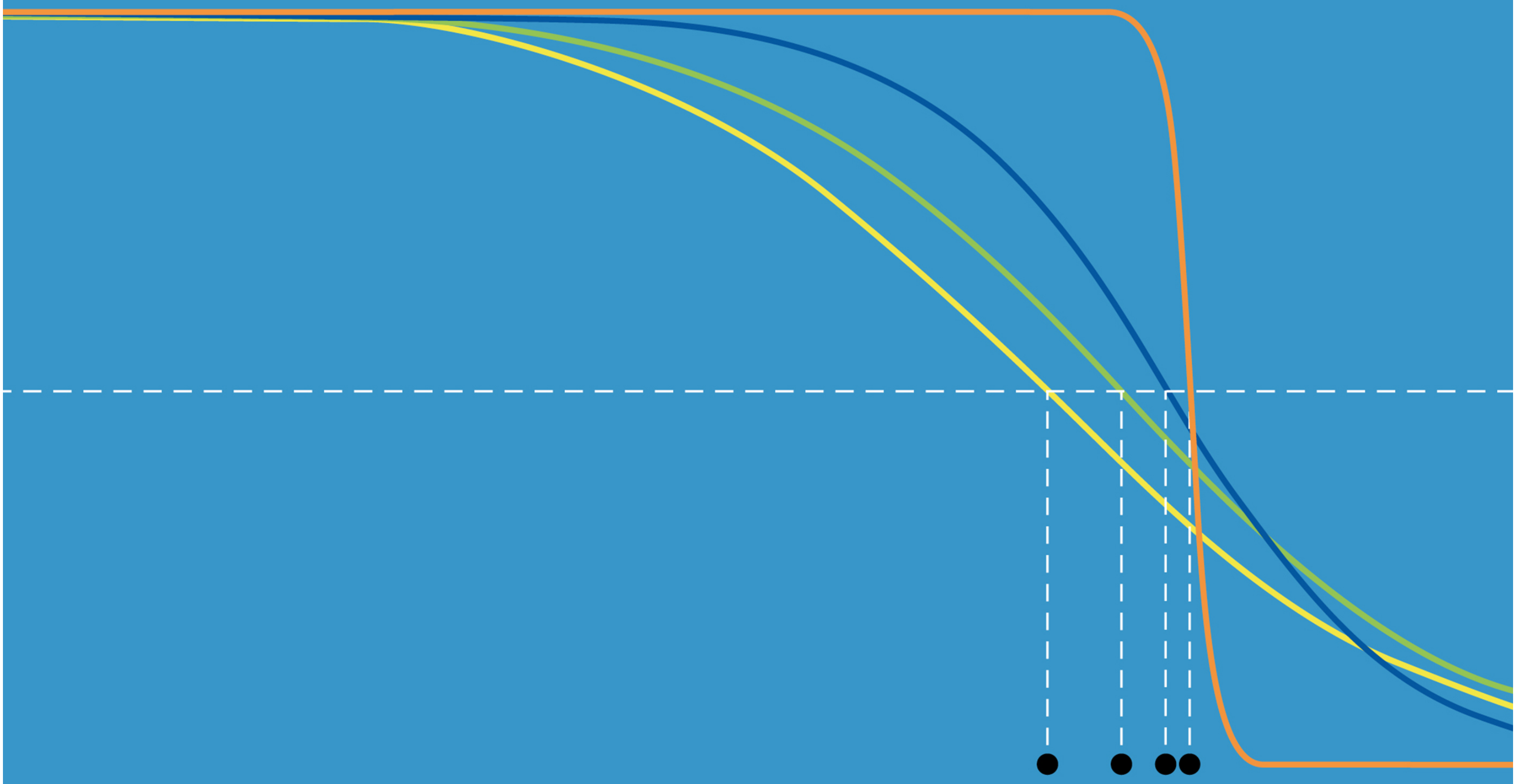
Statistical Physics of Condensed Matter Systems

A primer

Luciano Colombo



IOP | ebooks

Statistical Physics of Condensed Matter Systems

A primer

Online at: https://doi.org/10.1088/978-0-7503-2269-0

Statistical Physics of Condensed Matter Systems

A primer

Luciano Colombo
Department of Physics, University of Cagliari, Cagliari, Italy

IOP Publishing, Bristol, UK

ISBN 978-0-7503-2269-0 (ebook)
ISBN 978-0-7503-2266-9 (print)
ISBN 978-0-7503-2267-6 (myPrint)
ISBN 978-0-7503-2268-3 (mobi)

DOI 10.1088/978-0-7503-2269-0

Version: 20221101

IOP ebooks

British Library Cataloguing-in-Publication Data: A catalogue record for this book is available from the British Library.

Published by IOP Publishing, wholly owned by The Institute of Physics, London

IOP Publishing, No.2 The Distillery, Glassfields, Avon Street, Bristol, BS2 0GR, UK

US Office: IOP Publishing, Inc., 190 North Independence Mall West, Suite 601, Philadelphia, PA 19106, USA

Contents

Part I Classical statistical physics

Part II Quantum statistical physics

Part III Concluding remarks

Part IV Appendices

Foreword

The creation of a common European space for higher education, where students can easily train in a network of different European academic institutions has and continues to require new textbooks, separating the basics concepts taught in the first three years from the more complex and technical applications that characterize the last two years. These three volumes, primers in atomic and molecular physics, in solid-state physics, and in statistical mechanics greatly contribute to this need. As well explained by the author in his preface, these three volumes arise from the desire to expose students, even those who will leave their studies after the first three years, not so much to the casuistry of applications but to the beauty of the fundamental principles ruling the structure of matter. Forging thinking, rather than providing notions, is the spirit fuelling this editorial project.

It is not easy to write a primer on a subject as important and as vast as the statistical description of matter. It takes considerable effort to decide which topics to touch on and how to condense them into a clear and concise form, while maintaining the organic nature of the project. Luciano Colombo limits himself mostly to the description of non-interacting particles, whether classical or quantum, although a desire to go further transpires here and there in the text. This choice allows him to focus on the foundations of statistical mechanics, on how to *partition* particles among the energy levels, on how the seemingly simple concept of equal probability of microstates is transformed through the writing of a partition function into a powerful tool for predictions at the macroscopic level.

The primer in statistical mechanics is composed of five chapters. The first discusses the basic ideas defining statistical mechanics, the concept of ensemble, the correspondence between the average over time and the average over microstates, the process of equilibration as the search for the distribution that maximizes the appropriate probability conditioned to the proper constraints, for example, the conservation of the number of particles or the energy of the system. Special attention is given to the concept of temperature and entropy and how entropy teaches us about our ignorance of the microscopic state of systems. By the end of the chapter, the student has definitely understood that statistical mechanics offers a quite elegant microscopic explanation of thermodynamics and its axioms.

In chapter 2, the student learns how to build on the foundation just laid and how to predict the behaviour of systems of non-interacting particles (either atoms or molecules), even in the presence of electric and magnetic fields. Here, too, Luciano prefers to skip the commonly used description based on momenta and generalized coordinates, directly focussing on the alternative description based on energy levels, elegantly showing how the quantization of energy reveals its presence in the progressive unfreezing of rotational and vibrational degrees of freedom.

In chapter 3, Luciano introduces the quantum features associated with the particle wave function symmetry. You can perceive the amazement that students must feel as they realize that the simple imposition of the wave function particle exchange symmetry completely revolutionizes what had been learned for systems of distinguishable particles. Attraction or repulsion between particles arises as a consequence of quantum indistinguishability! Novel important concepts as the Fermi energy and the ground state energy are introduced and their physical role explained. The effect of quantum features on the temperature dependence of thermodynamic quantities and the difference with respect to the classical ideal gas emerge clearly and comprehensively.

Chapter 4 plays a role analogous to chapter 2. The power of the formalism learned in the previous chapter is shown in action, this time for quantum statistics. Here the temperature dependence of the properties of a free electron and phonon gases are discussed in detail and compared with experimentally measured trends. A short chapter 5 extends the fourth to two very important topics, black-body radiation and Bose–Einstein condensation.

I alluded that Luciano had to make a difficult choice in selecting the topics in this Primer. As I also said, despite the focus on non-interacting particles, here and there in the text, some important physics originating by the interaction between atoms and molecules peeps out. The effect of turning on attraction between atoms in classical gases is treated as a small interlude that stimulates and broadens the reader's horizon. Of course, these small inclusions cannot do justice to the many fields in which statistical mechanics has found remarkable applications in recent decades. It is nice, therefore, to find at the end of the Primer a list (chapter 6) of 'classical' topics not previously covered. Equally important is the short list of fields where statistical mechanics has played or is playing a fundamental role (critical phenomena, complex systems, non-equilibrium).

Finally, I want to emphasize the significant effort that has been made not only in the selection of the material but also in the way the material is presented. Each chapter is opened by a short syllabus which very clearly outlines the chapter content. The text is accompanied by a few figures, particularly impressive and well annotated, certainly helpful in conveying their message. Figures are also used to explain experimental techniques allowing scientists to measure the calculated quantities or to compare theoretical predictions with experimental results. Derivations are carried out step by step. The student is never left with that bitter sense of incomprehension engendered by missing steps. When the mathematical load is excessive, it is not suppressed but moved to the different appendices. This simplifies the reading of the text but at the same time offers the possibility of enjoying the continuity of the logical/mathematical process.

In short, an excellent text for students at their first encounter with statistical mechanics, which I am sure it will be adopted in several physics courses. Thank you Luciano for writing it.

Francesco Sciortino
Department of Physics
'Sapienza', University of Rome
July 2022

Presentation of the ‘Primer series’

This is the third volume of a series of three books that, as a whole, account for an introduction to the huge field usually referred to as ‘condensed matter physics’: they are respectively addressed to atomic and molecular physics, to solid-state physics, and to statistical methods for the description of classical or quantum ensembles of particles. They are based on my 20-year experience of teaching undergraduate courses on these topics for bachelor-level programs in physical and engineering sciences at the University of Cagliari (Italy).

The volumes are called ‘Primers’ to underline that the pedagogical aspects have been privileged over those of completeness. In particular, I selected the contents of each volume so as to keep limited its number of pages and so that the topics actually covered correspond to the *syllabus* of a typical one-semester course.

More important, however, was the choice of the style of presentation: I wanted to avoid an excessively formal treatment, preferring instead the exploration of the underlying physical features and always placing phenomenology at the centre of the discussion. More specifically, the main characteristics of this book series are:

- emphasis is always given to the physical content, rather than to formal proofs, i.e. mathematics is kept at the minimum level possible, without affecting rigour or clear thinking;
- an in-depth analysis is presented about the merits and faults of any approximation used, incorporating as well a thorough discussion of the conceptual framework supporting any adopted physical model;
- prominence is always on the underlying physical basis or principle, rather than to applications;
- when discussing the proposed experiments, the focus is given to their conceptual background, rather than to the details of the instrumental setup.

Despite the tutorial approach, I nevertheless wanted to follow the Italian academic tradition, which provides even the elementary introduction to condensed matter physics at a quantum level. I hope that my efforts have optimally combined ease of access and rigour, especially conceptual.

The intentionally non-encyclopaedic content and the tutorial character of these ‘Primers’ should facilitate their use even for students not specifically enrolled in a university *curriculum* in physics. I hope, in particular, that my textbooks could result accessible to students in chemistry, materials science and also of many engineering branches. In view of this, I have included a brief outline of non-relativistic quantum mechanics in the first ‘Primer’, a subject that does not appear in the typical engineering *curricula*. For the rest, classical mechanics, elementary thermodynamics and Maxwell theory of electromagnetism are used, to which all students of natural and engineering sciences are normally exposed.

