1” signal from Figure 9.4 again, without the carrier. This signal lasts for 10 seconds, and the first 20 Hz of its frequency spectrum above the carrier is shown at the top right in Figure 9.5.⁶ We have plotted the frequencies relative to the carrier, whose value is immaterial, and thus

⁶ I produced this spectrum by running a “discrete Fourier transform” on a set of samples taken at one-millisecond intervals from the signal. The nature of the discrete Fourier transform means it returns a discrete set of frequencies: consider these to be a discrete approximation of the actual spectrum, which is continuous.

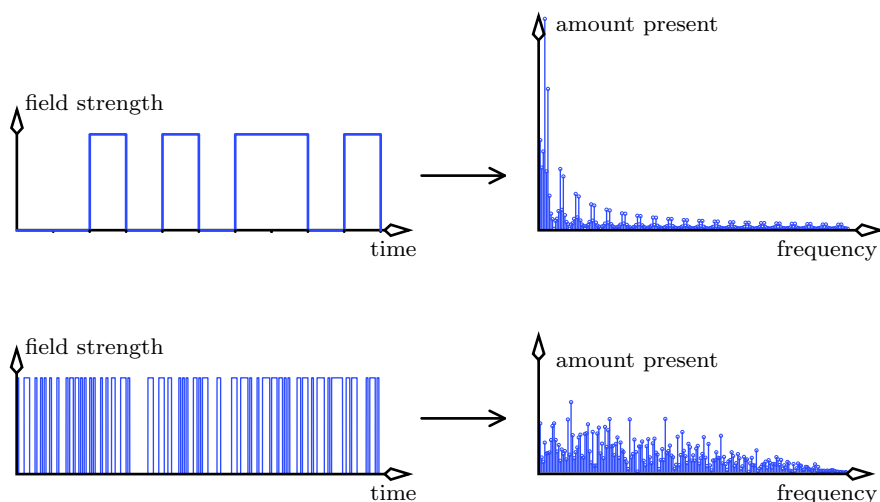


Fig. 9.5 **Top left:** A short sequence of zeroes and ones. **Top right:** The relatively simple “discrete Fourier spectrum” of this sequence (with carrier removed), showing only frequencies above the carrier frequency, because the spectrum is symmetrical about the carrier. **Bottom left:** A signal that delivers many more zeroes and ones in the same time as the one above it, and which thus has more structure than the top signal. **Bottom right:** The spectrum of this signal draws from a broader band of frequencies than the top signal

is being placed at zero frequency; also, because the spectrum is symmetrical about the carrier, we have shown only frequencies greater than the carrier. In fact, many higher frequencies are needed to build the signal, but their weights drop off as their frequencies increase. Thus, the first 20 Hz suffices to show the spectrum’s shape.

At the bottom left in Figure 9.5 is shown a new signal that squeezes in twenty times as many zeroes and ones as the top signal in the same transmission time, equating to twenty times the data rate. The first 20 Hz of its spectrum is shown at bottom right in the figure. It’s apparent that the weights of the higher frequencies have increased compared with those of the simpler signal, and so the more complicated signal exhibits a broader band of frequencies than the simpler signal. This need for a “broader band” of frequencies to send data at a higher rate has given rise to the term *broadband*, which is used often in modern data communications.

So, although reducing the value of the bandwidth B in (9.29) reduces the overall circuit noise, increasing B allows for higher data rates—but increases the noise. This noise can introduce errors into the transmission. If we are content with very low data rates, then we can always encode the data in such a way that the error rate is arbitrarily low. There will then be a maximum rate C at which we can send the data. The governing principle here is the *Shannon–Hartley theorem*. This states that the maximum transmission

rate C that a data-transmitting channel can have, below which we can always arrange for an arbitrarily low error rate, is a function of the available bandwidth B , and the ratio of the generated signal power to the resulting noise power, S/N :

$$\underline{C} = B \log_2(1 + S/N). \quad (9.30)$$

bits per unit time

That is, C is the maximum data rate that we can ever guarantee to be error free. (It is also known as the *channel capacity*.) By the phrase to “arrange for an arbitrarily low error rate” above, we mean the following. When signals are sent down a line, errors at the receiving end can always be introduced by noise en route and in the receiver itself. Sophisticated error-correction algorithms can find and correct some of these errors; but the higher the percentage of errors we wish to remove, the more sophisticated the algorithm must be. The Shannon–Hartley theorem puts an upper bound on the amount of information that we can ever send, even if we have an all-powerful error-correction algorithm that finds and corrects 100% of the errors.

For example, suppose we wish to send a signal with a signal-to-noise ratio of $S/N = 10$ (where S is signal power and N is noise power), and we have a bandwidth of $B = 1$ MHz at our disposal. Then, (9.30) gives the channel capacity (the maximum throughput for which we can ever hope to arrange an arbitrarily low error rate), as

$$C = 1 \text{ MHz} \times \log_2 11 = 3.46 \text{ megabits per second}. \quad (9.31)$$

The Shannon–Hartley expression (9.30) shows that to achieve a high channel capacity, we require a high bandwidth B and a high signal-to-noise S/N . But Nyquist’s result (9.28) says that increasing the bandwidth will add more noise to the signal, and thus *lower* the signal-to-noise S/N . That puts the brakes on the gain we were hoping to make by increasing B . Because (9.28) says that $N \propto B$, suppose we write (9.30) as

$$C \propto B \ln(1 + \alpha/B), \quad (9.32)$$

for some constant α . Then, in the limit of large B , (9.32) yields

$$C \rightarrow B \times \alpha/B = \alpha. \quad (9.33)$$

That is, the channel capacity tends toward a constant as we increase the bandwidth. Clearly, *some* bandwidth is necessary, but there is no point in having an excessively large amount of it as far as channel capacity is concerned. Nonetheless, a high bandwidth is worth a great deal in the world of data transmission.

Bandwidth in Radar and GPS

In the early days of radar in World War II, the range to an enemy aircraft could be found by “pinging” the aircraft with a short pulse of radio waves, then measuring the time from the ping’s emission from the radar set to its reception back at the radar set. An accurate measurement of the range required a well-defined, loud ping that could be heard above the receiver’s electronic noise. Such an impulsive type of ping requires a great many frequencies to build it (that is, it requires greater bandwidth), compared to a lazier ping constructed of just a few sinusoids, which has no well-defined start or end, and hence does not give an accurate range measurement.

In our era, digital signal processing has revolutionised radar. An aircraft need no longer be pinged by a single, loud, short-and-sharp analogue radio pulse. Instead, a sequence of very quiet pulses that effectively makes up a sequence of numbers is bounced from the aircraft. This sequence can have very low power, and thus be covert. After sending the sequence, the radar receiver gathers everything it can “hear” in a possible return, and searches for the emitted signal in that return. It does this by *convolving* the emitted sequence of numbers with the digitised return. We encountered the convolution of continuous functions in (4.95); that equation has a discrete counterpart that is used in digital signal processing. This discrete convolution is, in fact, identical to long multiplication (without the “carrying” procedure), or multiplying polynomials—although, in practice, it is often implemented more efficiently by using the discrete Fourier transform. This convolution can pinpoint the start and end of the emitted sequence of numbers very precisely in a very noisy return, and can thus determine the aircraft’s range precisely, even when a low-powered radar signal is being sent out. An emitted signal that has a very high structure (which requires more bandwidth than a simpler signal) is naturally easier to locate in a noisy return, in a similar way to humans being able to pick a quiet sentence out from a loud hubbub, when they are familiar with the speaker’s choice of words and tone of voice.

The same idea applies to signals sent by satellites of, for example, the Global Positioning System (GPS). The signals broadcast to Earth from these satellites are *exceptionally* weak—in fact, they lie below the electronic noise level in the receivers. But those signals have a high bandwidth that endows them with an extraordinarily complex structure. Because the receiver knows what signals are being sent by the satellites, it can search for those complicated signals in what it receives from the sky. Once it locates the signal precisely, it can calculate the range to the satellite accurately. Then, given knowledge of where several satellites are currently located, it is able to triangulate its own position on Earth. It does not even need a high-accuracy clock to establish signal flight times.

The satellites themselves contain atomic clocks, and the unknown time at the receiver's location is treated as a variable, like its position, that is determined as part of the triangulation procedure.

Thus, in radar and signal processing, high bandwidth is everything. The equations of signal processing lean heavily on bandwidth. A way to understand this is to realise that it is not bandwidth itself that is somehow key to those equations; instead, high bandwidth means an ability to construct *more complicated* signals—and more complicated signals are easier to search for in a noisy data set than are simple signals.

9.3 The Three-Dimensional Oven

In Section 9.1, we set out to calculate an oven's spectral energy density $\varrho(f)$. This is the amount of electromagnetic energy present in the oven, per unit frequency f and per unit oven volume. We first tackled the simpler one-dimensional problem of finding $\varrho(f)$ for a resistor in (9.13). Treating the resistor as a one-dimensional oven meant we could explore the calculation of the density of wave states $g(f)$ in one spatial dimension, which is simpler than the full three-dimensional treatment.

At this point, we have all the tools and have done most of the work needed to calculate $\varrho(f)$ for a three-dimensional oven of volume V . Return to (9.3), which related $\varrho(f)$ to $\varepsilon(f)$ and $g(f)$. For $\varepsilon(f)$, we have the full expression (9.9). Only $g(f)$ remains to be found for the oven. We'll calculate it by mimicking the last section's wave number analysis for the resistor.

Just as for the resistor, we calculate the oven radiation's density of wave states $g(f)$ by defining and counting the total number of those states in the frequency range 0 to f , and then applying the definition of a density, $g(f) = \Omega'_{\text{tot}}(f)$. We will reproduce the result found back in (2.107) that resulted from a different approach to counting states.

Any particular wave in the oven has a wave vector $\mathbf{k} = (k_x, k_y, k_z)$. We will define and count the number of states $\Omega_{\text{tot}}(f)$ in the frequency range 0 to f via a similar procedure to that of the resistor, now by considering the volume of the region of the three-dimensional $\mathbf{k}$ space that the allowed wave vectors occupy. Analogous to (9.16), write

$$\begin{aligned}\Omega_{\text{tot}} &= 2 \times \text{number of cells in } \mathbf{k} \text{ space} \\ &= 2 \times \frac{\text{volume of } \mathbf{k} \text{ space}}{\text{cell volume}}.\end{aligned}\tag{9.34}$$

We group these waves into cells by treating each spatial dimension independently. The disconnectedness of the oven's shape with its spectrum enables

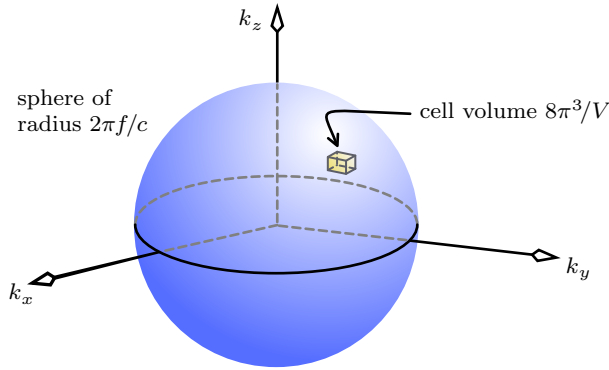


Fig. 9.6 The three-dimensional version of Figure 9.3. Just as the one-dimensional space encompasses all cells out to a “radius” of $2\pi f/c$, so too, the three-dimensional space encompasses all cells out to a radius of $2\pi f/c$, creating a sphere of cells. And just as the cells in the one-dimensional space have length $2\pi/L$, the cells in the three-dimensional space have volume $2\pi/L_x \times 2\pi/L_y \times 2\pi/L_z = 8\pi^3/V$

us to specify any shape that will make the analysis easy. Suppose, then, that the oven is a rectangular box with side lengths L_x, L_y, L_z , and so has volume $V = L_x L_y L_z$. In the x direction, follow the same analysis as in the one-dimensional case above, but replace the one-dimensional case’s λ with λ_x , the wavelength of the wave fronts’ projections onto the x axis. Exactly the same argument as in one dimension gives a constant cell width along this axis of $2\pi/L_x$ [recall (9.17)]; similarly, the cell widths along the y and z axes are $2\pi/L_y$ and $2\pi/L_z$. Each cell’s volume is the product of these:

$$\text{volume of each cell} = \frac{2\pi}{L_x} \frac{2\pi}{L_y} \frac{2\pi}{L_z} = \frac{8\pi^3}{V}. \quad (9.35)$$

Equation (9.34) also demands the volume of $\mathbf{k}$ space. The frequency range 0 to f defines this volume. Remember that $|\mathbf{k}| = 2\pi/\lambda = 2\pi f/c$, and so the frequency range 0 to f maps to $|\mathbf{k}|$ values of 0 to $2\pi f/c$. Given that the waves can now travel in all directions, the relevant portion of $\mathbf{k}$ space is a *sphere* of radius $2\pi f/c$, shown in Figure 9.6.

The number of wave states is, from (9.34),

$$\begin{aligned} \Omega_{\text{tot}} &= 2 \times \frac{\text{volume of a sphere of radius } 2\pi f/c}{\text{cell volume } 8\pi^3/V} \\ &= 2 \times \frac{\frac{4}{3}\pi(2\pi f/c)^3}{8\pi^3/V} = \frac{8\pi f^3 V}{3c^3}. \end{aligned} \quad (9.36)$$

Finally, the density of wave states is

$$g(f) = \Omega'_{\text{tot}}(f) = 8\pi f^2 V/c^3. \quad (9.37)$$

We found the same result in (2.107) by a different argument.

