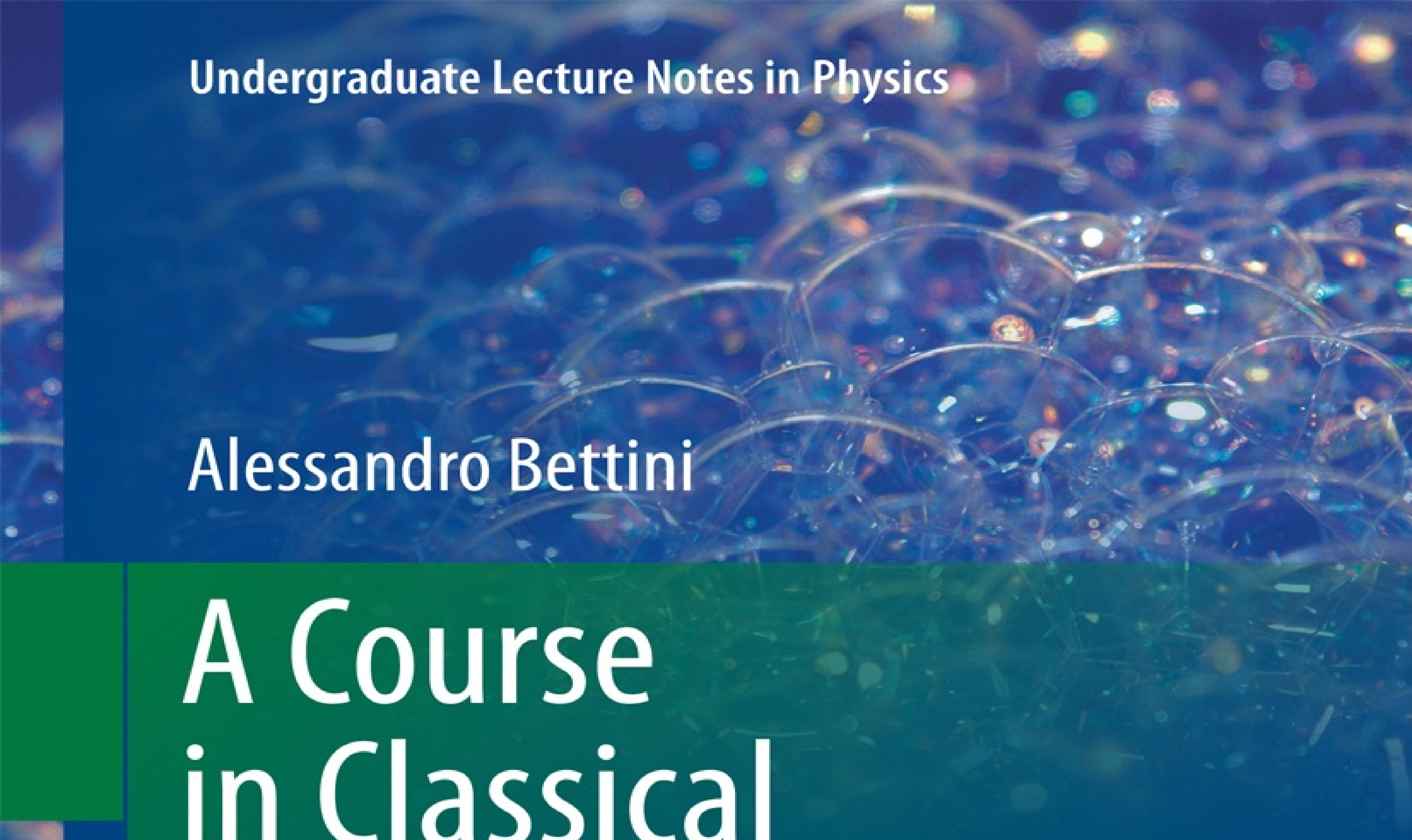
Undergraduate Lecture Notes in Physics



Alessandro Bettini

A Course in Classical Physics 2—Fluids and Thermodynamics

Springer

Undergraduate Lecture Notes in Physics

Undergraduate Lecture Notes in Physics (ULNP) publishes authoritative texts covering topics throughout pure and applied physics. Each title in the series is suitable as a basis for undergraduate instruction, typically containing practice problems, worked examples, chapter summaries, and suggestions for further reading.

ULNP titles must provide at least one of the following:

- An exceptionally clear and concise treatment of a standard undergraduate subject.
- A solid undergraduate-level introduction to a graduate, advanced, or non-standard subject.
- A novel perspective or an unusual approach to teaching a subject.