Each ‘Primer’ is organised in parts, divided into chapters. This structure is tailored to facilitate the planning of a one-semester course: these volumes aim at being their main teaching tool. More specifically, each part identifies an independent teaching module, while each chapter corresponds to about two weeks of lecturing.

I cannot conclude this general introduction without thanking the many students who, over the years, have attended my courses in condensed matter physics at the University of Cagliari. Through the continuous exchange of ideas with them I have gradually understood how best to organise my teaching and the corresponding study material. As a matter of fact, the contents that I have collected in these volumes were born from this very fruitful dialogue.

Luciano Colombo
Cagliari, June 2019

Acknowledgements

I am really indebted with Ms C Mitchell—Commissioning Editor of IOP Publishing—and Mr Daniel Heatley—ebooks Editorial Assistant of IOP Publishing—for their great enthusiasm in accepting my initial proposal and for shaping it into a suitable editorial product. I also warmly acknowledge Ms Emily Tapp—Commissioning Editor of IOP Publishing—and Mr R Trevelyan—Editorial Assistant of IOP Publishing—for supporting my writing efforts of the second and the third Primers with great professionalism, always promptly replying to my queries and clarifying my doubts. Overall, their assistance has been really precious.

Author biography

Luciano Colombo

Luciano Colombo received his doctoral degree in physics from the University of Pavia (I) in 1989 and then he was a post-doc at the École polytechnique fédérale de Lausanne (CH) and at the International School for Advanced Studies (I). He became assistant professor (tenured) at the University of Milano (I) in 1990, next moving to the University of Milano-Bicocca (I) in 1996 for an equivalent position. In 1999 he was appointed associated professor at the University of Cagliari (I) where in 2002 he became full professor of theoretical condensed matter physics. Since 2015 he has been a fellow of the 'Istituto Lombardo—Accademia di Scienze e Lettere' (Milano, I). In 2021 he was appointed as Vice-Rector for Research of the University of Cagliari (I) and 'Managing Editor' of the 'Materials Physics' section of *The European Physical Journal—Plus*. He has been the principal investigator of several research projects addressed to solid-state and materials physics problems, the supervisor of more than 80 students (at bachelor, master, and PhD level), and the mentor of about 20 post-docs. He is the author, or coauthor, of more than 280 scientific articles and 9 books (this included). More about him can be found at: http://people.unica.it/lucianocolombo.

Introduction to: *Statistical Physics of Condensed Matter Systems : A primer*

This is the third and final volume in the series of three Primers I devoted to condensed matter physics. In the first two I developed the phenomenological, conceptual and formal bases for a quantum description of atoms and molecules (namely, its elementary constituents) and the crystalline state (that is: its specific form characterised by translational invariance), respectively.

A simplistic mechanistic-reductionist approach would lead one to consider the work concluded: once the fundamental principles are known, the rest is merely their systematic application to systems of increasing complexity. In fact, this way of understanding things is inefficient and badly fails to grasp a fundamental aspect of the physics of aggregates. On the one hand, as the number of elementary constituents (be they atoms or molecules) increases, a calculation that claims to explicitly consider all their degrees of freedom would soon become unmanageable, even with the aid of the most powerful high-performance computing systems. On the other hand, and even more significantly, such a brute-force approach would indeed be pointless: most of the details that dominate the atomic scale are averaged at the meso- and macro-scale, becoming irrelevant.

The conceptually important point is this: the macroscopic properties of an aggregate are just an average of single-constituent characteristics. Making a macroscopic prediction, therefore, means calculating the average value of any given physical observable which can be represented by microscopic (that is: defined at the atomic or molecular level) quantities. This is the conceptual core of the statistical description of condensed matter systems.

One might think that this challenge is already addressed by classical thermodynamics. Indeed, thermodynamics describes the thermal properties of matter (as well as their related processes) through the use of a small number of macroscopic variables, either intensive or extensive: it is, therefore, a coarse-grain theory. The statistical approach conceptually lies somewhere between the mechanistic-reductionist and thermodynamical approaches: its ambition is to describe matter at the most fundamental level (in other words: taking explicit account of its atomic-scale structure), without, however, claiming to describe in detail all the degrees of freedom associated with its elementary constituents. In practice, statistical physics predicts the macroscopic observables through appropriate ensemble averages of microscopic properties.

The statistical approach is very elegant since it provides results of paradigmatic importance through a very effective mathematical formalism; it is also effective in predicting the emergent phenomena occurring in aggregates, providing their most fundamental explanation (when dealing with atomistic properties, one can also operate at the quantum level). Because of these characteristics it cannot be missing from any tutorial introduction to condensed matter physics and for this reason I devoted the present third Primer to this topic.

This Primer is divided into two parts, respectively addressed to the statistical physics of classical and quantum systems. Each part initially contains a general chapter where the basic concepts and related mathematics are developed leading to the Boltzmann (for classical systems), Fermi–Dirac (for fermion quantum systems), and Bose–Einstein (for boson quantum systems) distribution laws, namely the three cornerstones of statistical physics. Next, the distribution laws are applied to a thorough investigation of the thermal properties of paradigmatically important systems such as the classical ideal gas, the electron gas and the phonon gas. Some specific phenomenologies are also discussed, the statistical foundations of which are treated in detail. They include, in particular, the paramagnetism, the black-body radiation (photon gas), and the Bose–Einstein condensation.

Eight appendices are added to the text, each focussed on some technical development which, at first reading, can be skipped without compromising the general understanding of the arguments developed in the main text. A bibliography is added to each chapter as a guideline for further reading. The volume contains many figures, most of which are ‘conceptual’, that is: they are basically intended to provide a graphical representation of the main ideas and results developed in the written part. Tables with numerical values of important physical properties are included as well, in the attempt to provide the reader with information useful to ‘quantify’ the physical results presented. Finally, a list of all the mathematical symbols used in the volume is given at the beginning as an orientation guide while reading.

I started writing the first of three Primers during the Fall 2018, encouraged by my wife, no longer with us now. As I close this experience, my thoughts and deepest gratitude go out to her.

Luciano Colombo
Cagliari, July 2022

Acknowledgements

I am really indebted with many friends who helped me by critically reading (in part or totally) the pre-editorial version of the book (in particular, Dr A Cappai, University of Cagliari, Italy). The care these friends have put in checking my original manuscript has corrected several unclear passages and many misprints, thus greatly improving my presentation. If there are still errors or omissions they should be attributed solely to me.

Symbols

λ	vibrational quantum number for a quantum harmonic oscillator
μ_{B}	Bohr magneton
μ_{c}	chemical potential
τ_{e}	scattering relaxation time for electrons
$\tau_{s\mathbf{q}}$	scattering relaxation time for phonons
χ_{C}	Curie paramagnetic susceptibility
Ω	the number of different ways a partition can be realised
A	atomic mass number
$\mathbf{B}$	magnetic field
$\mathcal{C}_{\mathrm{P}}$	molar constant-pressure heat capacity
$\mathcal{C}_{\mathrm{V}}$	molar constant-volume heat capacity
e	electron charge
$\mathbf{E}$	electric field
E_{F}	Fermi energy
$f(E, T)$	Fermi–Dirac occupation number
$\mathcal{F}$	Helmholtz free energy
g	Lande' g-factor
$G(E)$	density of energy levels
$G(\nu)$	spectral density of the electromagnetic radiation
$\mathcal{G}$	Gibbs free energy
h	Planck constant
$\hbar$	normalised Planck constant $\hbar = h/2\pi$
$\mathcal{H}$	enthalpy
j	quantum number associated to the atomic total (orbital+spin) angular momentum
$\mathbf{J}$	total (orbital + spin) atomic angular momentum
k_{B}	Boltzmann constant
l	quantum number associated to the atomic orbital angular momentum
m_{e}	electron mass
m_j	total magnetic quantum number
$\mathcal{M}$	magnetisation (magnetic moment per unit volume)
n	number of moles
$n(E, T)$	Bose–Einstein occupation number
$\mathcal{N}_{\mathrm{A}}$	Avogadro number
$\langle O \rangle$	ensemble average of the observable O
P	pressure
P_{BB}	total black-body electromagnetic power emitted per unit area
$\mathcal{P}$	polarisation (electric moment per unit volume)
$\mathcal{Q}$	heat
r	rotational quantum number for a diatomic molecule
R	universal gas constant
s	quantum number associated to the spin angular momentum
S	Boltzmann entropy
S_{Gibbs}	Gibbs entropy
T	temperature
T_{BE}	transition temperature for the Bose–Einstein condensation
T_{F}	Fermi temperature (electron gas)

T_{room}	room temperature (such that $k_{\mathrm{B}}T_{\mathrm{room}} = 0.025$ eV)
T_{rot}	rotational temperature (diatomic molecule)
T_{vib}	vibrational temperature (diatomic molecule)
$u_{\mathrm{BB}}(\nu)$	black-body spectral energy density
$\mathcal{U}$	total internal energy
$\mathcal{U}_{\mathrm{rot}}$	rotational internal energy of a gas
$\mathcal{U}_{\mathrm{tr}}$	translational energy of a gas
$\mathcal{U}_{\mathrm{vib}}$	vibrational energy of a gas or crystalline solid
v	particle velocity (ideal gas)
v_{mp}	particle most probable velocity (ideal gas)
$\langle v \rangle$	particle average velocity (ideal gas)
$\langle v^2 \rangle$	particle mean square velocity (ideal gas)
V	volume
Z_{v}	number of valence electrons per atom
$\mathcal{Z}$	partition function
$\mathcal{Z}_{\mathrm{GC}}$	grand canonical partition function
$\mathcal{Z}_{\mathrm{rot}}$	rotational partition function of a gas
$\mathcal{Z}_{\mathrm{vib}}$	vibrational partition function of a gas
$\overline{\mathcal{Z}}$	grand partition function
W	mechanical work

Part I

Classical statistical physics

IOP Publishing

Statistical Physics of Condensed Matter Systems

A primer

Luciano Colombo

Chapter 1

The statistical description of a classical system

Syllabus—*We introduce the basic concepts underlying the statistical description of a system made by a large number of classical (spinless) particles. At first, two fundamental challenges are taken up, namely how to find the most probable distribution law for its equilibrium state and how to estimate the corresponding macroscopic properties. Next, the concept of entropy is thoroughly presented and used to elaborate a robust microscopic foundation of classical thermodynamics. Eventually, the thermal physics of the monoatomic ideal gas is discussed, in particular deriving its equation of state and the Maxwell distribution laws for energy and molecular velocities. The statistical definition of irreversible process concludes this chapter.*

1.1 Basic concepts

Our modern physical picture of condensed matter defines ordinary systems as an assembly of interacting atoms or molecules [1, 2]. This holds for any aggregate, either solid, liquid or gaseous. According to this *atomistic picture*, any observed macroscopic property can be intended as the result of organised actions occurring among a really huge[1] number of elementary constituents. However, any phenomena emerging at the macroscale can hardly be described by considering each single constituent and its interplay with all the remaining ones, simply because there are too many degrees of freedom to be managed. Even more important: pretending to follow each atom or molecule would be a rather clumsy approach, since most of the atomic-scale details are in fact averaged out and, therefore, their overall inclusion in a physical model would be a pedantic solution adding no meaningful knowledge.