To gain a feel for the number of states Ω_{tot} , suppose the oven is full of yellow light of wavelength 600 nm. Equation (9.36) says, for a 1 m³ oven,

$$\Omega_{\text{tot}} = \frac{8\pi f^3 V}{3c^3} = \frac{8\pi V}{3\lambda^3} = \frac{8\pi \times 1}{3 \times (600 \times 10^{-9})^3} \simeq 4 \times 10^{19}. \quad (9.38)$$

Simply dividing the sphere's volume by the cell volume implies that we are also including "part cells" on the sphere's surface whose cubic shape has been somewhat "shaved off" by the sphere's curved surface. The incompleteness of these cells might be seen as problematic; but they are in a minority and can be ignored, because the volume of the sphere is so much larger than the total volume of these cells, which are confined to its surface. How much larger? Equation (9.36) shows that the ratio is

$$\frac{\text{volume of sphere}}{\text{volume of cell}} = \frac{\Omega_{\text{tot}}}{2} \approx 10^{19}. \quad (9.39)$$

Using a Shell Instead of a Sphere in $\mathbf{k}$ Space

Just after (9.20), we gave a slightly different mathematical approach to calculating $g(f)$ in one dimension. For completeness, we will do the same here for three dimensions. Recall that calculating the density of states $g(f)$ by differentiating $\Omega_{\text{tot}}(f)$ is equivalent to calculating the infinitesimal $d\Omega_{\text{tot}}(f)$. But, whereas $\Omega_{\text{tot}}(f)$ requires the calculation of the volume of a sphere in $\mathbf{k}$ space, $d\Omega_{\text{tot}}(f)$ implies a calculation of that sphere's surface area.

The analogy to an everyday sphere of radius R is that its volume is the integral of its surface area times an infinitesimal surface thickness:

$$\text{volume} = \int_0^R \text{surface area} \times dr = \int_0^R 4\pi r^2 dr = \frac{4}{3}\pi R^3. \quad (9.40)$$

Differentiating the volume with respect to R returns the surface area.

Our calculations of $g(f)$ above were effectively counting the number of cells in a thin spherical shell of radius k centred at the origin of $\mathbf{k}$ space. This could be done by expressing the shell's volume as its surface area $4\pi k^2$ times its thickness dk . In effect, what we did was calculate the area of the spherical shell by differentiating the sphere's volume $\frac{4}{3}\pi k^3$ with respect to k to arrive at $4\pi k^2$. In the same way, $g(f)$ can be calculated by considering this shell of $\mathbf{k}$ space. Do this in analogy to (9.21), by writing

$$g(f) df = \left[\begin{array}{c} \text{number of possible} \\ \text{polarisations} \end{array} \right] \times \left[\begin{array}{c} \text{number of cells in spherical shell} \\ \text{of radius } k \text{ and thickness } dk \end{array} \right]$$

$$\begin{aligned}
&= 2 \times \frac{\text{volume of spherical shell of radius } k \text{ and thickness } dk}{\text{cell volume}} \\
&= \frac{2 \times 4\pi k^2 dk}{8\pi^3/V} = \frac{2 \times 4\pi \left(\frac{2\pi f}{c}\right)^2 \frac{2\pi df}{c}}{8\pi^3/V} = \frac{8\pi f^2 V df}{c^3}. \quad (9.41)
\end{aligned}$$

This result is (9.37) again, found without any mention of Ω_{tot} .

A Standing-Wave Picture in Three Dimensions

Just as in the one-dimensional case, some texts treat the oven as having conducting walls, and thus supporting standing waves only. The resulting analysis mimics the discussion of one-dimensional standing waves around (9.23), but now for three-dimensional standing waves. Just as in the one-dimensional analysis, the cell width along each axis in $\mathbf{k}$ space turns out to have half its continuum-of-frequencies value. Thus, the cell volume is reduced by a factor of 8.

Also, with each standing wave being a sum of two waves moving in opposite directions, the wave number components need only be all positive. This reduces $\mathbf{k}$ space to the “ k_x, k_y, k_z all positive” octant in Figure 9.6, and hence also reduces the relevant $\mathbf{k}$ space volume by a factor of 8. A glance at (9.34) then shows that the value of Ω_{tot} is unchanged from its value in the continuum-of-frequencies picture, because the two new factors of 1/8 cancel. The value of $g(f)$ is then also unchanged from its value in (9.37). Hence, just as in the one-dimensional case, the standing-wave picture gives the same result as the continuum-of-frequencies picture. But it’s the continuum of frequencies that is more physically meaningful in a general oven, and that’s why we have focussed more on that picture in this chapter.

Most discussions of the radiation in an oven treat the oven as being full of standing waves. This runs counter to the idea that an ideal oven is presumably made of perfectly black material, which is hardly a mirror surface and whose walls would not reflect waves at all, and so would not produce an environment in which only standing waves existed. But we see here that this assumption of standing waves is not actually necessary. In analogy, we saw, in Section 3.8.4, that when a (micro)state is defined to be a cell in phase space, any multiple of Planck’s constant can be used to define that cell’s extent, because what results is a unique expression for *increases* in entropy. The real difficulty in defining microstates was due to our insistence on *counting* those microstates, which required some notion of a *discrete* microstate.

The same ideas of counting apply to waves in the oven. Defining and then counting their states by constructing discrete cells in wave-number space can be problematic: do we choose a continuum of waves or standing

waves, and why do both choices give the same result? But no matter which choice we make, we have some kind of counting procedure that gives a seemingly unique expression for $g(f)$, and this expression turns out to produce the experimentally verified expression for the spectral energy density $\varrho(f)$. It is interesting and nontrivial as to why this should be. Ideas of counting wave states are intimately related to the quantum mechanical idea of representing the waves by a “gas” of photons. Even so, a basic difference between photons and the particles of the ideal gas in Chapter 2 is that the number of photons in an oven is continuously changing, whereas the number of ideal gas particles in a container remains constant.

9.4 The End Product: Planck’s Law

Having obtained $\varepsilon(f)$ in (9.9) and $g(f)$ for the three-dimensional oven in (9.37), we can now place them into (9.3) to write *Planck’s law* for the oven’s spectral energy density $\varrho(f)$, its electromagnetic energy per unit frequency f , per unit oven volume:

$$\varrho(f) = \frac{1}{V} \times \frac{hf}{\exp \frac{hf}{kT} - 1} \times \frac{8\pi f^2 V}{c^3}, \quad (9.42)$$

so that

$$\boxed{\varrho(f) = \frac{8\pi h f^3 / c^3}{\exp \frac{hf}{kT} - 1}}. \quad (9.43)$$

Planck’s law is shown in Figure 9.7. His energy density $\varrho(f)$ reduces to zero as the frequency increases, and this agrees closely with experiment. Planck formulated this law in 1900, by introducing the revolutionary postulate that the oscillators in the oven walls could only radiate their energy in quantised amounts hf proportional to their frequency of oscillation f : the constant of proportionality h that allowed the law to match experimental data became known as Planck’s constant.

Planck’s result replaced a slightly earlier expression credited to Rayleigh, and later re-derived by Jeans. Rayleigh and Jeans used the classical expression $\varepsilon(f) = kT$ that was based on equipartition with two quadratic energy terms, referred to just after (9.9). This gave them an energy density of

$$\varrho_{\text{RJ}}(f) \stackrel{(9.3)}{=} \frac{kT \times 8\pi f^2 V / c^3}{V} = 8\pi f^2 kT / c^3. \quad (9.44)$$

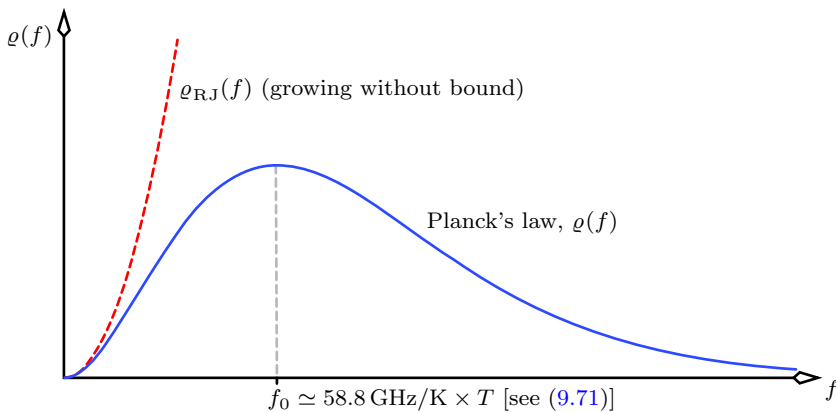


Fig. 9.7 The solid curve is Planck's law (9.43). Compare this with the earlier result of Rayleigh and Jeans (9.44), the dashed curve, which grows without bound

Because $\rho_{\text{RJ}} \propto f^2$ (as shown in Figure 9.7), this energy density predicts the existence of ever-larger amounts of radiation at high frequencies—meaning the oven is expected to contain an infinite amount of energy. When Rayleigh first derived his expression, this clearly wrong prediction at high frequencies became known as the *ultraviolet catastrophe*. In contrast, Planck's successful prediction of the energy density rested on the new idea of energy quantisation, and so marked the beginning of quantum theory. As expected, Planck's expression reduces to that of Rayleigh and Jeans in the low-frequency limit:

$$\rho(f \approx 0) \stackrel{(9.43)}{=} \frac{8\pi h f^3 / c^3}{1 + \frac{hf}{kT} - 1} = 8\pi f^2 kT / c^3 = \rho_{\text{RJ}}(f). \quad (9.45)$$

The real difference between the Rayleigh–Jeans expression and that of Planck comes down to their choices of $\varepsilon(f)$. Planck used the full version (9.9): recall, from (9.5), that because this version originates in the Boltzmann distribution, it incorporates factors of $e^{-\beta hf}$ whose effect is to suppress high-frequency contributions to $\varepsilon(f)$. In the language of quantum mechanics, oscillators with a high frequency have a large spacing hf between energy levels [see (5.74)]; hence, excited levels that are able to release a quantum of energy hf into the oven are very poorly populated.

In contrast, the Rayleigh–Jeans expression $\varepsilon(f) = kT$ is independent of frequency: each wave oscillation in the oven was tied, via the equipartition theorem, to an energy of $1/2 kT$, irrespective of its frequency. Thus, because an arbitrary number of high-frequency waves could exist inside their oven, Rayleigh and Jeans were effectively allowing an infinite amount of energy to exist in the oven.

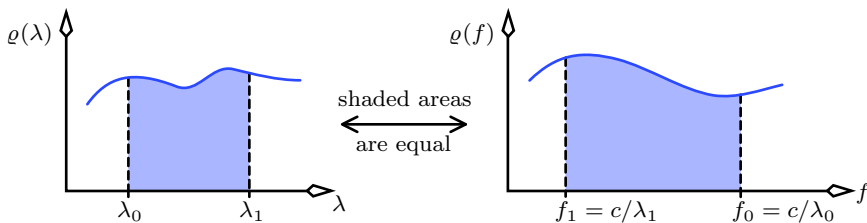


Fig. 9.8 Left: The area under the left-hand curve from λ_0 to λ_1 is defined to be the energy present in an oven of unit volume, in the wavelength range λ_0 to λ_1 . **Right:** That energy equals the area under the right-hand curve between f_0 to f_1 , where the frequencies are those that correspond to the given wavelengths. To be generic, the above curves are deliberately drawn so as not to follow the Planck distribution

9.4.1 Planck's Law Expressed Using Wavelength

We have expressed the oven's spectral energy density as a function $\varrho(f)$ of frequency: by definition, the area of a vertical strip under the graph of $\varrho(f)$ versus f is the energy present in an oven of unit volume, in the frequency range of the vertical strip. This energy density can also be written as a function of wavelength λ :

$$\varrho(\lambda) \equiv \left[\begin{array}{l} \text{amount of electromagnetic energy per unit} \\ \text{wavelength } \lambda, \text{ per unit oven volume} \end{array} \right]. \quad (9.46)$$

The area of a vertical strip under the graph of $\varrho(\lambda)$ versus λ is the energy present in an oven of unit volume, in the *wavelength* range of that vertical strip.⁷ How are the functions $\varrho(f)$ and $\varrho(\lambda)$ related?

Figure 9.8 shows the situation. We are given a band of wavelengths λ_0 to λ_1 within which we require the energy in the oven. First, create a band of corresponding frequencies f_0 to f_1 , where each frequency corresponds to a wavelength such that their product is the speed of light c . By definition,

$$\left[\begin{array}{l} \text{energy in unit-volume} \\ \text{oven in } \lambda_0 \text{ to } \lambda_1 \end{array} \right] = \int_{\lambda_0}^{\lambda_1} \varrho(\lambda) d\lambda = \int_{f_1}^{f_0} \varrho(f) df. \quad (9.47)$$

To isolate the required $\varrho(\lambda)$ here, we must eliminate the integral sign. So, consider an infinitesimal band of wavelengths λ to $\lambda + d\lambda$, shown in Figure 9.9.

⁷ As mentioned in Sections 2.6 and 9.2.1, when writing functions with similar meanings but different arguments, it's usual to economise the notation by using the same symbol (in this case ϱ) for what are really two different functions, and rely on its argument to indicate which function is meant. Of course, simply writing " $\varrho(5)$ " would be ambiguous here, since we have no way of knowing whether a frequency of 5 or a wavelength of 5 is meant. But we will always indicate the nature of the argument explicitly. If you are ever in a situation of needing to indicate explicitly which function is meant when a numerical argument is used, you can always write $\varrho_f(5)$ and $\varrho_\lambda(5)$.