ULNP especially encourages new, original, and idiosyncratic approaches to physics teaching at the undergraduate level.

The purpose of ULNP is to provide intriguing, absorbing books that will continue to be the reader's preferred reference throughout their academic career.

More information about this series at http://www.springer.com/series/8917

Alessandro Bettini

A Course in Classical Physics 2—Fluids and Thermodynamics

Springer

Alessandro Bettini
Dipartimento di Fisica e Astronomia, Università di Padova, Padova, Italy

ISSN 2192-4791 e-ISSN 2192-4805

ISBN 978-3-319-30685-8 e-ISBN 978-3-319-30686-5
DOI 10.1007/978-3-319-30686-5

Library of Congress Control Number: 2016938649

The text is partially based on the book "Meccanica e termodinamica", A. Bettini, Zanichelli, 1995. The author own all rights in the former publications.

Printed on acid-free paper

This Springer imprint is published by Springer Nature The registered company is

Springer International Publishing AG Switzerland

Preface

This is the second in a series of four volumes, all written at an elementary calculus level. The complete course covers the most important areas of classical physics, such as mechanics, thermodynamics, statistical mechanics, electromagnetism, waves and optics. This second volume deals with fluid mechanics, thermodynamics and statistical mechanics.

The laws of Physics, and, more generally, of Nature, are written in the language of mathematics. The reader is assumed to have previous knowledge of the basic concepts of calculus: vectors, functions, limits and the derivative and integration operations.

Physics is an experimental science, meaning that it is based on the experimental method, which was developed by Galileo Galilei in the seventeenth century. He taught us, in particular, that to try to understand a phenomenon, one must simplify the relevant working conditions as thoroughly as possible, understanding which aspects are secondary and eliminating them as far as possible. The understanding process is not immediate, but rather, it proceeds by trial and error, through a series of experiments, which might lead, with a bit of fortune and a lot of thinking, to the discovery of the governing laws. Induction process of the laws of physics goes back from the observed effects to their causes, and, as such, cannot be purely logical. Once a physical law is found, it is necessary to consider all its possible consequences. This is now a deductive process, which is logical and similar to the mathematical one. Each of the consequences, the predictions, of the law must then be experimentally verified. If only one prediction is found to be false by the experiment, even if thousands of others have been found to be true, it is enough to prove that the law is false or, better yet, to show the limits of its validity. This implies that we can never be completely sure that a law is true; indeed, the number of its possible predictions is unlimited, and in any historical moment, a number of them may be uncontrolled. However, this is the price we must pay in choosing the experimental method, which has allowed humankind to advance much further in the last four centuries than in all the preceding millennia.

Thermodynamics and statistical mechanics are amongst the great intellectual constructions of Physics. Their laws are well established as well as the limits of their validity. Consequently, it can be exposed in an axiomatic way, as a chapter of mathematics. We can start from a set of propositions whose axioms are assumed to be true by definition, and deduce from them a number of theorems using only logics, as the Euclidean geometry theorems are deduced from the

Euclid postulates.

We shall not follow this path. The reason for this is that, while it allows a shorter and quicker treatment and is also logically more satisfactory for some, it also hides the inductive historical trial and error process through which the postulates and the general laws have been discovered. These are arrival rather than starting points. This path has been complex, laborious, and highly nonlinear. Errors have been made, hypotheses have been advanced that turned out to be false, but finally the laws were discovered. The knowledge of at least a few of the most important aspects of this process is indispensable for developing the mental capabilities that are necessary to anybody contributing to the progress of natural sciences, whether they pursue applications or teach them. In any case, we shall mention the names of those that contributed most to the achievements that we will be discussing, along with the date of the discovery and, the first time we meet him, the life span of the author.

A large fraction of the book deals with the physics of fluids. We shall start with their mechanical properties, continue with their thermodynamic aspects, and end up with the statistical mechanics of their molecular structure. In Chap. 1 , we shall study the statics and the dynamics of fluids, called hydrostatics and hydrodynamics, respectively. This is, rigorously speaking, a chapter of mechanics, but fluids are much more complex mechanical systems than, for example, the rigid bodies studied in the first volume. As a matter of fact, fluids have an infinite number of degrees of freedom. We shall describe several aspects of the complex fluid phenomenology aimed at understanding the physics rather than the mathematics. The latter requires a knowledge of partial differential equations that is beyond the level of this course.