---

[1] In normal temperature and pressure conditions, one cm^3 of a gas typically contains as many as $\mathcal{O}(10^{19})$ constituents.

doi:10.1088/978-0-7503-2269-0ch1

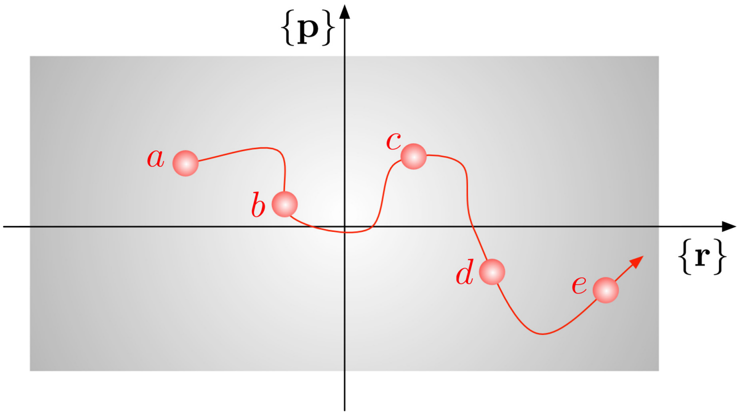


Figure 1.1. Pictorial representation of the phase space of a classical system. Different microstates a, b, … , e are shown. The sequence $a \to b \to \cdots \to e$ defines the trajectory in the phase space followed by the system during its time evolution. Axes define the full set of momenta $\{\mathbf{p}\}$ and positions $\{\mathbf{r}\}$ of the particles forming the system.

A rather different approach will be developed in this Primer, where processes involving very many elementary constituents (whose mutual interactions are known) will be described without considering the individual behaviour of each of them. We name this approach the *statistical description of a condensed matter system*. Basically, we aim at understanding the collective or macroscopic properties of an aggregate, without describing the motion of its constituents in full detail. The key point of this statistical description is that, while the atomistic picture is still adopted in describing the nature of the given aggregate, the single-particle mechanical approach will be replaced by a statistical one, providing predictions for the average (macroscopic) value of physical quantities.

In order to explain the basic concepts underlying statistical physics more rigorously, let us consider a classical system containing a *large number* N of particles. In this framework by 'classical system' we mean any aggregate of *spinless particles*; also, we will make no use of its quantum mechanical wavefunction in addressing the description of its physical state. If we assign an exact value to the space coordinates and momenta of each particle we define a specific *microstate* of the system. Such a microstate is associated with a single point in the $6N$-dimensional *phase space* of the system, whose axes represent all the coordinates and momenta, as shown in figure 1.1. If the system state undergoes a time evolution, its representative point follows a trajectory in the phase space. For the system to follow a specific trajectory, a small number of thermodynamical (that is macroscopic) variables must be assigned; for instance, we might fix the total internal energy $\mathcal{U}$, the volume V, and the total number of particles N of the system[2]. This reduced set of parameters defines the *macrostate* of the system. Formally this is equivalent to saying that we are

---

[2] If we are dealing with a multi-component system, that is, a system containing two or more chemical species, we should define the number of particles for each chemical species. This is just an annoying detail, which does not affect our reasoning and, therefore, will be neglected.

forcing the trajectory to lie on a specific surface of the phase space, defined by the assigned thermodynamical parameters.

If we wait a long enough time, it is reasonable to assume that the system will fully explore the surface defining a macrostate, that is, *it will visit all the microstates consistent with the imposed macroscopic constraints*. The actual value of any physical observable is simply the average among all the values it assumed in each visited microstate. The averaging operation is referred to as an *ensemble average*, where the word 'ensemble' simply identifies the assembly of all microstates. This situation is summarised by stating that *an ensemble average is the same as a time average*. Systems obeying this condition are said to be *ergodic*. While this subtle concept is treated more in detail elsewhere [3–5], here we remark a key notion of the statistical approach, namely: under the ergodic assumption, the observed value of a property corresponds equally well to a time average performed over a long enough time or to an ensemble average performed over a rich enough assembly of microstates[3]. In other words, we could select a specific microstate, follow its dynamical evolution on the surface defined by the assigned macroscopic parameters, and average over the values it assumed in the visited microstates; alternatively, we could randomly select a large enough number of microstates on the same surface and take the average over them. The result will be just the same[4].

In order to proceed, we need to establish a link between a macrostate and its corresponding microstates. By keeping the situation as clean as possible, let us assume that the N particles forming the system are identical and structureless. Let us further assume that they can occupy states with energy E_i with $i = 1, 2, 3, \ldots$ (with $E_1 < E_2 < E_3 < \ldots$). Then, it is straightforward to write

$$N = \sum_i n_i \quad \text{and} \quad \mathcal{U} = \sum_i n_i E_i \tag{1.1}$$

where n_i is the number of particles with energy E_i; this number is unrestricted since we are treating an aggregate of classical (spinless) particles. We need to remark that, strictly speaking, the above definition of internal energy is only valid for a system of non-interacting particles, since only in this case can we unambiguously attach to each particle its energy; however, to a good approximation, it could be extended to a system of weakly interacting constituents (like e.g. a dilute gas at high temperature). In order to manage the case where interactions are not negligible, we could adopt a *single-particle approximation* [2]: the many-body problem is reduced to N single-particle problems, where each ith particle undergoes a self-consistent potential V_i^{scp} only depending on its

---

[3] It is understood that both averages are taken under the same thermodynamical constraints.

[4] This conclusion has very important practical consequences. By implementing a computational approach to statistical physics, we could equally well perform ensemble averages by following the deterministic trajectory (we are developing a classical theory!) of the representative point of the system in the phase space or by randomly selecting microstates lying on the same surface. The first approach is named 'molecular dynamics', the second one 'Monte Carlo method' [6]: they are the kernel of contemporary computational statistical physics.

coordinates. Such potential V_i^{scp} is local and effectively describes the interaction with all the remaining $(N - 1)$ particles in the system. Under this approximation $E_i = E_i^{\text{kin}} + V_i^{\text{scp}}$ where both the kinetic and the potential contribution to the single-particle energy appear. For strongly interacting systems where the single-particle approximation fails, other techniques must be developed, explicitly including interactions. This situation falls beyond the scope of this Primer.

If the system is isolated, its energy $\mathcal{U}$ is conserved. However, collisions[5] and interactions affect the *distribution of the particles among the energy states* which is referred to as *partition*; this implies that the occupation numbers n_i may change in time and the partition (or, equivalently, the microstate) may accordingly change. A natural question arises at this point: *which is the most probable partition, once the macrostate*[6] *of the system has been assigned*? It is of fundamental importance to answer this question, since *a system in its most probable partition is said to be at equilibrium*. This statement can be easily justified since it is quite obvious to assume that the free overall evolution of an isolated system is always from a partition with lower probability to a partition with higher probability.

When an equilibrium condition is reached, there is no change of macrostate unless a disturbing external action is applied to the system. On the other hand, statistical fluctuations may occasionally affect the partition (or, equivalently, the microstate), but they are ultimately unable to lead the system far from equilibrium: in other words, the n_i can fluctuate around their equilibrium values, without sizeable macroscopic effects. Once the fundamental question of statistical physics is solved[7] the next challenge is to develop theoretical methods for *predicting, by means of an ensemble average, the equilibrium values of the physical properties we are interested in*.

In concluding this introduction to the statistical approach, it is important to remark that the use of average-based arguments for predicting system properties does not mean that the physical picture derived from it is vague, or approximate or even just qualitative. On the contrary, we will develop a very quantitative and robust description of physical systems containing a large number of constituents. Rather, we should more correctly state that, whenever an observable is specified by a reduced number of macroscopic parameters, it will be described statistically, through averaging procedures. This is the reason why statistical physics is the basic theory underlying macroscopic thermodynamics. The first theoretical formulation of the statistical approach was developed by L Boltzmann, J C Maxwell, and J W Gibbs within classical physics in the years between the XIX and the XX century.

---

[5] Occurring even within a non-interacting system.

[6] That is, for the case we discussed so far, once its volume, number of particles, and energy have been assigned.

[7] In other words: once, given an isolated system with known composition, its most probable partition has been found.

1.2 Equilibrium distribution probability

In order to develop the quantitative calculation of the equilibrium partition for a classical system, we will restrict ourselves to considering a set of identical particles which, at least for the moment, will be considered also *distinguishable*[8].

Let us then consider an isolated system of N such *identical* and *distinguishable* particles and let us further assume that all their allowed states with energy E_i have the *same probability of being occupied*. It is important to remark that both this assumption and the distinguishability requirement will be soon removed, in order to elaborate a more general picture; at this stage they must be intended just as pedagogical tools. Any given partition can be realised in different ways, each corresponding to a specific distribution of the particles among the states. It is quite reasonable to consider that *the probability of each partition is proportional to the number of different ways it can be realised*. The statistical theory we are going to develop is basically funded on this consideration. It is therefore mandatory to correctly count the number of distinguishable different ways we can place $n_i \leqslant N$ particles on the state with energy E_i, once the remaining particles have been already placed.

If we start from the lowest level E_1 such a number is $N!/n_1!(N - n_1)!$, corresponding to the number of permutations of N objects taken n_1 at a time. For the second state with energy E_2 the number of available particles to accomodate is $(N - n_1)$ and, therefore, it is easy to understand that $(N - n_1)!/n_2!(N - n_1 - n_2)!$ is the number of distinguishable different ways we can place n_2 particles on it. By iterating this procedure for all the remaining states, we eventually obtain *the number of different ways a partition can be realised with* n_1 *particles on the state* E_1, n_2 *on the state* E_2, n_3 *on the state* E_3, *and so on*

$$\frac{N!}{n_1!n_2!n_3!\dots} \tag{1.2}$$

while, as stated above, the corresponding probability P for this partition is proportional to it.

We proceed by refining and generalising our reasoning and, first of all, we admit that *the allowed states have now different occupation probabilities which we will indicate by* p_i. A direct way to understand this is to admit the energy E_i is degenerate: states with higher degeneracy are more likely to be occupied. The true existence of different occupation probabilities implies that the number of different ways we can realise the above partition must be corrected as

$$\frac{N!p_i^{n_1}p_2^{n_2}p_3^{n_3}\dots}{n_1!n_2!n_3!\dots} \tag{1.3}$$

---

[8] More specifically, two particles are identical if they have the same structure and physical properties (two H atoms or two NaCl molecules in the same quantum state are so), while they can be distinguished for instance by looking at their position (at least according to classical physics).

since the probability of finding n_i particles on the state E_i is precisely $p_i^{n_i}$. It is very important to remark that *we are not applying any restriction on such intrinsic (or state) probabilities* p_i. In other words, we are not assuming any exclusion principle related to the character of the total wavefunction. This is tantamount to stating that *we are developing a purely classical statistical theory*.

Next, we remove the distinguishability assumption (in other words, particle labelling is any longer relevant): this implies that the $N!$ permutations among particle pairs that occupy the different states actually provide the very same partition. Therefore, *the number* Ω *of different ways the partition can be realised* is eventually corrected in

$$\Omega = \frac{1}{N!}\frac{N! p_i^{n_1} p_2^{n_2} p_3^{n_3} \ldots}{n_1! n_2! n_3! \ldots} = \frac{p_i^{n_1} p_2^{n_2} p_3^{n_3} \ldots}{n_1! n_2! n_3! \ldots} = \prod_i \frac{p_i^{n_i}}{n_i!} \tag{1.4}$$

which corresponds to a probability

$$P = \xi \prod_i \frac{p_i^{n_i}}{n_i!} \tag{1.5}$$

where ξ is a convenient normalisation factor linking the number of different ways the partition can be realised and its corresponding probability $P \in [0.1]$. This factor does not play any role in the theory, as it will be clear very soon.