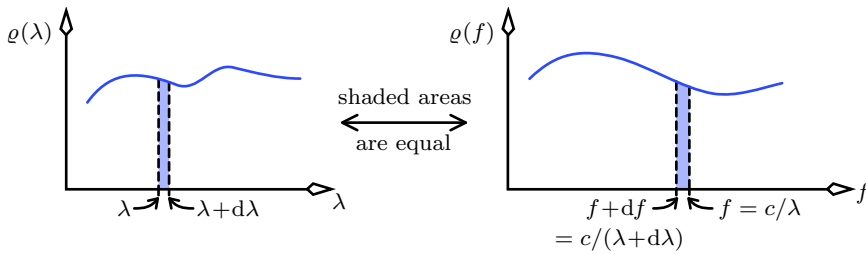


Fig. 9.9 The infinitesimal version of the areas in Figure 9.8

The area of the strip under the $\rho(\lambda)$ curve is $\rho(\lambda)$ times the width of the strip, $d\lambda$. This product, $\rho(\lambda) d\lambda$, equals the area of the strip under the $\rho(f)$ curve, which is $\rho(f)$ times the width of the strip on that curve—and this width is $-df$, not df . Hence,

$$\begin{aligned} \text{wavelength plot's strip area} &= \rho(\lambda) d\lambda \\ &\equiv \text{corresponding frequency plot's strip area} = \rho(f) \times -df. \end{aligned} \quad (9.48)$$

But wouldn't we have supposed that the frequency strip width was equal to $\rho(f) df$, as we wrote under the last integral sign in (9.47)?

Remember that df means “final f minus initial f ”. It is thus defined by a *process* that has initial and final states. In (9.47), df refers to the everyday process of integrating from f_1 to f_0 , where $f_1 < f_0$. As we form each product $\rho(f) df$ in the integration as the values of f move from the smaller number f_1 to the larger number f_0 , the increase df always equals the right-hand number minus the left-hand number, and so df is positive. Contrast this with the role of df in (9.48). There, df was written *after* $d\lambda$ was defined, and so df is tied to $d\lambda$: frequency f corresponds to wavelength λ , and frequency $f + df$ corresponds to wavelength $\lambda + d\lambda$ via $f\lambda = (f + df)(\lambda + d\lambda) \equiv c$. In that case, these infinitesimals are related in the usual way as a derivative: because $f = c/\lambda$, it follows that

$$df/d\lambda = -c/\lambda^2 < 0. \quad (9.49)$$

So, df and $d\lambda$ have opposite signs. Because the width of the strip under the $\rho(\lambda)$ curve in Figure 9.9 is $d\lambda > 0$, it follows that $df < 0$, and so the width of the corresponding strip under the $\rho(f)$ curve in that figure must be $-df > 0$.

In Figure 9.9, as we “grow” a strip's thickness from left to right on the wavelength plot by increasing the wavelength from λ to $\lambda + d\lambda$, a corresponding strip of equal area will “grow” on the frequency plot from *right* to *left*, with values of frequency that *decrease* from f to $f + df$ (*not* $f - df$: remember that when the initial value of some quantity is x , the final value is *defined* to be $x + dx$: see Section 1.6). It then follows from (9.48) that

$$\varrho(\lambda) = \varrho(f) \times \frac{-df}{d\lambda} = \frac{8\pi hf^3/c^3}{\exp \frac{hf}{kT} - 1} \times \frac{c}{\lambda^2}, \quad (9.50)$$

or

$$\boxed{\varrho(\lambda) = \frac{8\pi hc/\lambda^5}{\exp \frac{hc}{\lambda kT} - 1}.} \quad (9.51)$$

This treatment of $\varrho(\lambda)$ ensures that, just as for $\varrho(f)$, the total energy in a unit-volume oven is given by the usual prescription of finding an area under a curve, where the integral is taken from the “left (smaller) value to the right (larger) value”, irrespective of whether that value is of frequency or wavelength.

A final comment: this attaching of a minus sign to df , that expresses the fact that the areas in Figures 9.8 and 9.9 are positive, is actually a simple one-dimensional example of why the *absolute value* of the “jacobian determinant” is required when we change variables in a multi-dimensional integration.

9.5 Total Energy of Radiation in the Oven

The spatial energy density, or total energy U of radiation in a unit-volume oven, is the spectral energy density integrated over all frequencies or wavelengths. We’ll choose frequency:

$$U = \int_0^\infty \varrho(f) df \stackrel{(9.43)}{=} \int_0^\infty \frac{8\pi hf^3/c^3}{\exp \frac{hf}{kT} - 1} df. \quad (9.52)$$

A change of variables $x \equiv hf/(kT)$ converts this to

$$U = \frac{8\pi k^4 T^4}{c^3 h^3} \underbrace{\int_0^\infty \frac{x^3 dx}{e^x - 1}}_{= \pi^4/15} = \frac{8\pi^5 k^4}{15c^3 h^3} T^4 \equiv \frac{4\sigma}{c} T^4. \quad (9.53)$$

Here, $\sigma = 5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$ is the *Stefan–Boltzmann constant*, defined separately from the factor of $4/c$ to simplify (9.61) ahead. Finally, the total electromagnetic energy in an oven of volume V and temperature T is UV :

$$\boxed{\text{total electromagnetic energy inside oven} = V \frac{4\sigma}{c} T^4.} \quad (9.54)$$

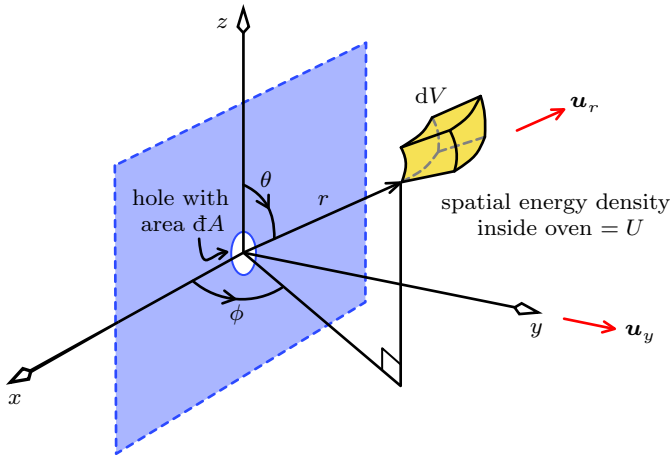


Fig. 9.10 A view from inside the oven of a small hole of area dA made in its side, through which energy can escape. We calculate the energy sent to the hole from a volume element dV

9.6 Letting the Radiation Escape the Oven

We began this chapter by enquiring into the amount of radiation emitted by a hot object. We made the analogy of a black body with a hot oven, and have now calculated the amount of energy inside this oven. For the next step in determining how much energy the hot object emits, we make a small hole in the oven and determine the energy's rate of escape through the hole.

Refer to Figure 9.10 for the scenario. It shows the hole of area dA as seen from inside the oven, along with an element of spatial volume dV that contains energy. Some of this energy is destined to pass through the hole during a time interval Δt . We will calculate how much energy exits the hole in time Δt , and then divide that by $dA \Delta t$ to arrive at the rate of energy emitted per unit hole area—that is, the power emitted per unit hole area. Denoting the area as “ dA ” emphasises that we will not be integrating an infinitude of infinitesimal holes; there is no total area “ A ” to be considered. But we *will* integrate over the volume elements dV .

Place the origin of a cartesian coordinate system at the hole, and let the wall containing the hole lie in the xz plane, with the y axis pointing into the oven, as in the figure. We will use spherical polar coordinates r, θ, ϕ to exploit the spherical symmetry of the scenario.

The energy passing through the hole in Δt is the energy of *some* of the photons that are within a distance $c \Delta t$ of the hole: the photons of interest are those moving in the correct direction to encounter the hole. The infinitesimal amount of energy dE exiting the hole from the volume dV at a distance $r \leq c \Delta t$ from the hole is then set by the solid angle subtended by dA as seen

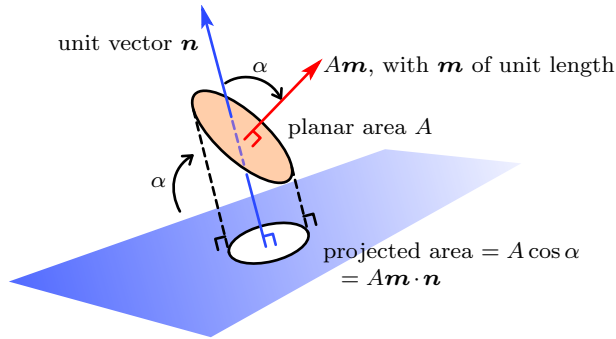


Fig. 9.11 A planar area whose plane forms an angle α to a plane surface can be treated as being formed of strips that each have their length shortened by $\cos \alpha$ when projected onto the surface, and whose projected width is unchanged. It follows that the projected area is $A \cos \alpha$. But $\cos \alpha$ is the dot product of two unit-length normals: $\mathbf{m}$, the normal to the area element, and $\mathbf{n}$, the normal to the surface. The projected area is thus $A\mathbf{m} \cdot \mathbf{n}$. The element's area and direction are conventionally written as a single vector, $A\mathbf{m}$, that can be “dotted” with the unit normal to any given plane to find the area projected onto that plane

from dV :

$$dE = \left[\begin{array}{c} \text{total energy} \\ \text{in } dV \end{array} \right] \times \frac{1}{4\pi} \times \left[\begin{array}{c} \text{solid angle subtended by } dA \\ \text{as seen from } dV \end{array} \right]. \quad (9.55)$$

The solid angle is the area that dA projects onto a sphere of radius r centred at dV , divided by r^2 :

$$dE = U dV \times \frac{1}{4\pi r^2} \times \left[\begin{array}{c} \text{projection of } dA \text{ onto sphere} \\ \text{of radius } r \text{ centred at } dV \end{array} \right]. \quad (9.56)$$

As shown in Figure 9.11, the projection of one area onto another is given by the dot product of the vector representing the area with the unit normal to the surface. For the case of Figure 9.10, the area element is represented by $-dA \mathbf{u}_y$, where $\mathbf{u}_y$ is the unit-length y basis vector. The unit normal to the sphere centred at dV is the negative of the unit-length radial basis vector $\mathbf{u}_r$, or $-\mathbf{u}_r$. The dot product can be evaluated by first expressing $\mathbf{u}_y$ and $\mathbf{u}_r$ in the cartesian coordinates of the figure:

$$\begin{aligned} [\mathbf{u}_y]_{\text{cart}} &= (0, 1, 0), \\ [\mathbf{u}_r]_{\text{cart}} &= (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta). \end{aligned} \quad (9.57)$$

The projection of dA onto the sphere is then

$$\begin{aligned}
\text{projected area} &= -\hat{\mathbf{d}}A \mathbf{u}_y \cdot -\mathbf{u}_r = \hat{\mathbf{d}}A \mathbf{u}_y \cdot \mathbf{u}_r \\
&= \hat{\mathbf{d}}A (0, 1, 0) \cdot (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta) \\
&= \hat{\mathbf{d}}A \sin \theta \sin \phi.
\end{aligned} \tag{9.58}$$

Equation (9.56) becomes, with $dV = r^2 \sin \theta dr d\theta d\phi$,

$$\begin{aligned}
dE &= U r^2 \sin \theta dr d\theta d\phi \times \frac{\hat{\mathbf{d}}A \sin \theta \sin \phi}{4\pi r^2} \\
&= \frac{U \hat{\mathbf{d}}A}{4\pi} \sin^2 \theta \sin \phi dr d\theta d\phi.
\end{aligned} \tag{9.59}$$

The total energy exiting the hole in a time Δt is the sum of the energies dE from all volume elements dV that lie within a distance $c \Delta t$ of the hole:

$$\begin{aligned}
\text{total energy exiting hole of area } \hat{\mathbf{d}}A \text{ in time } \Delta t &= \int dE \\
&= \underbrace{\int_0^{c \Delta t} dr}_{= c \Delta t} \underbrace{\int_0^\pi d\theta \sin^2 \theta}_{= \pi/2} \underbrace{\int_0^\pi d\phi \sin \phi}_{= 2} \frac{U \hat{\mathbf{d}}A}{4\pi} \\
&= c \Delta t \frac{U \hat{\mathbf{d}}A}{4}.
\end{aligned} \tag{9.60}$$

The energy radiated per unit hole area per unit time is the bottom line of (9.60) divided by $\hat{\mathbf{d}}A \Delta t$, and is $Uc/4$. In other words,

$$\text{power emitted per unit hole area} = \frac{Uc}{4} \stackrel{(9.53)}{=} \frac{2\pi^5 k^4}{15c^2 h^3} T^4 = \sigma T^4, \tag{9.61}$$

where σ is the Stefan–Boltzmann constant from (9.53).

Recalling that U is the total energy integrated over all frequencies per unit volume of the oven, the simple result of this section is that the power exiting a unit-area hole is found by multiplying U by $c/4$. But the above hole argument still holds true if we focus on an *infinitesimal* frequency or wavelength interval. It follows that we can replace U with the energy per unit volume *per unit frequency* $[\varrho(f)]$, or per unit wavelength $[\varrho(\lambda)]$, and the same multiplication by $c/4$ still applies:

$$\left[\begin{array}{l} \text{power emitted by oven} \\ \text{per unit hole area,} \\ \text{per unit frequency} \end{array} \right] = \frac{\varrho(f)c}{4} \stackrel{(9.43)}{=} \frac{2\pi h f^3 / c^2}{\exp \frac{hf}{kT} - 1}, \tag{9.62}$$

$$\left[\begin{array}{l} \text{power emitted by oven} \\ \text{per unit hole area,} \\ \text{per unit wavelength} \end{array} \right] = \frac{\varrho(\lambda)c}{4} \stackrel{(9.51)}{=} \frac{2\pi h c^2 / \lambda^5}{\exp \frac{hc}{\lambda kT} - 1}. \tag{9.63}$$

9.7 “Blackbody Radiation”: The Spectrum of a Black Body

We’re now in a position to address the original task of calculating how much radiation is radiated by a black body. We argued at the start of this chapter that when such a body is placed inside an oven, it must radiate what it absorbs. Hence, it must radiate the same (Planck) spectrum that is present in the oven:

$$\left[\begin{array}{l} \text{power radiated by a black body} \\ \text{per unit area of its surface} \\ \text{(whether across the whole} \\ \text{spectrum, or per unit frequency} \\ \text{or wavelength)} \end{array} \right] = \left[\begin{array}{l} \text{power in the same part of the} \\ \text{spectrum that emerges from a} \\ \text{unit-area hole made in the} \\ \text{side of the oven} \end{array} \right]. \quad (9.64)$$

This entails only a slight change to the left-hand sides of (9.62) and (9.63):

$$\left[\begin{array}{l} \text{power radiated by a black body} \\ \text{per unit surface area,} \\ \text{per unit frequency} \end{array} \right] = \frac{\varrho(f)c}{4} \stackrel{(9.43)}{=} \frac{2\pi hf^3/c^2}{\exp \frac{hf}{kT} - 1}, \quad (9.65)$$

$$\left[\begin{array}{l} \text{power radiated by a black body} \\ \text{per unit surface area,} \\ \text{per unit wavelength} \end{array} \right] = \frac{\varrho(\lambda)c}{4} \stackrel{(9.51)}{=} \frac{2\pi hc^2/\lambda^5}{\exp \frac{hc}{\lambda kT} - 1}. \quad (9.66)$$

These are “blackbody spectra” that refer to real black bodies, as opposed to ovens. (“Black-body spectra” is more correct grammatically, but the hyphen is generally omitted.)