In the first volume, we learned the fundamental conservation laws of energy, linear and angular momentum. We also saw that the total mechanical energy of an isolated system is not always conserved. It is not conserved in the presence of nonconservative forces. We had anticipated then, however, that the apparent nonconservation of energy is due to having neglected to include all the forms of energy in the balance. In the second chapter of this volume, we shall see that, indeed, energy is always conserved. Thermodynamics teaches us how one must take into account all the possible forms of energy exchange. Two systems can exchange energy not only in the form of work but also in the form of heat. Beyond the mechanical energy, we must also include the internal, or thermal, energy in the balance. The first law of thermodynamics is the law of energy conservation.

Thermodynamics deals with systems that are extremely complex from a mechanical point of view. In order to define the mechanical state of a fluid, for example, one should know the positions and velocities of all its molecules. This is not possible. Thermodynamics describes the state of the system under study with a small number of global variables, like volume, pressure, density and temperature, and the processes from one state to another considering the heat and work exchanges. In Chap. 2 , we shall study, in particular, two important classes of thermodynamic systems, the gases and the solids.

The second law of thermodynamics, discussed in Chap. 3 , deals with the irreversibility of natural phenomena. For example, if two bodies at different temperature are brought into contact, heat passes from the hotter to the colder one; the temperature of the former decreases, while that of the latter increases. The opposite process never happens spontaneously. As another example, if we drop a stone from a certain height, it stops when it hits the ground and its temperature increases. It never happens that a stone on the ground jumps up while cooling. We shall learn how entropy, a fundamental quantity of thermodynamics, rules the irreversibility.

In Chap. 4 , we shall apply the laws of thermodynamics to several relatively simple thermodynamic systems. After having given some information on the structure of matter and on its aggregation phases, we shall study the conditions for equilibrium between phases (liquid and vapor, liquid and solid, solid and vapor), the transitions between the phases and the surface phenomena.

In the final two chapters, we shall look at the thermodynamic processes from the microscopic point of view, namely considering that the bodies are made of an enormous number of molecules. We shall study statistically the kinematic variables, namely their probability distributions and their average values. In this way, we shall learn that (classical) thermodynamics laws are not independent of (classical) mechanics, but rather logical consequences of same. Historically, the most important steps forward in physics happen when fields that had been separated become unified in a single theory. This had been the case for terrestrial and heavenly mechanics with Galilei and Newton in the seventeenth century, as we saw in the first volume. Similarly, thermodynamics (and chemistry as a part of it) was unified with mechanics in the second half of the nineteenth century by, mainly, James Clerk Maxwell and Ludwig Boltzmann. The study of statistical mechanics will enlighten and give deep physical meaning to several findings within thermodynamics. It shall also lead us to discover the limits of classical mechanics, the limits at which quantum physics takes over.

Each chapter of the book starts with a brief introduction, to give the reader a preliminary idea of the arguments he/she will find. There is no need to fully understand these introductions at the first reading, as all the arguments are fully developed in the subsequent pages.

At the end of each chapter, the reader will find a number of queries, through which to check his/her level of understanding of the arguments put forward in the chapter. The difficulty of the queries is variable; some of them are very simple, some more complex, a few are true numerical problems. On the other hand, the book does not contain a sequence of full problems, owing to the existence of very good textbooks dedicated specifically to those.

The answers to the large majority of the queries are included. However, the solution to numerical problems (without looking at the answers) is mental gymnastics that are absolutely necessary for understanding the subject. Only the effort to apply concepts one has learned to specific cases will allow the reader to master them completely. The reader should be conscious of the fact that the solution of numerical problems requires mental mechanisms different from those engaged in understanding a text. The latter, indeed, has already been organized by the author; solving a problem requires much more active initiative from the student, a creative activity that is needed for advancing scientific knowledge and its technical applications as well. Consequently, the student should work on an exercise alone, without looking at the solution in the book. Even failed attempts to reach the solution autonomously, provided they are undertaken with sufficient persistence, yield important returns, because they aid in the development of processing skills. If, after several failed attempts, the solution has not yet been reached, it is a better practice to abandon the exercise momentarily, rather than looking at the solution, instead going on to another exercise and coming back to the previous one later.

The following working scheme is methodologically advisable:

1. Examine the conditions posed by the problem in depth. If it is appropriate, make a drawing containing the essential elements.

2. Solve the problem using letters in the formulas, not numbers, developing them up to the point when the requested quantities are expressed in terms of the known ones. Only then should you put numbers into the formulas.

3. Control the correctness of the physical dimensions.

4.
When necessary, transform all the data in the same system of units (prefer SI). Use the scientific notation, for example, 2.5×10^{3} rather than 2500, 2.5×10^{-3} rather than 0.0025. In general, two or three significant figures are enough.