According to the definition provided in section 1.1, the equilibrium state corresponds to the most probable partition and, therefore, it can be calculated by maximising the quantity provided in equation (1.5). This procedure, however, must fulfil two physical constraints provided in equations (1.1), namely (i) the conservation of energy and (ii) the conservation of the number of particles. In practice, the calculation proceeds by considering at first the natural logarithm[9] of the partition probability P

$$\ln P = \ln \xi + \ln\left(\prod_i \frac{p_i^{n_i}}{n_i!}\right) = \ln \xi + \sum_i \ln \frac{p_i^{n_i}}{n_i!} = \ln \xi + \sum_i (n_i \ln p_i - \ln n_i!) \tag{1.6}$$

which is easily calculated by means of the Stirling formula $\ln n! \simeq n \ln n - n$ valid for large enough n numbers (see appendix A)

$$\ln P = \ln \xi + \sum_i \left(-n_i \ln \frac{n_i}{p_i} + n_i\right) = \ln \xi + N - \sum_i n_i \ln \frac{n_i}{p_i} \tag{1.7}$$

Next, in order to find its maximum we must differentiate with respect to small changes dn_1, dn_2, dn_3,… of the occupation numbers

$$-d \ln P = \sum_i dn_i \ln \frac{n_i}{p_i} = 0 \tag{1.8}$$

---

[9] Since the logarithm is a monotonic function of its variable, the maximum of $\ln P$ corresponds to the maximum of P.

where we have taken into consideration that both ξ and N are constant since, respectively, the first one is just a numerical factor and the system is isolated. This latter condition further imposes that upon the same small changes we must have

$$\sum_i dn_i = 0 = \sum_i E_i \, dn_i \tag{1.9}$$

so that the maximum condition for the quantity appearing in equation (1.8) must be compensated according to the Lagrange method (see appendix A)

$$\sum_i \left(\ln \frac{n_i}{p_i} + \alpha + \beta E_i \right) dn_i = 0 \tag{1.10}$$

where α and β are the Lagrange multipliers. The maximum probability is then found by imposing

$$\ln \frac{n_i}{p_i} + \alpha + \beta E_i = 0 \tag{1.11}$$

which leads to

$$n_i = p_i \exp(-\alpha - \beta E_i) \tag{1.12}$$

the fundamental result paving the way to the statistical mechanics of classical identical particles.

1.3 The Boltzmann distribution law

In order to proceed further we need to determine the actual value of the Lagrange multipliers α and β appearing in equation (1.12). To this aim, let us calculate explicitly the total number of particles as

$$N = \sum_i n_i = \sum_i p_i \exp(-\alpha - \beta E_i) = \exp(-\alpha) \sum_i p_i \exp(-\beta E_i) \tag{1.13}$$

and introduce the *partition function* $\mathcal{Z}$

$$\mathcal{Z} = \sum_i p_i \exp(-\beta E_i) \tag{1.14}$$

which represents a key quantity of statistical physics. By combining equations (1.13) and (1.14) we obtain $\exp(-\alpha) = N/\mathcal{Z}$ and eventually write

$$n_i = \frac{N}{\mathcal{Z}} p_i \exp(-\beta E_i) \tag{1.15}$$

which is known as *the Boltzmann distribution law* describing the equilibrium statistics of an assembly of classical identical particles distributed on energy levels E_i with intrinsic probabilities p_i.

1.4 A statistical definition of temperature

For equation (1.15) to be a useful result for applications we still need to provide our theory with an explicit and computable expression of the β multiplier. To this aim we observe that for dimensional consistency β must be measured in reciprocal energy units; therefore, we guess that working on the total energy of the system could be the right way to proceed.

By inserting equation (1.15) into equation (1.1) we get

$$\mathcal{U} = \frac{N}{\mathcal{Z}} \sum_i p_i E_i \exp(-\beta E_i) = -\frac{N}{\mathcal{Z}} \frac{d}{d\beta} \left[\sum_i p_i \exp(-\beta E_i) \right] \tag{1.16}$$

which allows to express *the energy of a system of particles in terms of its partition function* as

$$\mathcal{U} = -N \frac{d}{d\beta} (\ln \mathcal{Z}) \tag{1.17}$$

and the *average energy per particle* as

$$\langle E \rangle = -\frac{d}{d\beta} (\ln \mathcal{Z}) \tag{1.18}$$

where $\langle E \rangle = \mathcal{U}/N$.

While equations (1.17) and (1.18) clearly confirm our guess that *the parameter β is linked to the system internal energy*, for historical reasons it has been preferred to introduce a new physical quantity T, referred to as *absolute temperature*, defined as

$$k_B T = \frac{1}{\beta} \tag{1.19}$$

where k_B is a suitable constant, discussed below, so that the product $k_B T$ is expressed in energy units. Equation (1.19) represents the *statistical definition of temperature*: it is a positive-definite quantity (the lowest absolute temperature value is therefore zero), whose relation with the phenomenological temperature defined in classical thermodynamics will soon be established.

It is important to remark that the absolute temperature introduced through the above statistical arguments is only defined at equilibrium: the fundamental reason is that the definition of T is based on the β parameter which, in turn, has been elaborated when looking for the most probable distribution of the system particles among the allowed energy state or, more concisely, when looking for the equilibrium condition of the system[10].

---

[10] In non-equilibrium situations is still possible to define a 'local temperature' for each portion of the system which is still large enough to validate a statistical approach, but comparatively much smaller than the full system size. This allows, for example, defining how the temperature varies within a material system in stationary condition, along the direction of an applied thermal gradient: indeed a non-equilibrium situation! While this comment is only qualitative, non-equilibrium thermodynamics makes this definition of temperature quite robust through the concept of 'local equilibrium' [7].

The physical meaning of the statistical temperature defined in equation (1.19) is fully understood by considering a system formed by two interacting subsystems containing $N^{(a)}$ and $N^{(b)}$ particles, respectively, so that $N = N^{(a)} + N^{(b)}$. Different energy states are allowed for the particles in the two subsystems, hereafter labelled as $E_i^{(a)}$ and $E_j^{(b)}$. We will consider a simplified, but meaningful, situation in which (i) the system as a whole is isolated, (ii) the two subsystems exchange energy through particle–particle interactions and collisions, and (iii) no phenomena could occur affecting the particle numbers. In short: the system total energy is conserved, as well as each individual particle number $N^{(a)}$, and $N^{(b)}$; on the other hand, the energy of each subsystem can vary. This situation is described by the following three constraints

$$N^{(a)} = \sum_i n_i^{(a)} \quad N^{(b)} = \sum_j n_j^{(b)} \quad \mathcal{U} = \sum_i n_i^{(a)} E_i^{(a)} + \sum_j n_j^{(b)} E_j^{(b)} \tag{1.20}$$

while, by extending the arguments developed in section 1.2, the probability P to observe a given partition with $n_i^{(a)}$ particles of the first subsystem on the energy levels $E_i^{(a)}$ and $n_j^{(b)}$ particles of the second subsystem on the energy levels $E_j^{(b)}$ is given by

$$P = \xi \left[\prod_i \frac{[p_i^{(a)}]^{n_i^{(a)}}}{n_i^{(a)}!} \right] \left[\prod_j \frac{[p_j^{(b)}]^{n_j^{(b)}}}{n_j^{(b)}!} \right] \tag{1.21}$$

where $p_i^{(a)}$ and $p_j^{(b)}$ are the occupation probabilities of the various energy states in the two subsystems, respectively. This equation is the straightforward extension of equation (1.5). In order to find the equilibrium distribution for the system we will follow the same procedure developed in section 1.2.

By differentiating the natural logarithm of the probability given in equation (1.21) with respect to small changes $dn_i^{(a)}$ and $dn_j^{(b)}$ in the occupation numbers within the two subsystems, we get

$$-d(\ln P) = \sum_i \ln \frac{n_i^{(a)}}{p_i^{(a)}}\, dn_i^{(a)} + \sum_j \ln \frac{n_j^{(a)}}{p_j^{(b)}}\, dn_j^{(b)} = 0 \tag{1.22}$$

where we set once again to zero such a variation since we are looking for the maximum probability distribution. The conservation conditions are now dictated as

$$\sum_i dn_i^{(a)} = 0 \quad \sum_j dn_j^{(b)} = 0 \quad \sum_i E_i^{(a)} dn_i^{(a)} + \sum_j E_j^{(b)} dn_j^{(b)} = 0 \tag{1.23}$$

leading to

$$\sum_i \left(\ln \frac{n_i^{(a)}}{p_i^{(a)}} + \alpha^{(a)} + \beta E_i^{(a)} \right) dn_i^{(a)} + \sum_j \left(\ln \frac{n_j^{(b)}}{p_j^{(b)}} + \alpha^{(b)} + \beta E_j^{(b)} \right) dn_j^{(b)} = 0 \tag{1.24}$$

once the three multipliers $\alpha^{(a)}$, $\alpha^{(b)}$, and β have been introduced. The above equation is satisfied provided that

$$\ln \frac{n_i^{(a)}}{p_i^{(a)}} + \alpha^{(a)} + \beta E_i^{(a)} = 0 = \ln \frac{n_j^{(b)}}{p_j^{(b)}} + \alpha^{(b)} + \beta E_j^{(b)} \tag{1.25}$$

or equivalently

$$n_i^{(a)} = \frac{N^{(a)}}{\mathcal{Z}^{(a)}} p_i^{(a)} \exp[-E_i^{(a)} \beta E_i^{(a)}] \quad \text{and} \quad n_j^{(b)} = \frac{N^{(b)}}{\mathcal{Z}^{(b)}} p_j^{(b)} \exp\left[-E_j^{(b)} \beta E_j^{(b)}\right] \tag{1.26}$$

where $\mathcal{Z}^{(a)}$ and $\mathcal{Z}^{(b)}$ are the respective partition functions of the two subsystems

$$\mathcal{Z}^{(a)} = \sum_i p_i^{(a)} \exp[-\beta E_i^{(a)}] \quad \text{and} \quad \mathcal{Z}^{(b)} = \sum_j p_j^{(b)} \exp\left[-\beta E_j^{(b)}\right] \tag{1.27}$$

which, interestingly enough, result in *having the same value of the parameter* β. In other words, thanks to this result and to the definition provided in equation (1.19) we conclude that *in equilibrium the two subsystems have the very same absolute temperature*. This is perfectly consistent with the *zeroth law of classical thermodynamics* [3, 8–10] and, therefore, makes the statistical definition of temperature fully compatible with its phenomenological counterpart. According to the present argument, it is understood that in equilibrium conditions the two subsystems may exchange energy at the microscopic level (occupation numbers in both subsystems do vary around their equilibrium values because of mutual interactions); however, such an exchange occurs in both directions so that statistically their energy remains constant. On the other hand, if the two subsystems are initially at different temperature, interactions cause energy exchanges until the same-temperature equilibrium condition is eventually reached.

The zeroth law has an important practical consequence. Its equivalent verification at the statistical and empirical level implies that *the statistical absolute temperature of a system is precisely the experimentally detected one*, since this latter is determined by placing in contact the system with a thermometer and waiting until equilibrium is reached. We recall that a thermometer is a device characterised by a thermometric parameter, that is, by a physical temperature-dependent property like, for instance, its volume or electrical resistance [9]. Different values of such a property are collected at corresponding conventional situations (for instance, at freezing and boiling conditions of water at normal pressure) and a suitable temperature scale is accordingly defined. Provided that absolute temperature is measured in degrees Kelvin, the conventional value of the constant k_B introduced in equation (1.19) is

$$k_\mathrm{B} = 1.3805 \times 10^{-5}\ \mathrm{eV\ K^{-1}} \tag{1.28}$$

and it is commonly referred to as the *Boltzmann constant*.