Figure 9.12 shows plots of $\varrho(f)c/4$ and $\varrho(\lambda)c/4$ for a range of temperatures. It’s apparent that the value f_0 of frequency for which the frequency plot peaks increases with temperature. The value λ_0 of wavelength for which the wavelength plot peaks decreases with temperature. We can easily find these values by setting $\varrho'(f_0)c/4$ and $\varrho'(\lambda_0)c/4$ each to zero. The most well-known expression here is for wavelength, and so setting $\varrho'(\lambda_0)$ to zero using (9.51) gives

$$\left(\frac{hc}{\lambda_0 kT} - 5 \right) \exp \frac{hc}{\lambda_0 kT} + 5 = 0. \quad (9.67)$$

Setting $x \equiv hc/(\lambda_0 kT)$ transforms (9.67) into the equation $(x - 5)e^x + 5 = 0$. This turns out to have a single root $x \simeq 4.96511$. In other words,

$$\lambda_0 = \frac{hc}{xkT} \simeq \frac{2.89777 \text{ mm K}}{T}. \quad (9.68)$$

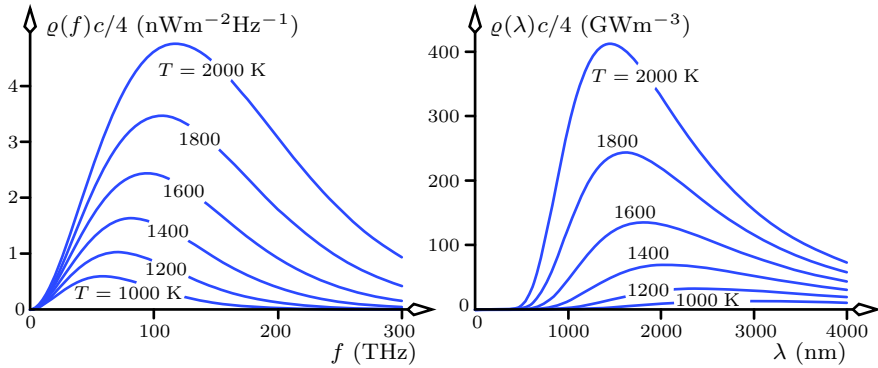


Fig. 9.12 Left: Power radiated from a black body per unit surface area, per unit frequency, $\varrho(f)c/4$. **Right:** Radiated power per unit surface area, per unit wavelength, $\varrho(\lambda)c/4$. The area under each curve is the total power radiated from a unit-area surface

Equation (9.68) is *Wien's law*, and returns what might be called the “most copiously emitted wavelength”, λ_0 . Notice that because λ_0 maximises the power density $\varrho(\lambda)c/4$ radiated by a black body, it also maximises the spectral density $\varrho(\lambda)$ inside an oven. That is, λ_0 is also the most common wavelength present inside an oven at temperature T .

For an example of Wien's law, consider that in the right-hand plot in Figure 9.12, the law says that the peak of the spectrum for $T = 2000$ K occurs at

$$\lambda_0 \simeq \frac{2.898 \text{ mm K}}{2000 \text{ K}} \approx 1450 \text{ nm.} \quad (9.69)$$

This value is evident in the figure.

Our Sun's power output is fitted well by a Planck spectrum whose peak lies at about 500 nm. Such a good fit to a Planck curve suggests that the Sun is well modelled as a black body, and so Wien's law can be applied to calculate its surface temperature:

$$T \simeq \frac{2.898 \text{ mm K}}{500 \text{ nm}} \approx 5800 \text{ K.} \quad (9.70)$$

This temperature is on the cooler side, as stars go. It is even thought to be very similar to that of Earth's core. The temperature of the Sun's core is in the million-kelvin range, as we found in Section 3.15.

The above procedure around (9.67) that found the most copiously emitted wavelength also serves to give the “most copiously emitted frequency” f_0 from (9.43), although the result is not widely written. We must solve $\varrho'(f_0)c/4 = 0$. Differentiating (9.43) gives an expression that can be written in terms of a dimensionless quantity $y \equiv hf_0/(kT)$. This turns out, numerically, to be $y \simeq 2.82144$. Hence, the most copiously emitted frequency is

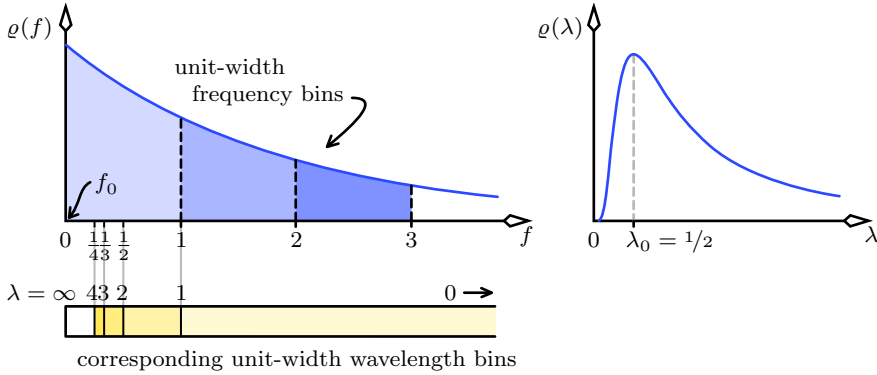


Fig. 9.13 **Left:** The frequency density $\varrho(f) = e^{-f}$. Unit-width frequency bins are delineated by dashed vertical lines. Below the plot are drawn the unit-width wavelength bins. **Right:** The corresponding wavelength density $\varrho(\lambda) = e^{-1/\lambda}/\lambda^2$

$$f_0 = ykT/h \simeq 58.8 \text{ GHz/K} \times T. \quad (9.71)$$

It might at first be thought that $f_0\lambda_0 = c$; but, in fact, it turns out that $f_0\lambda_0 \simeq 0.568c$. Why? The phrase “most copiously emitted wavelength” might suggest that the wavelengths present are like balls of various colours, and we are finding the ball of the most common colour—and similarly for frequency. But this is not really a true picture. Rather, we are taking the values of frequency and wavelength to be continuous, and are calculating two *densities*, of the total energy per unit frequency and per unit wavelength. But frequency and wavelength are not related linearly; recall (9.49). In that case, the equal-width frequency bins that we are essentially comparing to find the “most copiously emitted frequency” do not map one-to-one to the equal-width wavelength bins that we are essentially comparing to find the “most copiously emitted wavelength”. So, the phrase “most copiously emitted” should be taken with a grain of salt.

A simpler example can make the above analysis clearer. Suppose, for simplicity, that f and λ are *dimensionless* quantities, such that

$$f = 1/\lambda, \quad \text{and so} \quad df = -d\lambda/\lambda^2. \quad (9.72)$$

Suppose that we have a density of energy per unit frequency of $\varrho(f) = e^{-f}$. The corresponding wavelength density is $\varrho(\lambda)$. Applying (9.48) yields

$$\varrho(\lambda) = e^{-1/\lambda}/\lambda^2. \quad (9.73)$$

These densities are plotted in Figure 9.13. Note that the total energy present can be calculated by integrating either the frequency density or the wavelength density:

$$\int_0^\infty \varrho(f) df = \int_0^\infty \varrho(\lambda) d\lambda = 1. \quad (9.74)$$

The frequency density peaks at $f_0 = 0$, whereas the wavelength density peaks at $\lambda_0 = 1/2$ [as can be verified by solving $\varrho(\lambda_0) = 0$]. Clearly, $f_0 \neq 1/\lambda_0$ here. Studying the left-hand plot in Figure 9.13, we see that most of the energy resides in the first frequency bin ($f = 0$ to 1), with ever-decreasing amounts in the next bins $f = 1$ to 2, $f = 2$ to 3, and so on. Each of these frequency bins maps to a wavelength bin that contains the same energy, but these wavelength bins don't have unit width, and so are not appropriate for calculating the density in wavelength: this density is defined to be the amount of energy per *unit-width* wavelength bin.

Bins of unit width in wavelength are drawn below the left-hand plot in Figure 9.13. We see that a large amount of energy lies in the rightmost wavelength bin, $\lambda = 0$ to 1. Less energy lies in the next bin, $\lambda = 1$ to 2. And only tiny amounts lie in the remaining bins beginning with $\lambda = 2$ to 3, and so on. It's reasonable, then, that the peak of the wavelength density should lie in the wavelength bin $\lambda = 0$ to 1—and, in fact, we indeed find that $\lambda_0 = 1/2$. Relating different density plots to each other can sometimes stretch our intuition.

9.7.1 The Large-Scale Universe Is a Very Cold Oven

In the early 1960s, physicists Arno Penzias and Robert Wilson discovered a background of microwave radiation coming from all directions in the sky, with a fixed intensity. This radiation follows a Planck distribution to a high accuracy, with a peak wavelength of around 1 mm, corresponding to a temperature [via Wien's law in (9.68)] of about 2.73 kelvins. In line with modern cosmology's theory of the origin of the universe, this radiation seems to be a relic of the Big Bang, and fills the universe. It is called the *cosmic microwave background*.

Modern cosmology is based on solutions to Einstein's equation (6.178) for models of the large-scale universe. The field conventionally describes the universe as expanding.⁸ During this expansion, the light waves of the cosmic microwave background (the “CMB radiation”) get stretched as they traverse immense cosmic distances.

⁸ The amount of data supporting any flavour of cosmological theory is actually rather low, on account of the immense difficulties that astronomers face in making measurements of the extremely faint galaxies and quasars that cosmology is built on. “Anomalous redshifts” of some quasars is an example of a study that questions current cosmological ideas. Aside from questions of the correctness of the comparatively young subject of cosmology, the examination in this section assumes that the universe is expanding along conventional lines.

Many astronomy books say that a galactic *red shift*—the reddening of light from a distant galaxy—is a Doppler shift arising from the galaxy’s inferred recession from us. This idea has long been supplanted by general relativity, which views a *generic* galaxy as not moving through space away from us (although real galaxies can have such an additional motion). Rather, space itself is expanding, stretching the light travelling from the galaxy to us in the process. This same stretching is what gives CMB radiation its characteristic wavelength.

If the spectrum of these waves followed a Planck distribution at some time⁹ $t = 0$, will they follow a Planck distribution at a later time t ? The answer is yes, which can be seen as follows. Suppose that at $t = 0$, the universe had a volume V and was filled with radiation that followed a Planck distribution ϱ_0 at temperature T_0 . Refer to (9.51), where we now indicate the temperature dependence explicitly:¹⁰

$$\text{initial distribution} \equiv \varrho_0(\lambda, T_0) = \frac{8\pi hc/\lambda^5}{\exp \frac{hc}{\lambda k T_0} - 1}. \quad (9.75)$$

Suppose that over a long time, every photon’s wavelength gets stretched by a factor of some “ a ” due to cosmological expansion. The volume of the universe that is filled with the radiation grows to $a^3 V$. This stretching results in a new spectral energy density $\varrho_t(\lambda, T)$ at time t . We must determine whether ϱ_t follows a Planck distribution. Consider the following:

$$\begin{aligned} \frac{a^3 V \varrho_t(\lambda, T) d\lambda}{hc/\lambda} &= \text{number of photons in } \lambda \text{ to } \lambda + d\lambda \text{ at time } t \\ &= \text{number of corresponding “unstretched” photons at } t = 0 \\ &= \text{number of photons in } \lambda/a \text{ to } \lambda/a + d(\lambda/a) \text{ at } t = 0 \\ &= \frac{V \varrho_0(\lambda/a, T_0) d\lambda/a}{hc/(\lambda/a)} = \frac{1}{a^2} \frac{V \varrho_0(\lambda/a, T_0) d\lambda}{hc/\lambda}. \end{aligned} \quad (9.76)$$

Cancelling common factors on the left-hand side and in the last term of this expression yields

$$\varrho_t(\lambda, T) = \frac{1}{a^5} \varrho_0(\lambda/a, T_0) \stackrel{(9.75)}{=} \frac{1}{a^5} \frac{8\pi hca^5/\lambda^5}{\exp \frac{hca}{\lambda k T_0} - 1}$$

⁹ Here, “time” is understood to be a cosmological time that quantifies simultaneity in the large-scale universe. This idea is delicate, because simultaneity is not a uniquely defined idea in general relativity.