5. Once you have the final result, always verify if it is reasonable. For example, the mass of a molecule cannot turn out to be 30 mg, the speed of a bullet cannot be 10^{6} m/s, the distance between two towns cannot be 25 mm, etc.

Acknowledgments

The author is grateful to Andrej Gogala for his kind permission to use the photo in Fig. 4.17 .

Symbols

Table 1 Symbols for the principal quantities

Acceleration	$\mathbf{a}$, a_s
Absolute temperature	T
Angular acceleration	$\boldsymbol{\alpha}$, α
Angular frequency	ω
Angular momentum	$\mathbf{l}$, $\mathbf{L}$
Angular velocity	ω
Atmospheric pressure	p_a
Avogadro number	N_A
Boltzmann constant	k_B
Boundary layer thickness	δ
Cross section	σ
Debye temperature	θ_D
Density (mass)	ρ
Density (numerical)	n_p
Diameter	D
Diffusion coefficient	D
Drag coefficient	C_D
Efficiency (thermic engine)	η
Efficiency (refrigerator)	ξ
Electric charge (elementary)	q_e
Enthalpy	H
Entropy	S
Force	$\mathbf{F}$
Frequency	ν
Gas constant	R
Gas specific heats ratio	γ
Gravitational field	$\mathbf{G}$
Gravity acceleration	$\mathbf{g}$
Heat	Q
Heat capacity	C
Heat flux density	Φ_Q
Internal energy	U

Isothermal compressibility	κ
Latent heat vaporization	Q_{vap}
Kinetic energy	U_K
Mass	m , M
Mass flow rate	Q_m
Mass flux density (diffusion)	Φ_c
Mean free path	l
Mechanical equivalent of heat	J
Molar heat	C , C_p , C_V
Molar volume	V_m
Molecule kinetic energy flux density	Φ_T
Moment (mechanical)	M
Momentum	$\mathbf{p}$
Momentum flux density	Φ_p
Normal (to a surface) unit vector	$\mathbf{n}$
Number density	n_p
Number of moles	n
Plane angle	θ
Polar angle	θ , ϕ
Polar coordinates (space)	ρ , θ , ϕ
Position vector	$\mathbf{r}$
Potential energy	U_p
Power	w
Pressure	p
Radius	R , r
Reduced mass	μ
Reynolds number	Re
Shear stress	τ
Specific heat	c , c_p , c_V
Spring constant	κ
Surface	S , Σ
Surface tension	τ
Temperature	θ
Thermal compressibility	κ
Thermal conductivity	κ
Thermal diffusivity	χ
Time	t

Total (mechanical) energy	U_{tot}
Weight	$\mathbf{F}_w$
Work	W
Mean value, of x	$\langle x \rangle$
Angular velocity	$\boldsymbol{\omega}$, $\boldsymbol{\Omega}$
Velocity of light (in vacuum)	c
Velocity	$\mathbf{v}$, υ
Unit vector of $\mathbf{v}$	$\mathbf{u}_\upsilon$
Unit vectors of the axes	$\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$
Van der Waals parameters	a , b
Viscosity (dynamical)	η
Viscosity (kinematic)	ν
Velocity flux density	Φ_υ
Velocity, root mean square	υ_{rms}
Volume	V
Volumetric flow rate	Q_V
Volumetric thermal expansion coefficient	α

Table 2 Base units in the SI

Quantity	Unit	Symbol
Length	meter/metre	m
Mass	kilogram	kg
Time	second	s
Current intensity	ampere	A
Thermodynamic temperature	kelvin	K
Amount of substance	mole	mol
Luminous intensity	candela	cd

Table 3 Decimal multiples and submultiples of the units

Factor	Prefix	Symbol	Factor	Prefix	Symbol
10^{24}	yotta	Y	10^{-1}	deci	d
10^{21}	zetta	Z	10^{-2}	centi	c
10^{18}	exa	E	10^{-3}	milli	m
10^{15}	peta	P	10^{-6}	micro	µ
10^{12}	tera	T	10^{-9}	nano	n
10^{9}	giga	G	10^{-12}	pico	p

10^{6}	mega	M	10^{-15}	femto	f
10^{3}	kilo	k	10^{-18}	atto	a
10^{2}	hecto	h	10^{-21}	zepto	z
10	deka	da	10^{-24}	yocto	y