1.5 Calculating ensemble averages

Let us consider a system containing N particles and an observable O which can be expressed as a function of the particle energies E_i. Then, for any given partition its *average value* $\langle O \rangle$ is by definition calculated as

$$\langle O \rangle = \frac{1}{N} \sum_i n_i O(E_i) \tag{1.29}$$

which, by using the equilibrium distribution probability provided in equation (1.15) and the temperature definition provided in equation (1.19), is transformed into

$$\langle O \rangle = \frac{1}{\mathcal{Z}} \sum_i p_i \; O(E_i) \exp(-E_i/k_\mathrm{B}T) \tag{1.30}$$

corresponding to *the operational definition of ensemble average.*

A very special observable is of course the internal energy $\mathcal{U}$ defined in equation (1.17). Since equation (1.19) holds, we have

$$d\beta = -\frac{dT}{k_\mathrm{B}T^2} \tag{1.31}$$

and we obtain the *link between the partition function and the internal energy* as

$$\mathcal{U} = k_\mathrm{B}NT^2 \frac{d}{dT} \ln \mathcal{Z} \tag{1.32}$$

while the average energy per particle is

$$\langle E \rangle = k_\mathrm{B}T^2 \frac{d}{dT} \ln \mathcal{Z} = \frac{1}{\mathcal{Z}} \sum_i p_i E_i \exp(-E_i/K_\mathrm{B}T) \tag{1.33}$$

where we made use of equation (1.18). This result, fully consistent with the above definition of ensemble average, is of great conceptual relevance since *it establishes a direct link between the system temperature and the average energy per particle.* Using a more phenomenological language, by virtue of equation (1.33) we can say that the sensorial experience of 'hot' and 'cold' is nothing other than an effective (macroscopic) measure of the energy content of the system we entered in contact with.

We conclude this section by remarking that any explicit calculation of ensemble averages does require the knowledge of the partition function, which depends on the microscopic structure of the system as well as on the thermodynamical conditions the system is subjected to. Basically, the statistical description of a physical system is a two-step procedure: the partition function must be at first calculated; next, any observable can be straightforwardly evaluated according to equation (1.30). More specifically, different partition functions must be in principle considered for an isolated system or a system in contact with a thermal bath or a system which can vary its matter content. Averages calculated in the first case (where the system is characterised by a fixed energy and number of particles) correspond to the *microcanonical ensemble*, while in the second case (where the system can exchange energy, while its particles are still fixed in number) they define the *canonical ensemble*. In both the microcanonical and canonical case, the partition function $\mathcal{Z}$ is given by equation (1.14). In contrast, if the system can exchange energy and mass (that is, the number of particles is no longer fixed) the corresponding statistics is performed in the *grand canonical ensemble* for which a different form of the partition function must be elaborated. This topic falls beyond the scope of this Primer[11], as well as the details of the ensemble theory which can be found elsewhere [3, 10, 11]. The key point to retain for the following is that *the fundamental task of statistical physics is to evaluate the partition function of any specific system in assigned thermodynamical conditions*. Only after this mission has been accomplished, predictions about the (average) system properties can be elaborated.

1.6 Entropy

1.6.1 Thermodynamical entropy

In thermodynamics a function named *entropy* S is defined for any system in whatever equilibrium as well as non-equilibrium state, such that its variation when the system evolves depends only on the initial and final states occupied by the system and not on the specific process followed. Being a state function, S dictates the thermal physics of the system: any thermodynamical parameter can be calculated from it [8, 9, 11]. For instance, temperature T, pressure P and chemical potential μ_c, namely the three mostly relevant *intensive quantitites*, are directly calculated once the entropy of the system is known

$$\frac{1}{T} = \left.\frac{\partial S}{\partial \mathcal{U}}\right|_{V,N} \qquad \frac{P}{T} = \left.\frac{\partial S}{\partial V}\right|_{\mathcal{U},N} \qquad -\frac{\mu_c}{T} = \left.\frac{\partial S}{\partial N}\right|_{\mathcal{U},V} \tag{1.34}$$

where for sake of simplicity we have consider a mono-component system[12] whose entropy depends on just three parameters $S = S(\mathcal{U}, V, N)$. For an infinitesimal process we can therefore write

$$dS = \frac{1}{T}d\mathcal{U} + \frac{P}{T}dV - \frac{\mu_c}{T}dN \tag{1.35}$$

from which we immediately obtain the *first law of thermodynamics*: the variation of internal energy of the system

$$d\mathcal{U} = TdS - PdV + \mu_c dN = d\mathcal{Q} - dW + \mu_c dN \tag{1.36}$$

---

[11] The only exception is treated in section 3.4.

[12] The extension to multi-component systems is straightforward, but since it does not add any new physical feature it will not be considered here.

is given by the balance between the amount of heat $dQ = TdS$ exchanged with the environment[13], the mechanical work $dW = P\,dV$ caused by volume variations, and the chemical work $\mu_c\,dN$ due to the variation in the particle number of the system.

1.6.2 Boltzmann entropy

Entropy being such a fundamental object of thermodynamics, it deserves a *statistical definition*, which was in fact guessed by Boltzmann in the form

$$S = k_B \ln \Omega \tag{1.37}$$

where Ω is the number of microstates compatible with the selected macrostate of the system (sometimes Ω is also referred to as the 'multiplicity function'). It must be clear that *the definition provided in equation* (1.37) *is just a postulate* [3, 5], whose validity can only be proved heuristically, as thoroughly discussed below. Nevertheless, it is supported by robust and meaningful arguments. First of all, we observe that when a system evolves towards equilibrium Ω is maximised (see discussion in section 1.1): therefore, according to equation (1.37) *S is maximum for the equilibrium state*, consistently with thermodynamics [8, 9]. Next, we recall that statistical laws are multiplicative and, therefore, if a system is composed by two non-interacting subsystems, respectively, characterised by Ω_1 and Ω_2 number of microstates, it has $\Omega_1\Omega_2$ partitions. This implies that the entropy of the total system is $S = k_B \ln(\Omega_1\Omega_2) = k_B \ln \Omega_1 + k_B \ln \Omega_2 = S_1 + S_2$. In other words, equation (1.37) states that *S is an extensive quantity*, one again as dictated by thermodynamics [8, 9]. A more formal argument leading to the same definition as in equation (1.37) is outlined in section 1.6.3, where the concept of entropy is developed as *a measure of our ignorance on the actual microstate occupied by the system* once its macrostate has been assigned.

Let us now consider an isolated system out of equilibrium and follow its natural evolution: *we always observe a positive entropy variation* (simply because the free evolution drives the system to the macrostate with maximum number of realisations). On the other hand, if the isolated system was already at equilibrium, the only observed spontaneous processes are those characterised by a zero entropy variation. These arguments are tantamount to stating that *in isolated systems we always have* $dS \geqslant 0$, corresponding to the second law of thermodynamics[14]; $dS = 0$ processes are called *reversible*[15], while $dS > 0$ ones are said to be *irreversible*. The statistical nature

---

[13] We remark that $dQ = T\,dS$ holds only for quasi-static processes, that is for processes occurring through a sequence of equilibrium states, where the absolute temperature T can be defined. In practice, they correspond to evolutions proceeding by infinitesimal steps occurring over a time scale much longer than the time needed for the system to reach a new equilibrium state (also referred to as relaxation time). The relation $dQ = T\,dS$ was introduced for the first time by Clausius, long before the microscopic approach to thermal physics was elaborated; it represents the macroscopic operational definition for the entropy variation.

[14] If the system is not isolated, it is indeed possible to observe a negative entropy variation; however, $dS \geqslant 0$ for the larger complex formed by the system and its embedding environment.

[15] The notion of reversible process and of quasi-static process are here considered equivalent.

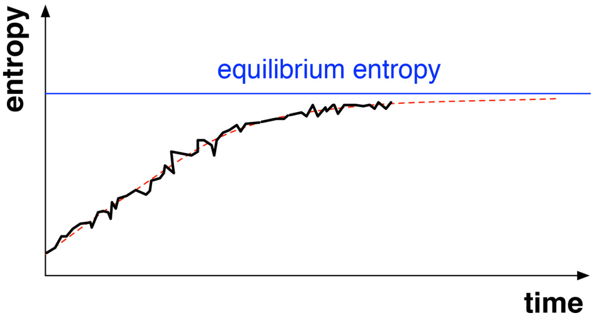


Figure 1.2. Entropy variation for an isolated system evolving towards equilibrium. Dashed line: a guide to the eye, corresponding to a macroscopic thermodynamical prediction. Noisy full line: actual entropy variation according to equation (1.37).

of the definition reported in equation (1.37) allows us to admit that, because of fluctuations, the entropy of an isolated system may occasionally decrease; however, the likelihood of such negative variations is inversely proportional to their amplitude. In figure 1.2 a pictorial representation of the entropy variation towards equilibrium for an isolated system is sketched.

Let us now proceed with the explicit calculation of the entropy for a classical system obeying the Boltzmann statistics. By inserting equation (1.4) into equation (1.37) and using the Stirling formula (see appendix A) we get

$$S = k_{\mathrm{B}}\left(\sum_i n_i \ln p_i - \sum_i n_i \ln n_i + \sum_i n_i\right) = -k_{\mathrm{B}}\sum_i n_i \ln \frac{n_i}{p_i} + k_{\mathrm{B}}N \tag{1.38}$$

where we used equation (1.1) for the total number of particles. The logarithm appearing in the sum on the right hand side of this equation is easily calculated by using equation (1.15)

$$S = k_{\mathrm{B}}\sum_i n_i \frac{E_i}{k_{\mathrm{B}}T} + k_{\mathrm{B}}\sum_i n_i \ln \frac{\mathcal{Z}}{N} + k_{\mathrm{B}}N = \frac{1}{T}\sum_i n_i E_i + k_{\mathrm{B}} \ln \frac{\mathcal{Z}}{N}\sum_i n_i + k_{\mathrm{B}}N \tag{1.39}$$

where it is easy to recognise the total internal energy of the system and the total number of particles in the two sums appearing, respectively, in the first and second term of the right hand side of this equation. By means of equations (1.1) we therefore obtain

$$S = \frac{\mathcal{U}}{T} + k_{\mathrm{B}}N \ln \frac{\mathcal{Z}}{N} + k_{\mathrm{B}}N \tag{1.40}$$

which eventually provides

$$S = \frac{\mathcal{U}}{T} + k_{\mathrm{B}} \ln \frac{\mathcal{Z}^N}{N!} \tag{1.41}$$

as the *entropy for a system of identical and indistinguishable particles following the Boltzmann statistics.*

The Boltzmann formulation for the entropy can justify the Clausius definition $dQ = TdS$ linking the amount of exchanged heat to the entropy variation observed during a reversible process. This result is so relevant that it deserves a proof, which is interesting either as a showcase application of the statistical formalism and for better establishing its link to thermodynamics. If we consider a reversible transformation of an isolated mono-component system obeying the Boltzmann statistics and use equation (1.40), we can write

$$dS = \frac{1}{T}d\mathcal{U} - \frac{\mathcal{U}}{T^2}dT + \frac{k_B N}{\mathcal{Z}}d\mathcal{Z} \tag{1.42}$$

where, since the system is isolated, we set $dN = 0$. We now focus on the last term appearing on the right hand side of this equation and we calculate at first

$$d\mathcal{Z} = -\sum_i \frac{p_i}{k_B T}\exp(-E_i/k_B T)dE_i + \sum_i \frac{p_i E_i}{k_B T^2}\exp(-E_i/k_B T)dT \tag{1.43}$$

by making use of the definition of partition function given in equation (1.14). Then we immediately obtain

$$\begin{aligned} \frac{k_B N}{\mathcal{Z}}d\mathcal{Z} &= -\frac{1}{T}\sum_i \underbrace{\frac{N}{\mathcal{Z}}p_i \exp(-E_i/k_B T)}_{=n_i} dE_i + \frac{1}{T^2}\sum_i \underbrace{\frac{N}{\mathcal{Z}}p_i \exp(-E_i/k_B T)}_{=n_i} E_i\, dT \\ &= -\frac{1}{T}\sum_i n_i dE_i + \frac{1}{T^2}\left(\sum_i n_i E_i\right)dT \end{aligned} \tag{1.44}$$

where we recalled the Boltzmann distribution law given in equation (1.15).