¹⁰ We make the standard assumption that the values of the constants h, c, k haven’t changed over time. Whether they might have, or not, is an area of research in cosmology.

$$= \frac{8\pi hc/\lambda^5}{\exp \frac{hc}{\lambda k T_0/a} - 1} = \left[\begin{array}{l} \text{Planck distribution} \\ \text{at temperature } T_0/a \end{array} \right]. \quad (9.77)$$

We see that the current distribution $\varrho_t(\lambda, T)$ is a Planck distribution at a reduced temperature of $T = T_0/a$. This is consistent with Wien's law (9.68), which says that the peak wavelength is inversely proportional to the temperature:

$$[\text{peak wavelength at } t] \propto \frac{1}{T_0/a}, \quad (9.78)$$

which implies that

$$[\text{peak wavelength at } t] = a \times [\text{peak wavelength at } t = 0]. \quad (9.79)$$

This is just as we expected, given that all wavelengths are being stretched by the factor of a as the universe expands.

Modern high-fidelity measurements of the CMB radiation show the presence of a dipole anisotropy on top of a Planck curve. The anisotropy is well-modelled as a Doppler shift arising from Earth's motion within a privileged cosmological frame: this privileged frame is defined as that in which a pure Planck distribution is seen. Earth's motion in this frame amounts to an average speed of about 368 km/s, and is a sum of its daily spin, its yearly orbit around the Sun, the Sun's orbit around the centre of our Milky Way Galaxy (one circuit every 200 million years), and our Galaxy's motion as part of the "Local Group" of galaxies.

When a best-fit of the anisotropy is subtracted from the observed CMB spectrum, what is left over matches a Planck curve extremely closely. Small departures from a perfect Planck curve are thought to signal the presence of fluctuations in the mass/energy distribution of spacetime in its earliest moments after the Big Bang: these fluctuations are thought to have crystallised into the galactic structure that we see around us now. The mechanism for generating these fluctuations is unknown: why, for example, would a Big Bang not generate a universe with complete spherical symmetry? It might be thought that spherical symmetry is a very special class of universe that is less likely to form in a Big Bang than a universe that lacks symmetry. That idea assumes that our universe was a statistical outcome of a kind of "throw of a die". But there is no primal reason to assume that any such cosmic die was thrown. Conversely, it might be thought that the very presence of fluctuations suggests that our universe *did* arise as a throw of a die. Such discussions lie at the speculative edge of cosmology.

9.7.2 Total Power Emitted by a Black Body

Returning to Earth, we can now examine the total power (that is, integrated over all frequencies) radiated by a black body:

$$\begin{aligned}
 \left[\begin{array}{l} \text{total power radiated by a} \\ \text{black body of area } A \end{array} \right] &= A \times \left[\begin{array}{l} \text{total power radiated by a} \\ \text{black body, per unit area} \end{array} \right] \\
 &\stackrel{(9.64)}{=} A \times \left[\begin{array}{l} \text{total power emitted from a} \\ \text{hole in the side of an oven,} \\ \text{per unit hole area} \end{array} \right] \\
 &\stackrel{(9.61)}{=} A\sigma T^4.
 \end{aligned} \tag{9.80}$$

The degree to which real emitters are not perfectly black is quantified by their *emissivity* $e(\lambda, T)$, usually measured experimentally. The emissivity is sometimes approximated as a constant e for the material. In that case,

$$\boxed{\text{total power radiated by an object with temperature } T \simeq Ae\sigma T^4.} \tag{9.81}$$

The *reflectivity* of an object is defined as

$$\text{reflectivity} \equiv 1 - \text{emissivity}. \tag{9.82}$$

Some representative emissivities are shown in Table 9.1.

Table 9.1 Emissivities of various materials

Material	Emissivity e	Material	Emissivity e
water	0.96	marble (polished)	0.90
candle soot	0.95	Pyrex glass	0.90
opaque plastic (any colour)	0.95	wood	0.85
asphalt	0.94	graphite	0.70
concrete	0.94	aluminium paint	0.55
paper (any colour)	0.94	aluminium (polished)	0.09
red brick	0.93	brass (polished)	0.03
earthenware ceramic	0.90	silver (polished)	0.02

The Sun as a Black Body

Our Sun's spectrum is fitted very well by a Planck curve for a black body. Given this fact, let's use a measurement of the power per unit area received from it on Earth to estimate its temperature.

The areal intensity of power received from the Sun at the top of Earth's atmosphere is about $I = 1366 \text{ W/m}^2$. Given that the Sun is about $D = 149.6$ million kilometres away, its total radiant power is

$$\begin{aligned} \text{total radiant power } P &= I \times 4\pi D^2 \\ &= 1366 \times 4\pi \times (149.6 \times 10^6)^2 \text{ W} \\ &\simeq 3.84 \times 10^{26} \text{ W}. \end{aligned} \quad (9.83)$$

Set this total radiant power equal to $A\sigma T^4$ and solve for T . The Sun's radius is $R = 696,000 \text{ km}$, so

$$T^4 = \frac{P}{A\sigma} = \frac{I 4\pi D^2}{4\pi R^2 \sigma}. \quad (9.84)$$

Inserting the appropriate SI units produces

$$T^4 = \frac{1366 \times (149.6 \times 10^6)^2}{(6.96 \times 10^8)^2 \times 5.67 \times 10^{-8}} \simeq 1.1130 \times 10^{15} \text{ K}^4.$$

Hence, $T \simeq 5780 \text{ K}$. This is consistent with the estimate in (9.70).

The Blackbody Spectrum and Wien's Law in Radio Astronomy

In the field of radio astronomy, the blackbody spectrum is often plotted in a way that, confusingly, misrepresents its meaning as a density. As a result, radio astronomers are obliged to use a slightly different form of Wien's law. We derive that different form here.

We've seen two flavours of Planck's law for a black body in (9.65) and (9.66). The frequency version plots $\varrho(f)c/4$ versus f , and the area under the curve gives the power radiated per unit emitter area in a given frequency range. Similarly, the wavelength version plots $\varrho(\lambda)c/4$ versus λ , and the area under the curve gives the power radiated per unit emitter area in a given wavelength range. The emitter area is usually unknown, and so astronomers do not try to plot, say, $\varrho(f)c/4$. Instead, they measure the power received per unit *receiver* area per unit frequency. The traditional unit is the *jansky*:

$$1 \text{ jansky} \equiv 10^{-26} \text{ W m}^{-2} \text{ Hz}^{-1}. \quad (9.85)$$

This is a valid thing to do; if we now plot janskys received versus frequency, the area under the plot is the power received in a given frequency interval.

But a problem of interpretation occurs when researchers choose to plot janskys versus *wavelength*, rather than frequency. Now a single plot mixes

frequency (janskys on the y axis) and wavelength (on the x axis). The resulting set of curves for various temperatures looks much like the wavelength spectra in the right-hand plot in Figure 9.12; so, this strange mixing of frequency and wavelength is certainly not being done to produce a differently shaped spectrum that might be more amenable to analysis. The choice of what to measure, janskys, is governed by considerations of receiver bandwidth; but the plotting of this versus *wavelength* is perhaps determined only by wavelength being arguably more well-known than frequency as a free variable.

Although plotting $\varrho(f)$ versus λ is not mathematically incorrect per se, such a plot *does* misuse the whole idea of a density. It is akin to plotting the linear mass density of a long wire versus, instead of the distance from one end, the *reciprocal* of that distance. When you portray linear mass density by plotting it versus the usual variable of *distance*, you can immediately estimate the mass contained in the wire between two points: it will be the area under the curve. In contrast, if you plot linear mass density versus the reciprocal of distance, the area under the curve has no physical meaning. You will have only a characterless and potentially misleading plot that doesn't exhibit the information that the concept of density was designed to portray.

The standard forms of Wien's law, (9.68) and (9.71), don't apply to a plot of $\varrho(f)$ versus λ . A new version of the law must then be used instead. Create it by differentiating $\varrho(f)$ with respect to λ : start with $\varrho(f)$ in (9.43), set $f = c/\lambda$, differentiate the resulting expression with respect to λ , and find the root $\tilde{\lambda}_0$ of the derivative numerically (where the tilde denotes relevance only to the janskys-versus-wavelength plot). The result is similar to (9.68), but with a different constant:

$$\tilde{\lambda}_0 = \frac{\text{constant}}{T} \simeq \frac{5.100 \text{ mm K}}{T}. \quad (9.86)$$

If you find yourself having to analyse a plot of janskys versus wavelength, you will have to leave your hard-earned mathematical and physical intuition at the front door step.¹¹

¹¹ I grew up reading books on astronomy, and I think it's a fine subject. But I suspect that an element of contrariness is part of the culture of modern *analytic* astronomy, requiring physicists to be on their guard when working in this field. Many astronomers define various quantities in ways that clash with mathematical and physical usage. Apart from the above (mis)use of Wien's law, they omit minus signs, define angles in the opposite direction to established mathematical/physical convention, and—in orbit theory—add certain angles in a mathematically ill-defined way (producing the “longitude of perifocus” and “mean longitude”, which are only used for tabulation, never analysis, and so do not affect predictions). Outdated units such as ergs and parsecs are widespread in the field, despite those units generally being more obscure than conventional ones such as joules and light-years. Some astronomers replace the conventional prefix “mega” with “Mio.” for no apparent reason. And the International Astronomical Union's recent and somewhat arbitrary redefinition of “planet” has only created inconsistency and dissension.

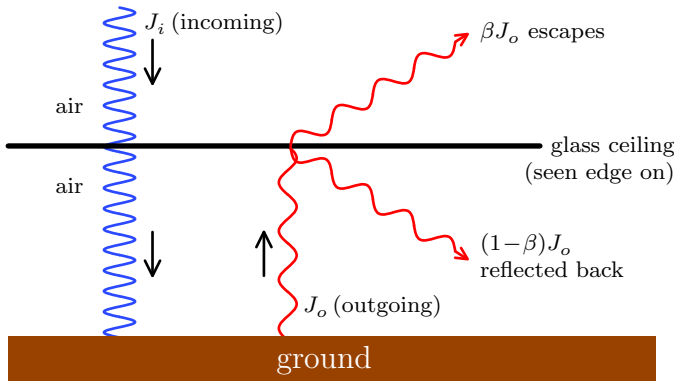


Fig. 9.14 Side view of the glass ceiling and interior of a greenhouse. To the left, incoming sunlight (mainly yellow) has flux density J_i , and passes through the glass with negligible reflection. This light is then absorbed by the ground, and heats the ground. The ground radiates an outgoing flux density J_o at longer wavelengths. These wavelengths are scattered appreciably by the glass. But only a fraction β of this outgoing radiation escapes the glass house; the rest is reflected back to the ground, heating it still further

9.8 The Greenhouse Effect

At ground level, Earth receives a mean flux density of solar energy of 175 Wm^{-2} , averaged over all latitudes and all times of the day and night. Of this, 90% or about 158 Wm^{-2} is absorbed, and the rest reflected. We ask: what would the average temperature on Earth's surface be if it had no atmosphere?

We assume that the absorbed 158 Wm^{-2} heats Earth's surface. This surface radiates a spectrum appropriate to its temperature and an emissivity of $e = 0.9$. Recall, from (9.81), that the total power emitted by a body is $Ae\sigma T^4$. For Earth, this means

$$A \times 0.9 \times \sigma T^4 = A \times 158 \text{ Wm}^{-2}. \quad (9.87)$$

Hence, if Earth had no atmosphere, its surface temperature would be

$$T(\text{no atmosphere}) = \left(\frac{158}{0.9 \times 5.67 \times 10^{-8}} \right)^{1/4} \text{ K} = 236 \text{ K} = -37^\circ\text{C}. \quad (9.88)$$

Lacking an atmosphere would make Earth inhospitable for humans to live on. But just how does our atmosphere moderate Earth's surface temperature?

We can analyse this question by referring to the glass greenhouse shown in Figure 9.14. Solar radiation comes down with an incoming flux density J_i : this can be the above figure of 175 Wm^{-2} . Almost all of the solar spectrum passes through the glass roof without being absorbed and re-scattered, and

contributes to heating the ground. The hot ground radiates an “outgoing” flux density J_o , but at much longer wavelengths, since it is not at the Sun’s temperature: these wavelengths are largely in the infra-red.

When this flux density J_o of mainly infra-red light encounters the glass, only an amount βJ_o passes through, where (of course) $0 < \beta < 1$. The rest, $(1 - \beta)J_o$, is scattered back to the ground, and thus further contributes to heating the ground. In equilibrium, it’s the combination of this “second” heating and the “initial” heating due to J_i that causes the ground to radiate J_o . In equilibrium, the total flow must be zero everywhere. In that case, picture an imaginary plane above the glass and parallel to it. The flux density down through this imaginary plane is J_i , and this must equal the flux density up through it, βJ_o :

$$J_i = \beta J_o. \quad (9.89)$$

Alternatively, place the imaginary plane between the glass and ground, but still parallel to the glass. What comes down through *this* imaginary plane, $J_i + (1 - \beta)J_o$, equals the flux density up through it, J_o . This equality gives us (9.89) again.

Equation (9.89) says that $J_i < J_o$. This doesn’t mean that more energy is being sent out from Earth than is coming in; it simply means that the ground must radiate (J_o) more than what came directly from the Sun (J_i), because it must also re-radiate the portion of light that was reflected back down onto it by the glass ceiling, $(1 - \beta)J_o$.

Now consider J_o with and without the glass ceiling present:

$$\frac{\sigma T^4(\text{glass})}{\sigma T^4(\text{no glass})} = \frac{J_o(\text{glass})}{J_o(\text{no glass})} = \frac{J_i/\beta}{J_i} = \frac{1}{\beta} > 1. \quad (9.90)$$

We see that the effect of the glass ceiling is to increase the ground temperature by a factor of

$$\frac{T(\text{glass})}{T(\text{no glass})} = \frac{1}{\beta^{1/4}} > 1. \quad (9.91)$$

This temperature increase is called the *greenhouse effect*.