In order to proceed, let us consider the most general expression for the variation of the total internal energy of the system occurring during the transformation. From equation (1.1) it is easily calculated

$$d\mathcal{U} = \sum_i n_i dE_i + \sum_{i=1}^{N} E_i dn_i \tag{1.45}$$

which represents the microscopic counterpart of the first law of thermodynamics stated in equation (1.36) and here referred to the case in which the total number of particles is not varied since, as commented above, the system is isolated (no contribution from chemical work). Consistently with the macroscopic version of this energy balance principle, we can attach to the two contributions a rather different physical origin: they, respectively, describe how the internal energy is affected (i) as the energies of single-particle states vary by dE_i and (ii) as their partition is modified by a particle redistribution dn_i. We conclude that $dW = -\sum_i n_i\, dE_i$ is the work done by the system upon a volume variation; the negative sign indicates that such a work is performed at the expense of the internal energy of the system, as sketched in figure 1.3 in the simple case of a one-dimensional assembly of free particles confined within a potential box whose width is enlarged. On the other hand, it is natural to identify $dQ = \sum_i E_i dn_i$ as the exchanged heat.

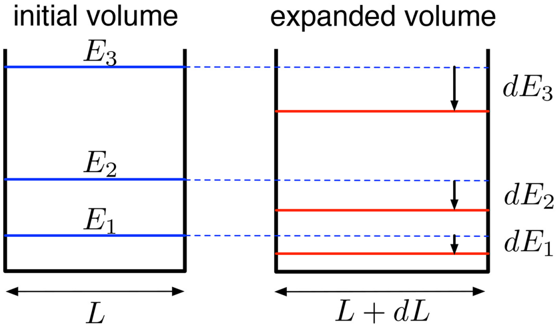


Figure 1.3. Pictorial representation of the change in the energy levels dE_i upon volume variation.

Using these results, equation (1.44) is finally recast in the following form

$$\frac{k_{\mathrm{B}}N}{\mathcal{Z}}d\mathcal{Z} = \frac{1}{T}\,dW + \frac{\mathcal{U}}{T^2}\,dT \tag{1.46}$$

which, if inserted into equation (1.42), immediately leads to the result we were looking for

$$dS = \frac{1}{T}(d\mathcal{U} + dW) = \frac{1}{T}d\mathcal{Q} \tag{1.47}$$

in virtue of the first law stated in equation (1.36).

In conclusion, this journey back-and-forth through thermodynamics and statistical physics provides clear evidence of how much it pays to develop an atomistic description of physical systems: phenomenological definitions introduced at the macroscopic level end up finding a satisfactory microscopic justification based on first principles.

1.6.3 Entropy and ignorance

As extensively discussed above, once a macrostate is assigned, the system can occupy a (possibly very large) variety of microstates. Physically, this reflects in the fact that *a macroscopic measurement cannot determine the actual microstate*. Let us consider a model system whose N particles can occupy just two discrete energy levels E_1 and E_2. As sketched in figure 1.4 the macrostate with energy $\mathcal{U} = NE_1$ has only one microstate since it corresponds to the sole partition with all particles on the ground state with energy E_1 (top panel of figure 1.4); on the other hand, if we consider the macrostate with energy $\mathcal{U} = (N-1)E_1 + E_2$ we have N different choices to select the particle to place on the higher-energy level, corresponding to as many different partitions (middle panels of figure 1.4); next, the macrostate with energy $\mathcal{U} = (N-2)E_1 + 2E_2$ corresponds to $N(N-1)/2$ different microstates, corresponding to the number of pairs we can select out of the set of N particles (bottom panels of figure 1.4); and so on. We immediately realise that *the full knowledge of the macrostate corresponds to different levels of ignorance about its compatible microstates*: in the first case we have no ignorance (just one microstate is possible), in the

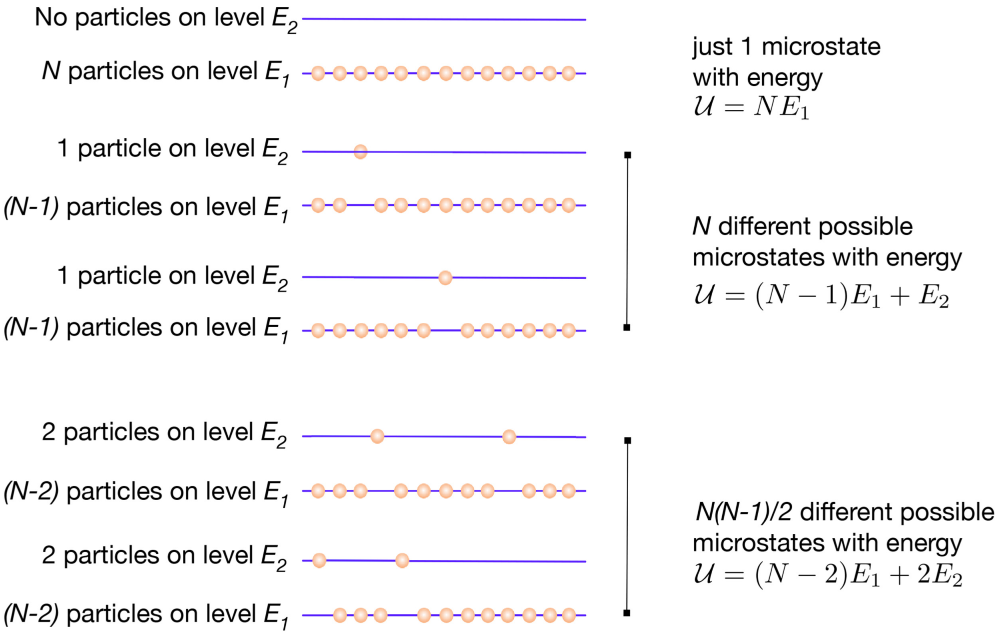


Figure 1.4. Pictorial representation of the microstate with $\mathcal{U} = NE_1$ (top) and some microstates with $\mathcal{U} = (N - 1)E_1 + E_2$ (middle) and $\mathcal{U} = (N - 2)E_1 + 2E_2$ (bottom).

second case we have some ignorance, in the third case we suffer an even worse ignorance (by constitutive hypothesis N is large and, therefore, $N(N - 1)/2 > N$), and so on. This argument suggests that we can try to link our ignorance about the system microstate to its entropy.

Let us start with the assumption that *all microstates corresponding to the same macrostate are equally likely*[16]. This assumption is physically quite reasonable: there is no reason to expect microstates with equal macroscopic properties to have different probability of occurrence. While in appendix B this concept will be critically readdressed, leading to an alternative definition of entropy due to Gibbs, here we will rely on such an assumption. Next, let us introduce a function quantifying our ignorance of the microstate, as defined above. For reasons to be clear in a while, we will name this function S; since all accessible microstates are equally likely, S can only be a function of their number Ω; formally, $S = S(\Omega)$. This function has some intriguing properties, namely: (i) whenever we have no ignorance, that is whenever $\Omega = 1$, it must be zero, or equivalently, we can set $S(\Omega = 1) = 0$; (ii) by considering two macrostates a and b with $\Omega_a > \Omega_b$ we set $S(\Omega_a) > S(\Omega_b)$ since we suffer more ignorance for the first system; (iii) finally, if we consider two independent systems Σ_1 and Σ_2, respectively, with a number of possible microstates Ω_{Σ_1} and Ω_{Σ_2}, then the total ignorance of the joint system $\Sigma_1 + \Sigma_2$ is the sum of the ignorances we have about its separated components and formally we write

---

[16] We remark that microstates corresponding to *different* macrostates are not constrained to such a same-likelihood condition.

$S(\Omega_{\Sigma_1+\Sigma_2}) = S(\Omega_{\Sigma_1}\Omega_{\Sigma_2}) = S(\Omega_{\Sigma_1}) + S(\Omega_{\Sigma_2})$. In summary, this ignorance function S must obey the following properties

$$\begin{aligned} S(1) &= 0 \\ S(xy) &= S(x) + S(y) \\ \text{if } x > y \rightarrow S(x) &> S(y) \end{aligned} \tag{1.48}$$

which are fulfilled by the function logarithm. Therefore, we straightforwardly set

$$S(x) = k \ln x + \text{constant} \tag{1.49}$$

that, since $S(1) = 0$, further reduces to $S(x) = k \ln x$. If we now set $k = k_B$ we eventually obtain equation (1.37). This formal reasoning completes the plausibility argument developed in the previous section to define entropy by means of equation (1.6.2) which is now linked to the content of information (or, reversely, to the degree of ignorance) we have about the system.

The conceptual link we just established between the entropy and the ignorance we have about the system allows for an interesting excursion in the realm of information theory. Here the key ingredients are not physical particles, but rather 'information quanta' (usually referred to as *bits*), which are restricted to just two values, namely 0 and 1. If we have N bits, we have 2^N possible information states which correspond to a physical entropy $S_{\text{init}} = Nk_B \ln 2$. If we want to erase the information stored in a computational device, we should set all bits to the same value (it is irrelevant whether 0 or 1): a situation which corresponds to $S_{\text{final}} = 0$. Erasing the stored information requires therefore an entropy variation $\Delta S = S_{\text{final}} - S_{\text{init}} = -Nk_B \ln 2$. As discussed in section 1.6.2, the second law of thermodynamics imposes that the entropy change must always be non-negative and, therefore, we conclude that we need to increase the environment entropy by an amount $k_B \ln 2$ (sometimes referred to as Shannon entropy) for any bit erased in the device memory. This corresponds to a minimum $k_B T \ln 2$ dissipated heat per erased bit. Deleting a file on our personal computer does increase the environment temperature!

1.7 The ideal monoatomic gas

The ideal monoatomic gas is a simple model system made by *identical, structureless, non-interacting particles*. This implies that the energy of each particle is just translational kinetic, since no interactions among them are present nor must internal roto-vibrational contributions be taken into account.

Although we are addressing an idealised physical situation, the results obtained for the ideal monoatomic gas are largely valid for real dilute gases (kept at intermediate or high temperature). More importantly, this model represents a benchmark system where calculations can be exploited analytically and the results can be directly compared with similar ones obtained by other theoretical approaches, like for instance the kinetic theory of gases [3].

In short, the physics of the ideal monoatomic gas provides paradigmatic results which exceed the limits of the model for relevance and general validity.

1.7.1 Partition function

Let us consider an assembly of N independent (that is, free) particles confined in a volume V. It is a well known fundamental result of quantum mechanics that the energy spectrum of a confined particle is discrete [12–14]; however, the spacing between two adjacent levels is inversely proportional to $V^{2/3}$. This means that for a large enough volume, successive levels are so close that their spectrum can be treated as continuous to a very good approximation. This is precisely the case we are treating: a gas of pointlike particles contained within a macroscopic volume.

The fact that the energy spectrum is continuous has a twofold implication. First of all, we can profitably make use of the concept of *density of energy levels* $G(E)$ for the confined particles, where it is understood that $G(E)\,dE$ provides the number of levels in the energy interval $[E, E + dE]$; a standard quantum mechanical calculation [12–14] proves that

$$G(E) = \frac{4\pi V}{h^3}(2m^3E)^{1/2} \tag{1.50}$$

where $h = 4.135\,7 \times 10^{-15}$ eV s is the Planck constant and m is the particle mass. Next, we proceed by replacing the partition function given in equation (1.14) with its continuous counterpart

$$\mathcal{Z} = \int_0^{+\infty} \exp(-E/k_\mathrm{B}T)G(E)dE \tag{1.51}$$

where in setting the limits of the energy integral we have duly taken into account that in the present case E is just kinetic energy and, therefore, is a positive quantity. By replacing equation (1.50) in equation (1.51) we easily get the *partition function of the ideal monoatomic gas*

$$\mathcal{Z} = \frac{V}{h^3}(2\pi m k_\mathrm{B}T)^{3/2} \tag{1.52}$$

where for the calculation we used appendix A reporting some integrals useful in statistical physics. This result shows that *the partition function of the gas depends both on its volume and temperature.*

1.7.2 The equation of state

In applying the general expression for the variation of the partition function given in equation (1.46), we must take into account that for an ideal gas the only form of mechanical work is related to volume variations or, more formally, $dW = P\,dV$ where P is the system pressure. This allows us to recast equation (1.46) in the form

$$\frac{k_\mathrm{B}N}{\mathcal{Z}}d\mathcal{Z} = \frac{P}{T}dV + \frac{\mathcal{U}}{T^2}dT \tag{1.53}$$

which, in the case of a constant-temperature process, allows us to find the *statistical definition of pressure*

$$P = k_B NT \frac{\partial \ln \mathcal{Z}}{\partial V}\bigg|_T \tag{1.54}$$

in the form of a relationship between temperature, pressure, and partition function. By inserting the explicit form of $\mathcal{Z}$ provided in equation (1.52) we immediately find

$$PV = k_B NT = nRT \tag{1.55}$$

a result known as *equation of state for the ideal monoatomic gas*; as is customary, the number of particles has been set $N = n\mathcal{N}_A$ where $\mathcal{N}_A = 6.022 \times 10^{23}$ mol^{-1} is the Avogadro number, corresponding to the number of particles in one mole of gas.

Through equation (1.55) we can define the experimental value of the Boltzmann constant anticipated in equation (1.28). Following the standard convention, we define as *the normal condition* the equilibrium of a mixture of liquid water and solid water (ice) at the standard pressure of 1 atm. The corresponding temperature is assigned a conventional value T_0 by choosing a convenient scale. If we now bring a gas in contact with the mixture and wait until it reaches the equilibrium, we can measure its pressure P_0 and volume V_0. The Boltzmann constant is accordingly defined as $k_B = P_0V_0/NT_0$ and its numerical value is straightforwardly obtained once we know the mass of each atom forming the gas.

1.7.3 Energy and heat capacity

In order to calculate the total internal energy $\mathcal{U}$ of an ideal gas in equilibrium condition we make use of equation (1.33) where the partition function is provided by equation (1.52)

$$\langle E \rangle = k_B T^2 \frac{d}{dT} \ln \mathcal{Z} = \frac{3}{2} k_B T \tag{1.56}$$

from which we immediately obtain

$$\mathcal{U} = \frac{3}{2} N k_B T \tag{1.57}$$

a result which can be alternatively derived within the kinetic theory [3]. A different useful formulation is obtained by considering n moles of ideal gas in equilibrium at temperature T: equation (1.57) easily leads to

$$\mathcal{U} = \frac{3}{2} nRT \tag{1.58}$$

where $R = k_B \mathcal{N}_A = 8.313\,4$ J K^{-1} mol^{-1} is the universal gas constant. This very popular result, consistent with classical macroscopic thermodynamics [8–10], states that *the internal energy on an ideal gas in equilibrium depends only on its temperature*. From this result we obtain immediately the *molar constant-volume heat capacity* $\mathcal{C}_V$ of the ideal monoatomic gas.

$$\mathcal{C}_V = \frac{1}{n} \frac{\partial \mathcal{U}}{\partial T}\bigg|_V = \frac{3}{2} R \tag{1.59}$$

The constant-pressure counterpart of this quantity is calculated from the system enthalpy $\mathcal{H} = \mathcal{U} + PV$, which as outlined in appendix C is a thermodynamical potential representing the work available when the system is kept in a P = constant condition. For the monoatomic ideal gas we have

$$\mathcal{H} = \frac{3}{2}nRT + PV = \frac{3}{2}nRT + nRT = \frac{5}{2}nRT \tag{1.60}$$

and therefore we can define the *molar constant-pressure heat capacity* $\mathcal{C}_\mathrm{P}$ as

$$\mathcal{C}_\mathrm{P} = \frac{1}{n}\frac{\partial \mathcal{H}}{\partial T}\bigg|_\mathrm{P} = \frac{5}{2}R \tag{1.61}$$

a result immediately leading to

$$\mathcal{C}_\mathrm{P} - \mathcal{C}_\mathrm{V} = R \tag{1.62}$$

which is known as the *Meyer relation for the ideal gas*. This equation states that the amount of heat required to increase the temperature by one degree Kelvin of an ideal gas at constant pressure is greater than that required to achieve the same increase at constant volume. The physical reason is that the heat supplied to the system in the first case is divided between temperature increase and mechanical work. More precisely, the work per mole produced at constant pressure is

$$W = \frac{1}{n}\int PdV = \frac{1}{n}\int_T^{T+1} nRdT = R \tag{1.63}$$

for a $\Delta T = 1$ K temperature increase.

1.7.4 The Maxwell distribution law

By means of equation (1.56), we can look at the superior Boltzmann theory as the statistical foundation of the kinetic theory of gases[17]. In order to further develop this concept, let us consider the Boltzmann distribution law given in equation (1.15) and replace the occupation probabilities p_i with their continuous counterpart $G(E)dE$; the resulting *number dN of particles with energy in the range* $[E,\ E + dE]$ is given by

$$dN = \frac{N}{\mathcal{Z}}\exp(-E/k_\mathrm{B}T)G(E)dE = \frac{N}{\mathcal{Z}}\frac{4\pi V}{h^3}(2m^3)^{1/2}E^{1/2}\exp(-E/k_\mathrm{B}T)dE \tag{1.64}$$

which, by using equation (1.52) for the equilibrium partition function, immediately leads to

---

[17] Incidentally, we remark that this result suggested the identification anticipated in equation (1.19). The kinetic theory (developed prior to statistical mechanics) already proved that $\langle E \rangle = 3k_\mathrm{B}T/2$ on the basis of purely mechanical arguments; on the other hand, the Boltzmann statistics without equation (1.19) would lead to the formulation $\langle E \rangle = 3/2\beta$. It is then straightforward to identify the Lagrange multiplier as $\beta = 1/k_\mathrm{B}T$.

$$\frac{dN}{dE} = \frac{2\pi N}{(\pi k_B T)^{3/2}} E^{1/2} \exp(-E/k_B T) \tag{1.65}$$

a result known as the *Maxwell energy distribution law* in an ideal gas. Since in the case under study the particle energy is purely kinetic, the energy only depends on the velocity v and we have

$$\frac{dN}{dv} = \frac{dN}{dE}\frac{dE}{dv} = \frac{dN}{dE}mv \tag{1.66}$$

where we made use of equation (1.65) by setting $E = mv^2/2$. By combining these intermediate results we obtain

$$\frac{dN}{dv} = 4\pi N\left(\frac{m}{2\pi k_B T}\right)^{3/2} v^2 \exp(-mv^2/2k_B T) \tag{1.67}$$

which represents the *Maxwell velocity distribution law* in an ideal gas, providing the number of particles with velocity in the range $[v,\ v + dv]$ (irrespective of the direction of motion) when the system is in equilibrium at temperature T.

The two Maxwell laws are pretty well confirmed by direct experimental results. Let us consider the experimental setup sketched in figure 1.5: the particles of a monoatomic gas kept at controlled temperature by a thermostat are effused through a nozzle placed on the surface of the containing box and collimated to form a thin beam. The beam is directed towards two slotted disks which rotate around their common axis at constant angular frequency ω; their slots are displaced by an angle θ, while the centres of the two disks are placed at distance s. The particles can pass through the two-disk chopper only if they propagate with translational velocity $v = s\omega/\theta$, since only in this case can the beam pass through both slots, one after the other; in this condition, they are collected by a detector. Therefore, by tuning ω or θ

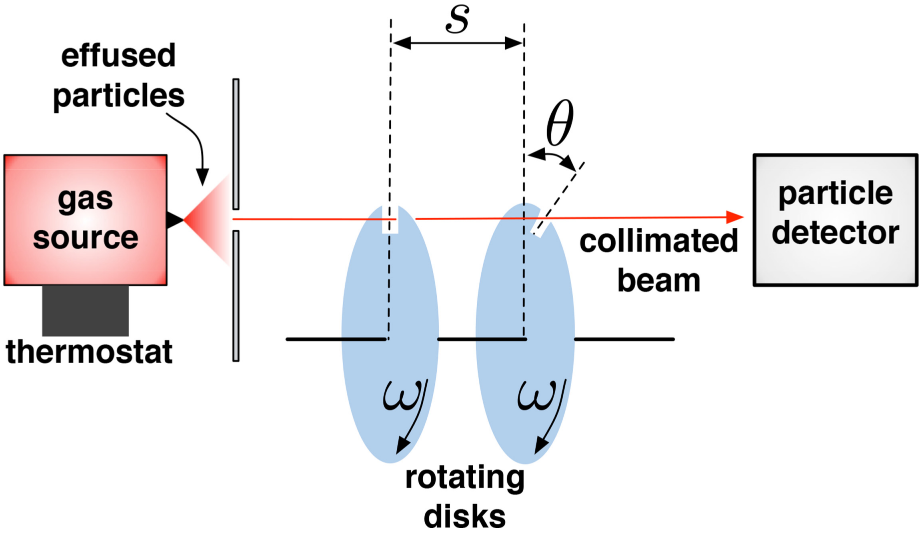


Figure 1.5. Experimental setup to determine the velocity distribution within a gas.

we can effectively select the velocity of the transmitted particles and count their number. The experiment consists in fixing the thermostat temperature and counting the number of transmitted particles at several different velocities. This provides the velocity distribution at that temperature, just in perfect agreement with equation (1.67). By now changing the gas temperature and repeating the measure, we can experimentally determine how the velocity distribution depends on T; once again, empirical data confirm the theoretical prediction by the Maxwell law[18].

The Maxwell laws are not monotonic: thus, it is meaningful to calculate *the most probable energy* E_{mp} *and velocity* v_{mp} for the particles in the ideal monoatomic gas. They are obtained by looking at the maximum of the distributions provided by equations (1.65) and (1.67)

$$E_{\mathrm{mp}} = \frac{1}{2}k_{\mathrm{B}}T \quad \text{and} \quad v_{\mathrm{mp}} = \sqrt{\frac{2k_{\mathrm{B}}T}{m}} \tag{1.68}$$

and turn out to only depend on temperature. Similarly, we can calculate *the average velocity* $\langle v \rangle$ of the particles. We simply apply the definition of average

$$\langle v \rangle = \frac{1}{N}\int_0^{+\infty} v\frac{dN}{dv}dv = 4\pi N\left(\frac{m}{2\pi k_{\mathrm{B}}T}\right)^{3/2}\int_0^{+\infty} v^3 \exp(-mv^2/2k_{\mathrm{B}}T)dv \tag{1.69}$$

where we used equation (1.67) for the term dN/dv. By integration by parts we obtain

$$\langle v \rangle = \sqrt{\frac{8k_{\mathrm{B}}T}{\pi m}} \tag{1.70}$$

proving that the average velocity is always larger that the most probable one. We can also evaluate *the mean square velocity* $\langle v^2 \rangle$ of the particles by a similar averaging procedure

$$\langle v^2 \rangle = \frac{1}{N}\int_0^{+\infty} v^2\frac{dN}{dv}dv \tag{1.71}$$

but in this case we can transform the integral as

$$\langle v^2 \rangle = \frac{2}{m}\frac{1}{N}\int_0^{+\infty} EdN = \frac{2}{m}\langle E \rangle \tag{1.72}$$

since $v^2 = 2E/m$. We eventually get

---

[18] A more indirect way to verify equation (1.65) is to consider a chemical reaction which can only occur by involving particles with energy above a known value E_{react}. Accordingly, the rate of this chemical reaction at a given temperature depends on the number of particles whose energy exceeds E_{react}. Now, we can measure the reaction rate at different temperatures and we can estimate by equation (1.65) the number of active particles at the same temperatures with energy higher than E_{react}. The key point is that if the observed rates correlated with the estimated particles number at each temperature, we could definitely validate the Maxwell law for the energy distribution. Good enough, this is indeed the case: data correlate and, therefore, the empirical evidence does support statistical theory predictions.