Earth’s atmosphere can play the role of the glass: it allows most of the Sun’s spectrum through to the ground, but scatters almost all of the predominantly infra-red radiation leaving the ground. Hence, about half of J_o escapes the atmosphere (and half is scattered back to the ground), and thus $\beta = 1/2$. In that case, what is Earth’s average temperature when this atmosphere is included? Equation (9.91) then says

$$\frac{T(\text{atmosphere})}{T(\text{no atmosphere})} = \frac{1}{\beta^{1/4}} = 2^{1/4}. \quad (9.92)$$

The atmosphere thus causes the surface temperature to increase to

$$T(\text{atmosphere}) \stackrel{(9.88)}{=} 2^{1/4} \times 236 \text{ K} = 281 \text{ K} = 8^\circ\text{C}. \quad (9.93)$$

This prediction compares well with the measured average temperature of around 14°C —this value is not very precise, because different ways of defining the average combine with huge variations across Earth’s surface, and from day to night, to produce different numbers. Aside from our requiring an atmosphere for our breathing, we see here how the greenhouse effect of our atmosphere blanketing Earth’s surface makes our planet far more habitable for humans than it would be without an atmosphere.

9.9 Photon Absorption and Emission: the Laser

The laser was developed in stages by various teams of researchers during the middle years of the twentieth century. The basic theory for it was introduced in 1917 by Einstein, who made use of the Boltzmann distribution to study a population of excited atoms. We will formulate Einstein’s approach in this section.

Start with a set of atoms in thermal equilibrium in an oven at temperature T , which is usually called a *cavity* in the context of lasers. Each of these atoms is able to occupy either of two energy levels E_1 and E_2 , where $E_1 < E_2$. The numbers of atoms in each level, N_1 and N_2 , follow the Boltzmann distribution:

$$N_i \propto \exp \frac{-E_i}{kT}. \quad (9.94)$$

The energy gap between the levels can be written as $E_2 - E_1 = hf$ for some frequency f . When immersed in a bath of radiation, atoms can jump between the levels: those at level 1 can absorb a photon of frequency f and jump to level 2, and (even if the radiation bath is not present) those at level 2 can emit a photon of frequency f and drop to level 1. In equilibrium, the rate of atoms jumping up a level must equal the rate of those dropping down. In a time dt , various numbers of atoms switch levels in any of the following three ways, as shown in Figure 9.15.

Three Non-Thermal Ways to Jump Between Energy Levels

1. **Absorption:** Radiation in the oven with frequency f can stimulate an atom in level 1 to jump to level 2, with that photon being absorbed. The number of atoms that make this jump in time dt is postulated to be proportional to the number of atoms in level 1, the length of time dt , and the electromagnetic energy per unit frequency f per unit cavity volume, which is the spectral energy density $\rho(f)$ of (9.43).

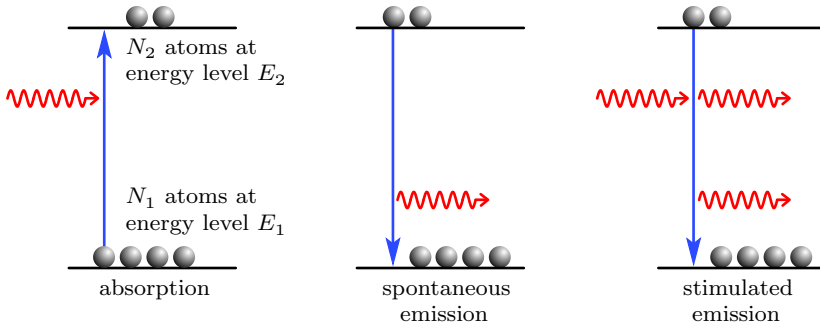


Fig. 9.15 The three ways in which a two-level atom can interact with (or create) an electromagnetic field

In this time, the decrease in the level-1 population N_1 due to this process is

$$-dN_1^{\text{abs}} = K^{\text{abs}} N_1 \varrho(f) dt, \quad (9.95)$$

with constant of proportionality $K^{\text{abs}} > 0$.

2. **Spontaneous Emission:** Atoms in level 2 can drop to level 1 spontaneously, emitting a photon of frequency f in the process. This number of atoms is postulated to be proportional to the number of atoms in level 2, and to the time interval dt . So, in this time, the increase in the level-1 population N_1 due to this process is

$$dN_1^{\text{spon}} = K^{\text{spon}} N_2 dt, \quad (9.96)$$

where $K^{\text{spon}} > 0$ is the constant of proportionality.

3. **Stimulated Emission:** Frequency- f radiation in the cavity can stimulate atoms in level 2 to drop to level 1, emitting a photon of frequency f in the process. This number of atoms de-exciting in time dt is postulated to be proportional to the number of atoms in level 2, the energy density $\varrho(f)$, and dt . So, in this time, the increase in the level-1 population N_1 due to this process is

$$dN_1^{\text{stim}} = K^{\text{stim}} N_2 \varrho(f) dt, \quad (9.97)$$

with constant of proportionality $K^{\text{stim}} > 0$.

In equilibrium, N_1 doesn't change, and so the total dN_1 from all sources equals zero:

$$dN_1 = dN_1^{\text{abs}} + dN_1^{\text{spon}} + dN_1^{\text{stim}} = 0. \quad (9.98)$$

Equations (9.95)–(9.97) convert the second equality in (9.98) to

$$-K^{\text{abs}} N_1 \varrho(f) + K^{\text{spon}} N_2 + K^{\text{stim}} N_2 \varrho(f) = 0. \quad (9.99)$$

It follows that

$$\begin{aligned} \varrho(f) &= \frac{K^{\text{spon}} N_2}{K^{\text{abs}} N_1 - K^{\text{stim}} N_2} = \frac{K^{\text{spon}}}{K^{\text{abs}} \frac{N_1}{N_2} - K^{\text{stim}}} \\ &= \frac{K^{\text{spon}}}{K^{\text{abs}}} \frac{1}{\frac{N_1}{N_2} - \frac{K^{\text{stim}}}{K^{\text{abs}}}}. \end{aligned} \quad (9.100)$$

But, with the two levels differing in energy by hf , their ratio of occupation numbers is

$$\frac{N_1}{N_2} \stackrel{(9.94)}{=} \frac{\exp \frac{-E_1}{kT}}{\exp \frac{-E_2}{kT}} = \exp \frac{E_2 - E_1}{kT} = \exp \frac{hf}{kT}. \quad (9.101)$$

This converts (9.100) to

$$\varrho(f) = \frac{K^{\text{spon}}}{K^{\text{abs}}} \frac{1}{\exp \frac{hf}{kT} - \frac{K^{\text{stim}}}{K^{\text{abs}}}}. \quad (9.102)$$

Compare this with (9.43), which is

$$\varrho(f) = \frac{8\pi hf^3}{v^3} \frac{1}{\exp \frac{hf}{kT} - 1}, \quad (9.103)$$

where we have written the speed of light as v rather than c , to emphasise that the light might not be travelling in a vacuum: v equals the vacuum-inertial value c divided by the refractive index of the medium. [See the grey box just after (9.106).] From (9.102) and (9.103), we infer that

$$K^{\text{spon}}/K^{\text{abs}} = 8\pi hf^3/v^3, \quad \text{and} \quad K^{\text{stim}} = K^{\text{abs}}. \quad (9.104)$$

The equality $K^{\text{stim}} = K^{\text{abs}}$ suggests that these symbols be replaced with a single symbol. Einstein used the symbols “ A ” and “ B ” in his original discussion of this subject, as follows:

$$A \equiv K^{\text{spon}}, \quad B \equiv K^{\text{stim}} = K^{\text{abs}}, \quad \text{and so} \quad \frac{A}{B} \stackrel{(9.104)}{=} \frac{8\pi hf^3}{v^3}. \quad (9.105)$$

These have come to be called the *Einstein A and B coefficients* that describe photon emission and absorption in a two-level system. With A and B , equations (9.95)–(9.97) are written more conventionally as

$$\begin{aligned}
dN_1^{\text{abs}}/dt &= -BN_1 \varrho(f), \\
dN_1^{\text{spn}}/dt &= AN_2, \\
dN_1^{\text{stim}}/dt &= BN_2 \varrho(f).
\end{aligned} \tag{9.106}$$

Light's Speed, Frequency, and Wavelength in a Laser Cavity

Because the electromagnetic waves in a laser cavity do not necessarily travel in a vacuum, their speed v is generally not the usual vacuum-inertial speed of light $c = 299,792,458$ m/s. Their speed can instead be written as $v = c/n$, where c is the vacuum-inertial speed of light and n is the refractive index of the “lasing medium” that fills the cavity.

In contrast, the frequency f of the light waves is not affected by the refractive index of the lasing medium. That is, when waves of frequency f enter a “linear” medium that has no effect other than to change their speed, their frequency doesn’t change. To see why, imagine tapping a long table at one tap per second, and allowing the taps to propagate through the table and “emerge” from the other end (perhaps to tap something else). The speed of waves through the table might be very fast or glacially slow, but—in a steady state—one tap per second *must* emerge from the other end, since, otherwise, taps would either be created or stored somewhere, which is not the way taps work. (The frequency of light *can* be changed by passing the light through exotic non-linear media, but that does not affect the above argument.)

But, whereas the tap frequency doesn’t change in the medium of the table, the tap wavelength (the distance between “tap waves” in the table) can well change, since it equals the propagation speed divided by the tap frequency: $\lambda = v/f$. So, the wavelength of the waves in the laser cavity *is* changed by the refractive index of the lasing medium.

Although we can write (9.105) as “ $A/B = 8\pi h/\lambda^3$ ”, it can be all too easy in a laser context (where the light might not travel in a vacuum) to forget that λ differs from the usual value that we might associate with, say, red light. In contrast, writing $A/B = 8\pi h f^3/v^3$ carries less risk, since f is independent of the lasing medium, as long we remember that v is not necessarily the vacuum-inertial speed of light.

As its acronym-name implies, the laser is a device for “light amplification by stimulated emission of radiation”: we wish to stimulate a set of excited atoms to drop to a lower energy level, causing them to emit a concentrated set of photons that are *coherent*. By “coherent”, we mean that the light waves emitted by successive de-exciting atoms have phase differences relative to each other that are not random. In Section 1.3.2, we discussed the relation of the random walk to the intense brightness of coherent light, but this brightness

is really a minor attribute of such a beam. Its main attribute is its coherence, which gives it a well-defined behaviour in procedures that make use of wave interference. This behaviour is the key to the laser's striking properties.

To investigate this stimulated emission, begin by comparing the various rates of emission and absorption. Refer to (9.106), to write

$$\begin{aligned}\text{absorption rate} &= -dN_1^{\text{abs}}/dt = BN_1 \varrho(f), \\ \text{spontaneous emission rate} &= dN_1^{\text{spont}}/dt = AN_2, \\ \text{stimulated emission rate} &= dN_1^{\text{stim}}/dt = BN_2 \varrho(f).\end{aligned}\quad (9.107)$$

Using (9.105) and (9.43) [and remembering to replace c in (9.43) with v], we have

$$\frac{\text{stimulated emission rate}}{\text{spontaneous emission rate}} = \frac{BN_2 \varrho(f)}{AN_2} = \frac{1}{\exp \frac{hf}{kT} - 1}. \quad (9.108)$$

At low temperatures ($kT \ll hf$), this ratio is approximately $\exp[-hf/(kT)]$. For an example, consider two-level atoms that emit 2 eV photons, which will emerge from the laser with a wavelength of about 620 nm, which is red light. When these atoms have a temperature of $T = 300$ K, the low- T version of (9.108) becomes (using $hf = 2$ eV, and SI units throughout)

$$\frac{\text{stimulated emission rate}}{\text{spontaneous emission rate}} \simeq \exp \frac{-hf}{kT} \simeq \exp \frac{-2 \times 1.602 \frac{-19}{\text{eV}}}{1.381 \frac{-23}{\text{eV}} \times 300} \approx 10^{-34}. \quad (9.109)$$

We conclude that stimulated emission can be completely ignored at room temperature for a collection of atoms in equilibrium. Only at much higher temperatures does this rate lift: at $T = 3000$ K, the ratio is about 10^{-4} ; and at 30,000 K, it is 0.86 [here, we must use the exact expression (9.108), not its low- T approximation].

The rate of stimulated emission compared to that of absorption is

$$\frac{\text{stimulated emission rate}}{\text{absorption rate}} = \frac{BN_2 \varrho(f)}{BN_1 \varrho(f)} = \frac{N_2}{N_1}. \quad (9.110)$$

In thermal equilibrium, the populations N_1 and N_2 are constant, with $N_2 < N_1$ as set by the Boltzmann distribution. But (9.110) suggests that if we can create an artificial *population inversion* $N_2 \gg N_1$, a high stimulated emission rate will result. This is the idea behind the operation of a laser. Stimulating the atoms to emit a high amount of coherent radiation means that an intense coherent beam of light exits the system, and this light can be put to use in ways that exploit this coherence.

This idea of producing a population inversion will not work with just two energy levels, since exciting atoms from level 1 to level 2 will be accompanied by an equal number de-exciting spontaneously from 2 back to 1. One

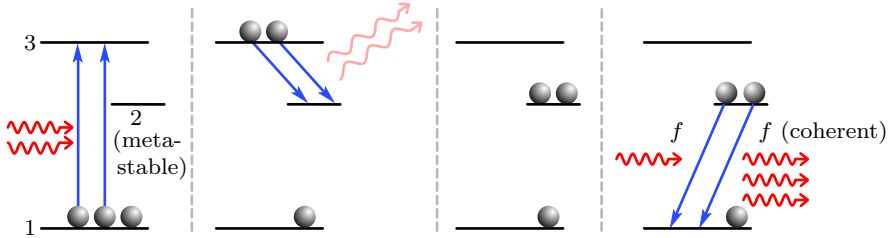


Fig. 9.16 The archetypal sequence of events that generate laser light. **First picture from left:** The three energy levels, with the first level occupied by some atoms. Incoming photons (possibly incoherent) of a carefully chosen frequency excite the atoms to energy level 3. **Second picture:** The atoms quickly de-excite to level 2, the “meta-stable level” that has a comparatively long occupancy lifetime. The light emitted is incoherent, since each atom de-excites in its own time. **Third picture:** The comparatively large number of atoms that are now in this meta-stable level is the required “population inversion”. **Fourth picture:** An incoming photon now stimulates these level-2 atoms to drop back to level 1, emitting photons in the process. Along with the incoming photon of frequency f , the emerging photons are coherent and of the same frequency

alternative is to find a material that has a third level with a comparatively long lifetime, a *meta-stable level*, shown in Figure 9.16. Level 2 in the figure is this meta-stable level. The process begins when we use light whose frequency matches the energy difference between levels 1 and 3 to “pump” atoms from level 1 to level 3. These quickly de-excite, dropping down to the meta-stable level 2, where they have a comparatively long lifetime of occupation. This is the required population inversion $N_2 \gg N_1$ that the Boltzmann distribution says would not be achieved if we could only rely on thermal interactions to pump the atoms to excited states. An incoming photon of frequency f just matches the energy difference hf between levels 1 and 2, and it stimulates atoms in level 2 to drop back to level 1. As they de-excite, they emit frequency- f photons that are coherent with each other and with the stimulating photon. Some of these emerging photons go on to stimulate other atoms at level 2 in a cascade of de-excitation that produces a flood of coherent photons. The laser cavity is enclosed by semi-reflecting mirrors. These ensure that some of the radiation bounces back and forth and builds up in intensity, while a fraction of this intense coherent light escapes to form the laser beam.

The parameters describing this laser can be related to give a condition for the device to operate. Let $\mathcal{N}$ be the number of frequency- f photons produced by any process. Refer to (9.107), noting that the absorption rate *decreases* $\mathcal{N}$, whereas the two emission rates *increase* $\mathcal{N}$. Additionally, using the conventional symbol γ for a photon, we introduce what might be called a “beam production rate” that is set by the length of time τ_γ that a photon remains in the cavity before it escapes.

The Beam Production Rate

What is this beam production rate? The escape of photons is a random process that follows the same mathematics as the discussion of radioactive decay in Section 6.6. There we saw, in (6.105), that the rate of radioactive decay of N atoms is $-dN/dt = fN/T$, with these symbols defined in that section. We also saw, in (6.109), that the mean lifetime of the atoms is T/f . It follows that $-dN/dt = N/(\text{mean lifetime})$. The same idea applies to the beam production rate: the rate of loss of photons from the cavity is $-d\mathcal{N}/dt = \mathcal{N}/\tau_\gamma$.

The four rates of increase of $\mathcal{N}$ are then as follows:

Process	$d\mathcal{N}/dt$
absorption	$-BN_1 \varrho(f)$
spontaneous emission	AN_2
stimulated emission	$BN_2 \varrho(f)$
beam production	$-\mathcal{N}/\tau_\gamma$

The total rate of increase of photon number in the cavity is the sum of the four terms in the above table:

$$\frac{d\mathcal{N}}{dt} = AN_2 + B(N_2 - N_1)\varrho(f) - \frac{\mathcal{N}}{\tau_\gamma}. \quad (9.111)$$

What is the spectral energy density $\varrho(f)$? Recall, from (9.1), that $\varrho(f)$ is the amount of electromagnetic energy in the cavity, per unit frequency f per unit cavity volume. The transition “level 2 to level 1” produces a non-zero spread of frequencies around f , known as the *line width* Δf . The cavity has a volume V containing $\mathcal{N}$ photons, each of energy hf . Thus,

$$\varrho(f) = \frac{\text{total energy}}{\text{line width} \times \text{volume}} = \frac{\mathcal{N}hf}{\Delta f V}. \quad (9.112)$$

Equation (9.111) is now

$$\frac{d\mathcal{N}}{dt} = AN_2 + \frac{B(N_2 - N_1)\mathcal{N}hf}{\Delta f V} - \frac{\mathcal{N}}{\tau_\gamma}. \quad (9.113)$$

Picture the laser as a long tube with a mirror at each end. The cascade of photons produced by stimulated emission bounces back and forth along its length axis. The $\mathcal{N}$ photons present are mostly a huge number $\mathcal{N}_{\text{ax}}$ of axially directed photons, with the remaining *much* smaller number of photons $\mathcal{N} - \mathcal{N}_{\text{ax}}$ leaking through the sides. We’ll ignore these leakage photons, so that $\mathcal{N} \simeq \mathcal{N}_{\text{ax}}$. Also, spontaneous emission creates photons moving in all directions, with very few of them emerging along the axis; hence, we can ignore

the contribution AN_2 to what is essentially $d\mathcal{N}_{\text{ax}}/dt$ in (9.113). Focussing on the axially directed photons, we extract the following from (9.113):

$$\frac{d\mathcal{N}_{\text{ax}}}{dt} \simeq \frac{B(N_2 - N_1)\mathcal{N}_{\text{ax}}hf}{\Delta f V} - \frac{\mathcal{N}_{\text{ax}}}{\tau_\gamma}. \quad (9.114)$$

Now recall, from (9.105), that

$$B = \frac{Av^3}{8\pi hf^3} = \frac{Ac^3}{8\pi hf^3 n^3}, \quad (9.115)$$

where we have written the photons' speed v as c/n , with c being the usual vacuum-inertial value (299,792,458 m/s), and n the refractive index of the medium in the laser cavity.

What is the value of A ? Recall that this constant represents spontaneous emission. If spontaneous emission were the only way for the N_2 atoms in level 2 to de-excite, we could write the middle line of (9.106) as

$$-dN_2/dt = AN_2. \quad (9.116)$$

This is solved easily, to yield

$$N_2(t) = N_2(0) e^{-At}. \quad (9.117)$$

This equation is just like radioactive decay, and if it *did* represent a radioactive decay, the mean lifetime of an atom before it decayed would be $1/A$: we proved this earlier in (6.106)–(6.109), but will prove it again here in the current context. Realise that $-dN_2$ laser atoms at level 2 de-excite in a time dt at each moment t , which implies that these $-dN_2$ atoms have survived for a time t . If we begin counting de-excitations at time $t = 0$, the mean lifetime of the $N_2(0)$ atoms in level 2 before they de-excite spontaneously will be some τ_2 , where

$$\begin{aligned} \tau_2 &\equiv \frac{\text{sum of lifetimes of all atoms}}{\text{total number of atoms}} = \frac{1}{N_2(0)} \int_0^{N_2(0)} (-dN_2 \times t) \\ &\stackrel{(9.116)}{=} \frac{1}{N_2(0)} \int_0^\infty AN_2(t) t dt \stackrel{(9.117)}{=} A \int_0^\infty e^{-At} t dt = \frac{1}{A}. \end{aligned} \quad (9.118)$$

The bottom line here is that $A = 1/\tau_2$, where τ_2 is the mean lifetime of atoms in level 2 that de-excite only spontaneously, not by being stimulated. This expression is now inserted into (9.115), to enable (9.114) to be written as

$$\frac{d\mathcal{N}_{\text{ax}}}{dt} \simeq \frac{c^3(N_2 - N_1)\mathcal{N}_{\text{ax}}hf}{8\pi hf^3 n^3 \tau_2 \Delta f V} - \frac{\mathcal{N}_{\text{ax}}}{\tau_\gamma}$$

$$= \frac{c^3(N_2 - N_1)\mathcal{N}_{\text{ax}}}{8\pi f^2 n^3 \tau_2 \Delta f V} - \frac{\mathcal{N}_{\text{ax}}}{\tau_\gamma}. \quad (9.119)$$

For laser action to occur, we require the number of axially directed photons $\mathcal{N}_{\text{ax}}$ to either remain constant or grow. So, we require $d\mathcal{N}_{\text{ax}}/dt \geq 0$, meaning the right-hand side of (9.119) must be greater than or equal to zero. Some rearrangement of that inequality then gives

$$\frac{N_2 - N_1}{V} \geq \frac{8\pi f^2 n^3 \tau_2 \Delta f}{c^3 \tau_\gamma}. \quad (9.120)$$

(Note that $\mathcal{N}_{\text{ax}}$ has now vanished.) The right-hand side of (9.120) is called the laser's *critical inversion per unit volume*. The smaller we can make this number, the easier it will be for the laser to operate in the laboratory. For example, putting good mirrors at each end of the cavity holds photons inside for longer, producing a large τ_γ that reduces the right-hand side of (9.120). But, of course, such mirrors also prevent laser light from escaping, and so we must seek a trade-off between ease of operation and amount of light produced.

What is a value for τ_γ ? Suppose the laser cavity has length L . The time it takes for a photon to traverse this length is $L/v = Ln/c$. In this time, *all* of the photons will collide once with a mirror. If the mirrors have reflectivity R , then (by definition of reflectivity), after this time, a fraction $1 - R$ of the equilibrium number of photons $\mathcal{N}_{\text{ax}}$ will have exited the cavity:

$$(1 - R)\mathcal{N}_{\text{ax}} \text{ photons exit the cavity in time } Ln/c. \quad (9.121)$$

It follows that

$$\mathcal{N}_{\text{ax}} \text{ photons exit the cavity in a time } \frac{Ln}{c(1 - R)}. \quad (9.122)$$

But if, say, all $\mathcal{N}_{\text{ax}}$ photons escape (and are replenished) every 3 seconds, we can ignore the tiny time it takes to create them, and state that each photon resides in the cavity for 3 seconds. This 3 seconds is then the lifetime of a photon in the cavity. That is,

$$\tau_\gamma = \frac{Ln}{c(1 - R)}. \quad (9.123)$$

This allows the critical-inversion equation (9.120) to be written as

$$\frac{N_2 - N_1}{V} \geq \frac{8\pi f^2 n^2 \tau_2 \Delta f (1 - R)}{c^2 L}. \quad (9.124)$$

Recall that we seek to make the right-hand side of (9.124) small. This equation places engineering limits on the laser: we see, for example, that a longer cavity length L makes for easier lasing. Also, cooling the whole system will

reduce the atoms' thermal motion, which has the effect of reducing Doppler broadening of the spectrum. This reduces the line width Δf , which then makes for easier lasing. Reducing the line width can also produce a better laser when we require the laser light to have as close as possible to a single frequency f . This is a common requirement in practical applications of the laser.

We commenced this last chapter with a speculative analysis of ovens that was based on the principle of detailed balance. The shape of the resulting spectrum has matched observations that cover electronic noise, hot glowing objects, our ability to survive on Earth, the noise that our radio receivers pick up from distant parts of the universe, and the lasers of modern technology. Planck's work a century ago was based on what were then new ideas, which would eventually become quantum mechanics; and the spectrum that he derived has proved to be a most useful tool of physics.

Lasers and radio receivers are a long way from the basic ideas that appeared at the start of this book. Statistical mechanics is sometimes described as a simple theory, because it rests on a single, straightforward proposition: that an isolated system in equilibrium is equally likely to be found in any of its microstates. In practice, making sense of that proposition requires the great effort that we have followed in this book. We have needed to define microstates, determine how to count them, describe the laws of thermodynamics, and introduce quantum concepts. Along the way, the subject constantly has had to calibrate itself against the "real world", and tackle problematic devil-in-the-detail ideas, such as the growth of entropy in complex systems, the proper use of the Boltzmann distribution, and the introduction of quantum concepts.

But, despite its success in explaining and predicting so many experimental results, the pure numbers game that is the entropy growth lying at the heart of statistical mechanics should not be accorded too much explanatory power. While it can tell us with effectively complete certainty which way a movie should be run—the so-called "arrow of time"—it certainly does not explain why time seems to us to flow; rather, the growth of entropy occurs *within* a flowing time. And although entropy growth goes hand in hand with the operation of various forces in Nature (Section 3.13), I think that most physicists would consider far-fetched the idea that life itself is nothing more than a vacuity arising from the blind growth of entropy. After all, it is entirely unreasonable to suppose that an incredibly finely tuned initial condition plus nothing more than effectively random billiard-ball collisions of molecules and the completely random effects of quantum mechanics over several thousand million years have given rise to you who are reading these words that were written by me. Life is certainly far more than the result of such randomness; and although physics has long pondered our apparently free will, it has never settled on any really concrete ideas in this area. For now, at least, statistical mechanics has resisted being pushed to such extremes. But nonetheless, when applied to the world we see around us, the subject has a tremendous power to explain and predict a vast array of phenomena.

Index

Italicised page numbers denote where the entry has been defined. For subjects that are referred to many times closely following their initial definition, I have pointed only to the defining page.