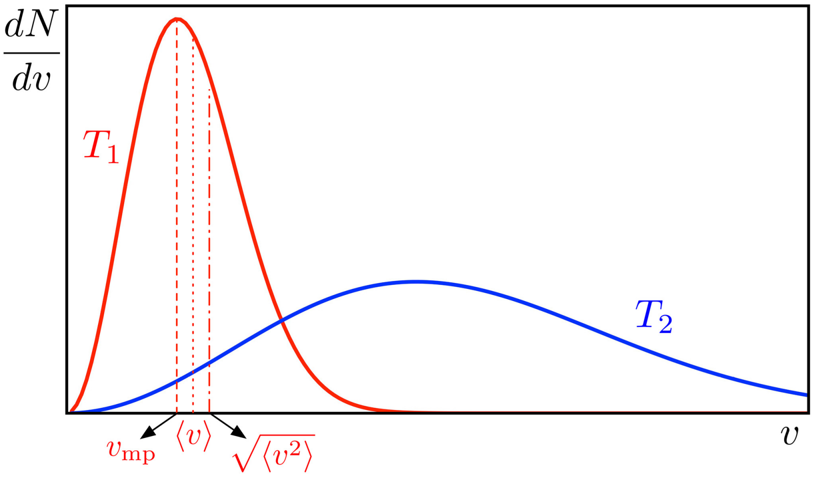


Figure 1.6. Maxwell law for the velocity distribution (see equation (1.67)) calculated for the same system at temperature T_1 (red line) and T_2 (blue line) with $T_1 < T_2$. In the case of lower temperature, the most probable $v_{\rm mp}$ velocity, the average velocity $\langle v\rangle = 1.13 v_{\rm mp}$, and the (square root of the) mean square velocity $\sqrt{\langle v^2\rangle} = 1.15 v_{\rm mp}$ are, respectively, indicated by a dashed, dotted, and dash-dotted line.

$$\langle v^2\rangle = \frac{3k_{\rm B}T}{m} \tag{1.73}$$

where equation (1.56) has been used. We further calculate

$$\sqrt{\langle v^2\rangle} = \sqrt{\frac{3k_{\rm B}T}{m}} \tag{1.74}$$

which allows us to compare this value with the average and most probable velocities, as reported in figure 1.6 where we have made use of the numerical estimations $\langle v\rangle \simeq 1.13\ v_{\rm mp}$ and $\sqrt{\langle v^2\rangle} \simeq 1.25\ v_{\rm mp}$.

1.7.5 The entropy state function

As a matter of fact, for a monoatomic system three different equations of state $T = T(S, V, N)$, $P = P(S, V, N)$, and $\mu = \mu(S, V, N)$ can be formulated: equation (1.55) is just one of them. It must be stressed, however, that each single of such relations does not imply a complete knowledge of the thermodynamical properties of the system: this is only provided by the full set of them or, alternatively, by the single entropy state function. It is equally possible to formulate the thermodynamical theory by making use of the internal energy, which is as well a state function. We will, respectively, refer to the 'entropy representation' or to the 'energy representation' [8]. In this section we are calculating the entropy state function.

By inserting the partition function of the ideal monoatomic gas given in equation (1.52) into the general expression for entropy provided through equation (1.40) or (1.41) we get

$$S = \frac{5}{2}k_{\rm B}N + k_{\rm B}N \ln \frac{V(2\pi m k_{\rm B}T)^{3/2}}{Nh^3} \tag{1.75}$$

a result referred to as the *Sackur–Tedrode state equation*. We remark that during the early stages of development of the statistical approach it has been cumbersome to elaborate this result since the subtleties underlying the concept of distinguishable or indistinguishable particles were not all that clear. The discussion that arose around this problem and the apparent contradictions of the theory were referred to as the 'Gibbs paradox' [3, 4].

In order to provide an example of the powerfulness of the entropy state equation, let us consider an *adiabatic process*, namely a process not involving any exchange of heat between the system and its environment (we will also assume a constant mass). We assume that the ideal monoatomic gas undergoes an expansion from the initial volume V_{init} to the final one V_{fin}; its temperature accordingly varies from T_{init} to T_{fin}. By calculating explicitly the corresponding entropy values S_{init} and S_{fin} and setting $\Delta S = S_{\text{fin}} - S_{\text{init}} = 0$ we obtain

$$\ln V_{\text{init}} + \frac{3}{2} \ln T_{\text{init}} = \ln V_{\text{fin}} + \frac{3}{2} \ln T_{\text{fin}} \tag{1.76}$$

and therefore

$$VT^{3/2} = \text{constant} \tag{1.77}$$

which is the equation describing an adiabatic transformation, as phenomenologically found in macroscopic thermodynamics [8, 9].

1.8 On the statistical concept of irreversible process

As a final argument of this tutorial approach to the statistical description of a classical system, we will address the very concept of *irreversible process*[19]. To this aim, we consider an ideal monoatomic gas initially confined within a container with volume V_{init}. Let us also suppose that such a container is linked through a removable wall to a second *empty* container m-times larger, i.e. with volume $V_{\text{final}} = mV_{\text{init}}$. By some means we now remove the separating wall and, therefore, we observe a *free expansion* of the gas from the initial volume V_{init} to the final one V_{final}, as pictorially shown in figure 1.7. During this expansion the gas temperature does not change, because the internal energy of the gas does not vary (see equation (1.57)): simply, the

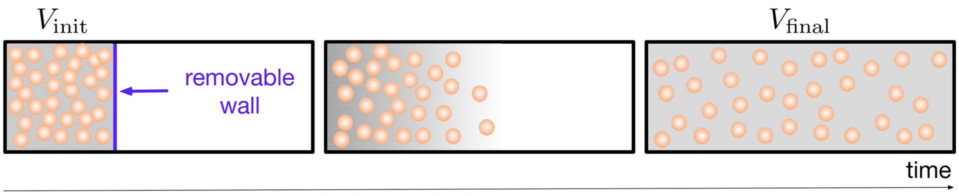


Figure 1.7. Pictorial representation of the free expansion process of a ideal monoatomic gas from the initial volume V_{init} to the final one V_{final}.

---

[19] We recall that in section 1.6 we basically proclaimed as *reversible* any process occurring quasi-statically, that is displaying over a time scale much larger than the intrinsic relaxation time of the system.

atoms move in a larger space and the gas does not need to spend work in expanding. From the macroscopic point of view this quintessentially is an irreversible process: no one reasonably expects that the gas, without external actions, can spontaneously return to the initial condition in which it occupied only the volume V_{init}. From the statistical point of view, the interpretation is more subtle and it is fully enlightened by proceeding with the calculation of the *entropy variation for a free expansion*.

Let us consider equation (1.75) and introduce the shortcut notation

$$\bar{V} = \frac{V}{N} \quad \text{and} \quad \bar{S} = \frac{5}{2}k_B N + k_B N \ln \frac{(2\pi m k_B)^{3/2}}{h^3} \tag{1.78}$$

which allows us to calculate the gas entropy in the initial and final conditions as

$$S_{\text{init}} = k_B N \ln(\bar{V}_{\text{init}} T^{3/2}) + \bar{S} \quad \text{and} \quad S_{\text{final}} = k_B N \ln(\bar{V}_{\text{final}} T^{3/2}) + \bar{S} \tag{1.79}$$

and the corresponding *entropy variation for the free expansion* as

$$\Delta S = S_{\text{final}} - S_{\text{init}} = k_B N \ln\left(\frac{\bar{V}_{\text{final}}}{\bar{V}_{\text{init}}}\right) = k_B N \ln m > 0 \tag{1.80}$$

since $m > 1$. This result ensures that the statistical description of the free expansion process is compatible with its thermodynamical counterpart: following whichever of the two approaches, the variation of the system entropy for a free expansion is consistently found to be positive. However, by remembering equation (1.37) we can recast equation (1.80) in the different form

$$\Delta S = k_B \ln \frac{\Omega_{\text{final}}}{\Omega_{\text{init}}} \tag{1.81}$$

which leads immediately to

$$\ln \frac{\Omega_{\text{final}}}{\Omega_{\text{init}}} = N \ln m \quad \rightarrow \quad \frac{\Omega_{\text{final}}}{\Omega_{\text{init}}} = m^N \tag{1.82}$$

where we exploited the link between the value of the entropy of any given macrostate and the corresponding number of compatible microstates. In general N is very large and, therefore, $\Omega_{\text{final}} \gg \Omega_{\text{init}}$ that is: *the final state is much more probable than the initial one*. On the other hand, if the reverse process $V_{\text{final}} \rightarrow V_{\text{init}}$ occurred, then we would get $\Omega_{\text{init}}/\Omega_{\text{final}} = m^{-N}$ which corresponds to an extremely small quantity indeed! In conclusion, *from the statistical point of view the free compression is not strictly speaking forbidden, rather it is 'just' very unlikely*.

While we developed this concept by studying a very simple model system, the conclusion is valid in general: *irreversible processes are processes that very unlikely occur in the opposite direction*. However, the smaller the number of particles forming the system, the higher is the probability to observe a process to occur in both directions. In this case, however, the statistical approach lacks a fundamental constitutive hypothesis (namely, that we are dealing with large systems) and, therefore, it is more fruitful to follow a mechanical one.

References

[1] Demtröder W 2010 *Atoms, Molecules and Photons* (Berlin: Springer)

[2] Colombo L 2021 *Solid State Physics: A Primer* (Bristol: IOP Publishing)

[3] Reif F 1987 *Fundamentals of Statistical and Thermal Physics* (New York: McGraw-Hill)

[4] Kennett M P 2021 *Essential Statistical Mechanics* (Cambridge: Cambridge University Press)

[5] Chandler D 1987 *Introduction to Modern Statistical Mechanics* (Oxford: Oxford University Press)

[6] Frenkel D and Smit B 1996 *Understanding Molecular Simulations: From Algorithms to Applications* (San Diego, CA: Academic)

[7] Lebon G, Jou D and Casas-Vázquez J 2008 *Understanding Non-equilibrium Thermodynamics* (Berlin: Springer)

[8] Callen H 1985 *Thermodynamics and an Introduction to Thermostatistics* (New York: Wiley)

[9] Dittman R and Zemansky M 1997 *Heat and Thermodynamics* 7th edn (New York: McGraw-Hill)

[10] Swendsen R H 2012 *An Introduction to Statistical Mechanics and Thermodynamics* (Oxford: Oxford University Press)

[11] Glazer M and Wark J 2001 *Statistical Mechanics: A Survival Guide* (Oxford: Oxford University Press)

[12] Eisberg R and Resnick R 1985 *Quantum Physics of Atoms, Molecules, Solids, Nuclei, and Particles* 2nd edn (Hoboken, NJ: Wiley)

[13] Miller D A B 2008 *Quantum Mechanics for Scientists and Engineers* (New York: Cambridge University Press)

[14] Bransden B H and Joachain C J 2000 *Quantum Mechanics* (Englewood Cliffs, NJ: Prentice-Hall)