Symbols

- Ω_{tot}
 - for identical-classical particles 114
 - for complex molecules 109
 - for free point particle in one dimension 96
 - for free point particle in three dimensions 98
 - for ideal gas of point particles 101
 - for ideal gas of rotating non-point particles 104
 - for ideal gas of rotating, oscillating diatomic molecules 106
 - for lattice of point oscillators in one dimension 107
 - for lattice of point oscillators in three dimensions 108
 - for massless particles 119
 - summarised for gas and lattice 114
- dQ and Q , but never ΔQ 138
- ∇ in any coordinates 55
- ∇^2 in heat equation 228
- Δ = increase in, not change in 39, 270
- for thermal interaction 148
- spectrum of 117
- adiabat 237
- “adiabatic” mis-used 140
- adiabatic wall 139
- “amplitude”, use of 395
- analytic continuation and functions 16
- angular velocity and momentum 94, 292
- arrangements and selections 5
- arrow of time 187, 534
- atmosphere
 - gaseous makeup 358
 - height, weight, pressure 250, 278
 - layers 336
 - pressure at Everest’s peak 261
 - role in greenhouse effect 523
 - temperature at great height 161, 333
 - temperature gradient 336, 353
- atomic bomb 350
- atomic sizes 296, 376
- average over time versus over ensemble 275
- Avogadro’s number 71, 157, 376

A

- absolute deviation 19
- accessible states
 - counting 83
 - defining entropy using 177
 - for system + bath 277

B

- band theory of solids 468
- bands
 - allowed and forbidden 469
 - valence and conduction 469
- bandwidth of signal 497

basis vectors 48
 bath denotes “environment” 275
 “the bends” in diving 252
 bias voltage in semiconductors 476
 Big Bang and early universe 517
 binomial theorem 22
 Birkhoff’s theorem 383
 bit rate in signals 499
 bits (binary digits) 329
 black body 482
 blackbody radiation 481, 514
 Bohr magneton 283
 Bohr radius 301
 Boltzmann’s constant 152
 Bose–Einstein condensation 430
 Bose–Einstein statistics 416
 boson
 in Debye’s model 410
 bosons 413
 Brandeis dice 319
 broadband in communications 499
 bulk modulus 45
 of water 252
C
 caloric 136, 173
 canonical quantity 276
 Carathéodory and reversible processes 175
 Carnot cycle 236
 Carnot’s theorem 234
 Celsius degrees versus degree Celsius 153
 central force 126
 chain rule of differentiation 57
 channel capacity 500
 chemical equilibrium 270
 chemical potential 146
 in phase changes 247
 zero for photons 487
 Clausius–Clapeyron equation 257, 267
 cloud production in atmosphere 358
 CMB radiation 517
 coefficient of
 isothermal compressibility 45, 198
 thermal conductivity 220
 thermal expansion 198
 viscosity 375
 coherent light waves 528
 collision cross section 366
 collision frequency 367
 combinations and permutations 9, 441
 compressibility (isothermal) 45

conductance 75
 conduction and conductivity
 electrical 458
 thermal 220, 376
 conductors, semiconductors, insulators 468
 conservation, local and global 224
 constant of motion 128
 constant-volume gas thermometer 159
 continuity equation 224
 convolution
 circular 232
 standard 230, 501
 cooking on mountains 261
 cooling a system to absolute zero 217, 246
 coordinate system, choosing 81
 cosmic microwave background 517
 Count Rumford experiments 136
 countability 323
 counting
 generalised 314
 gives particle statistics 432
 coupled pendula 395
 covariance matrix 39
 critical inversion in a laser 533
 cumulative probability 51
 Curie’s law 283
 current and flux 374
 current density
 electric 460
 thermal 220

D

data rates 496
 data-transmission theory 324
 de Broglie wavelength 88, 405
 Debye energy 397
 Debye temperature and function 399
 degeneracy 86
 of states 286
 degree of freedom 112
 delta function 43, 229, 392
 density
 of mass 50, 116, 117
 of probability 51
 of water at Titanic 254
 density of states 117
 for massive classical particles 419
 for monatomic gas 118
 inside oven 502
 inside resistor 489
 diamagnetism 279

Dickens, *A Christmas Carol* 330
 differentials 39
 exact and inexact 61
 differentiating under the integral sign
 35, 348
 diffusion
 constant 227
 equation 174, 227
 dimension versus unit 67
 dimensionless units 71
 diode 474
 directional derivative 222
 distinguishable particles 6, 91
 distributions
 binomial 8
 fluctuations in 17
 gaussian approximation of 25
 Boltzmann 278
 Bose–Einstein 416
 Fermi–Dirac 416
 gaussian (normal) 25
 Maxwell velocity and speed 338
 multinomial 10
 divergence theorem 224
 diving in the ocean 252
 Doppler shift 518
 Dulong–Petit law 210, 385

E

Earth
 motion in universe 519
 temperature with and without
 atmosphere 523
 eigenvalue and eigenvector of inertia
 tensor 293
 Einstein *A* and *B* coefficients 527
 Einstein tensor 382
 electric dipole and moment 142
 electric field 45
 electron volt 69
 emissivity of surface 520
 endothermic reaction 195
 energy
 fluctuation for system + bath 306
 Gibbs 196, 270
 Helmholtz 193, 312
 in orbital motion 128
 internal 87
 kinetic 87, 128
 level, state, band 283
 non-quadratic terms 305
 of harmonic oscillator 484
 potential 87, 128

 potential and force 132
 potential is additive 134
 quadratic terms 112
 ensemble 23
 types of 275
 enthalpy 195
 in Joule–Thomson process 241
 of vaporisation 259
 entropy 168
 additivity 168
 canonical example of growth 171
 Gibbs' expression 313
 growth versus cause and effect 190,
 266
 increase and interaction direction
 187
 of ideal gas of point particles 170
 of system + bath 312
 of water 218
 of written English 330
 Planck's constant only a guide 176
 Shannon 326
 using Ω_{tot} in place of Ω 170
 equilibrium
 chemical constant 271
 of isolated system 85
 thermal 152
 equipartition theorem 154
 for non-quadratic energy terms 305
 for resistor 488
 for system + bath 303
 ergodic assumption 2, 275
 error function 30
 numerical evaluation 36
 escape speed 359
 exothermic reaction 195
 expanding universe 518
 expected/expectation value 18

F

factorial in number theory 15
 Fermi
 energy 422
 sea 422, 446, 449, 454
 speed 449
 temperature 449
 Fermi–Dirac statistics 416
 fermions 413
 ferromagnetism 280
 flow directions from First and Second
 Laws of Thermodynamics 187
 fluctuation

- in accessible microstates for thermally interacting systems 152
- in binomial distribution 17
- in large systems 28
- relative, in particle number 22
- flux density 373
 - of solar radiation 523
 - of thermal energy 220
- four-current and vectors 226
- Fourier transform in signal processing 232, 498, 501
- free expansion 161
- free expansion of a gas 139, 183
- frequency-to-wavelength spectral conversion 45
- friction as irreversible process 175
- full width at half maximum 150, 348
- fundamental postulate of statistical mechanics 85

G

- galactic recession 518
- gamma function 16
- gas constant R 157
- Gauss's theorem 224
- gaussian function and delta function 229
- gaussian integral 30
- general theory of relativity 382, 518
- generalised force and displacement 145
- Gibbs and reversible processes 175
- Gibbs–Duhem equation 198, 258, 267
- Gibbs' paradox 404
- GPS satellites 384, 501
- gradient operator 55
- Green function for heat equation 230
- greenhouse effect 523, 524

H

- heat
 - certainly a noun 174
 - current 221
 - engine 233
 - equation 227
 - lost through roof 223
 - sink 236
 - transfer
 - quantifying flow 220
 - used in defining temperature 148
- heat capacity 73, 207
 - Debye's model 394
 - Einstein's model 385

- of aluminium and water 213
- of copper 400, 453
- of diatomic gases 291
- of helium gas 211
- ratio for P and V 211
- specific and molar 209
- valence electrons' contribution 445
- helium, liquid 431
- histogram 29, 339
- Hooke's Law 133
- hydrogen energy levels 284, 287
- hydrostatic support, equation of 203

I

- ideal gas 101
 - law derived 154, 185
- identical-classical particles 6, 114, 404, 411, 443
- incoherence of light waves 24
- indistinguishability of money in bank 138
- infinitesimals 39
- information theory 324
- ink mixed in water 2, 324
- interaction types in First Law 137
- isentropes 237
- isotherm 236
- isothermal compressibility 198

J

- jacobian matrix 57
- jansky 521
- Jaynes, E.T. 319
- Johnson noise 498
- Joule and caloric 137
- Joule–Thomson process 238
 - coefficient 241

K

- kelvin (SI unit) 153
- Kirchhoff's laws 45
- Kruskal–Szekeres coordinates 383

L

- labelling plot axes 78
- Lagrange multipliers 321, 437
- lapse rate of atmosphere 336
- latent heat 259
 - of fusion 208, 218, 269
 - of vaporisation 259

laws of thermodynamics

Zeroth 137

First 138

Second 169, 234

not really a law 169, 187, 233

Third 216

LED (light-emitting diode) 479

Legendre transform 193

Lenz's law 280

light-emitting diode 479

line width of laser 531

linear operator 20

logarithm 30

better to fit a peak 150

used in analysing distributions 26

Lorentz transform 382

Lorenz number 467

Loschmidt, J.J. 157, 376

M

macrostate 83

magnetic dipole and moment 144, 177, 280

magnetisation 282

mass action law in chemistry 272

maximum inversion temperature 245

Maxwell

relation 198

speed and velocity distributions 337

Maxwell–Boltzmann statistics 417

mean

arithmetic 123

cubic harmonic 123, 401

free path 365

generalised 123

geometric 123

harmonic 123

speed of molecules 347

median speed 347

meta-stable level in laser 530

microstate 83

as cell of phase space 95

number determined by volume of phase space 96

Miller and Kusch experiments 345

molar mass 72

mole 71

moment of inertia 105, 106, 292

momentum

canonical 91

in special relativity 119

most likely speed 347

N

noise factor of resistor 498

normal mode of vibration/oscillation 389, 395

notation for calculations 64

Nyquist's theorem for noise 498

O

occupation number for crystal 390

ocean, pressure and density 251

Ohm's rule 45, 222, 460

one-forms and infinitesimals 48

orbital mechanics 125, 181, 205, 301, 522

osmosis, osmotic pressure 190, 264

oven, idealised 484

model for resistor 488

P

paramagnetism 280

partial derivatives 53

partition function 284, 307

Pauli exclusion principle 405

“per unit”, care needed 74, 372

permutations and combinations 9

pH 262

phase space 91, 410

phase transitions 257

phonon 119, 396

photon 119, 486

piston in car engine 85

Planck

Vorlesungen über Thermodynamik 176

quantum mechanics 484

reversible processes 176

Planck's constant 89

used in counting states 176

Planck's law for radiation 506

polarisation of photons and phonons 120

population inversion 180, 529

position–momentum and phase space 91, 114

potential 128

chemical 146

gravitational, electrostatic 128

pressure at core of Sun 200

principal axes of rotation 104, 292

principle of detailed balance 481

processes

adiabatic 237
 cyclic 175
 diffusive 196, 247
 endo- and exothermic 195
 isobaric nondiffusive 195
 isothermal isobaric 196
 isothermal nondiffusive 193
 quasi-static 85, 140
 reversible 174, 233
 role of friction 175
 transport 365
 PV diagram 184

Q

quadratic energy terms 110
 quadrature addition 349
 quantum cosmology 6
 quantum nature, extent of for a system 405
 quasi-stasis 142
 quasi-static process 85, 140
 quaternion 48

R

radar 501
 radioactive decay 362
 rain production in atmosphere 358
 random walk 23
 Rayleigh and Jeans 506
 reaction, direction of 270
 rectifier in a circuit 478
 red shift in cosmology 518
 reduced mass of molecule 107, 295
 reflectivity of surface 520
 relativistic mass 119, 295
 relaxation time 85, 162
 removable discontinuity 327
 representation of an object 81, 292
 reservoir denotes “environment” 275
 resistance and resistivity
 electrical 459
 thermal 223
 resistor, resistance, resistivity 459
 rest mass 119, 161
 reverse-biasing a diode 478
 reversible process 174, 233
 R -factor in building trade 223
 Riemann sum 12
 Riemann zeta function 17, 428
 rms
 deviation 19
 speed 348

rocket equation 46
 rotation of molecules 292

S

Sackur–Tetrode equation 170, 255, 279, 311, 403, 408
 salinity of humans 266
 salt water
 drinking 265
 melting and boiling points 268
 scalar 162, 379
 field 220
 Schrödinger equation’s solution
 for hydrogen 284
 for lattice of atoms 468
 for particle in a box 89
 Schwarzschild spacetime and black hole 383
 semiconductor 470
 series, asymptotic and convergent 13
 Shannon, C. 324
 Shannon–Hartley theorem 499
 shuffling cards 5
 signal processing 232
 simple harmonic motion 91
 smoothening, not smoothing 231
 solutes and solvents 262, 265
 specific heat 209
 spectral energy density 482
 speed of sound 397, 401
 speeds of molecules 304, 346
 spherical polar coordinates 38, 512
 spin
 and polarisation 120
 in quantum mechanics 416, 419
 standard deviation 19
 state variable 61
 states
 density of 116, 117
 for monatomic gas 118
 number accessible at given energy 87
 of a system 83
 spacing in energy 90
 various definitions 410
 statistical quantities related geometrically 350
 statistics of particles 409
 from counting argument 432
 steam engine invention 136
 Stefan–Boltzmann constant 510
 stellar birth and composition 201
 Stirling’s rule 12

stoichiometric coefficients 271
 suction cup 250
 Sun as a black body 515
 support of a function 229, 352

T

temperature 152
 negative? 177
 of excitation 290
 of onset of rotation 293
 of onset of vibration 298
 of star 289
 throughout Sun 200
 tensor 82, 379
 energy–momentum 380
 moment of inertia 105, 292
 rotation 292
 thermal expansion 198
 thermal wavelength 408
 Tolstoy, War and Peace 330
 transport process 371
 triple point of water 153

U

ultraviolet catastrophe 507
 uniformly accelerated observer 188, 383
 units and dimensions 64

V

validity of kinetic/atomic models 378
 van der Waals' equation 165
 constants in molar form 167
 in Joule–Thomson process 242
 table of a and b constants 246
 vapour pressure 258

variable names in computer programming 68
 variable stars 201
 variables
 conjugate pairs in First Law 182
 intensive and extensive 181, 311
 pairs used to construct phase space 95
 variance 19
 vector 162, 379
 velocity space 339
 Vermeer, Girl with a Pearl Earring 2
 vibration of harmonic oscillator 298
 virial
 expansion for gases 164
 theorem 163, 205
 viscosity 371
 vis-viva equation 301
 volume of hypersphere 102

W

water
 compressible and incompressible 255
 vapour pressure and boiling point calculation 261
 wave number and vector 491
 wave states for light 484, 490, 502
 ways to sample 441
 weight versus mass 131
 width of probability distribution 17, 341
 Wiedemann–Franz–Lorenz law 467
 Wien's law 515
 altered in radio astronomy 521
 work done on system 130

Z

zero-point energy 192, 387, 485