Table 4 Fundamental constants

Quantity	Symbol	Value	Uncertainty
Speed of light in vacuum	c	299,792,458 m s^{-1}	Defined
Elementary charge	q_e	$1.60217653(14) \times 10^{-19}$ C	85 ppb
Electron mass	m_e	$9.1093826(16) \times 10^{-31}$ kg	170 ppb
Proton mass	m_p	$1.67262171(29) \times 10^{-27}$ kg	170 ppb
Newton constant	G_N	$6.67384(80) \times 10^{-11}$ m^{3} kg^{-1} s^{-2}	120 ppm
Gas constant	R	8.3144598(48) J mol^{-1} K^{-1}	580 ppb
Water triple point temp.		273.16 K	Defined
Avogadro number	N_A	$6.0221415(10) \times 10^{23}$ mole^{-1}	170 ppb
Boltzmann constant	k_B	$1.3806505(24) \times 10^{-23}$ J K^{-1}	1.8 ppm

Table 5 Greek alphabet

alpha	α	Α	iota	ι	Ι	rho	ρ	Ρ
beta	β	Β	kappa	κ	Κ	sigma	σ, ς	Σ
gamma	γ	Γ	lambda	λ	Λ	tau	τ	Τ
delta	δ	Δ	mu	μ	Μ	upsilon	υ	Υ, ϒ
epsilon	ε	Ε	nu	ν	Ν	phi	ϕ, φ	Φ
zeta	ζ	Ζ	xi	ξ	Ξ	chi	χ	Χ
eta	η	Η	omicron	ο	Ο	psi	ψ	Ψ
theta	θ, ϑ	Θ	pi	π	Π	omega	ω	Ω

Contents

Alessandro Bettini, A Course in Classical Physics 2—Fluids and Thermodynamics, Undergraduate Lecture Notes in Physics, DOI 10.1007/978-3-319-30686-5_1

1. Fluid Dynamics

Alessandro Bettini[1] ✉

(1) Dipartimento di Fisica e Astronomia, Università di Padova, Padova, Italy

✉ **Alessandro Bettini**
Email: alessandro.bettini@pd.infn.it

Both liquids and gases are called fluids. As opposed to solids, fluid substances do not have a definite shape. Liquids do have a definite volume, as, in a very good approximation, they are not compressible. On the other hand, gases occupy all the volume at their disposal and can be easily compressed.

In this chapter, we shall study the motions of fluids, which may be very complex. In the first volume of the course, we studied the motions of material points, which have three degrees of freedom, and of rigid bodies, which have six. Fluids, on the other hand, have an infinite number of degrees of freedom.

In the first five sections of the chapter, we shall study fluids at rest, in equilibrium conditions. This chapter of physics is called hydrostatics. We shall see, in particular, examples of equilibrium in the fields of weight and centrifugal (pseudo) force.

We shall then introduce the concept of viscosity, which is the internal friction in fluids in motion, and discuss its dynamical effects. Its physical meaning, from the microscopic point of view, will be studied in Chap. 6. In Sects. 1.7–1.10, we shall consider the motions of an idealized system, the ideal fluid, which is a fluid that is incompressible and with zero viscosity. Even if such a fluid does not exist, under certain conditions, real fluids (even gases, in some instances) behave approximately like that ideal one.

In the last four sections, we shall consider real fluids. We shall see how the presence of viscosity often radically changes the flow. Next, we shall consider the flow in cylindrical pipes and its different regimes, with the laminar, ordered one being examined in Sect. 1.11, and the chaotic, turbulent one in Sect. 1.12. Following that, we shall study the flow of a fluid around a body (equivalent to the motion of a body in a fluid), which we shall take to be spherical. Again, we shall start with the laminar regime in Sect. 1.13, and finally move to the turbulent one in Sect. 1.14. The situations rapidly become complicated. However, it is possible to employ dimensional arguments to analyze the principal characteristics of complex motions in a simple way. We shall see, in particular, how the different regimes are characterized by different values of a dimensionless quantity, the Reynolds numbers.

1.1 Fluids

Gases and liquids are collectively called *fluids* . As opposed to solids, fluids do not have a proper shape; rather, they can be easily deformed. Liquids and gases have some common and some different characteristics. Let us start with the latter.

The volume of liquids is limited by sharp surfaces, both those in contact with the container and the free one. The density of liquids is much larger than that of gases under the usual conditions. For example, the density of water is 10^3 kg/m^3, while the density of air at normal pressure and temperature is 1.3 kg/m^3. Liquids are practically incompressible, as are solids, and we have said that their volume is definite. Contrastingly, gases can be easily compressed and their density strongly depends both on pressure and temperature.

In liquids, the distances between the molecules are similar to their diameters, we can say they touch each other, while in gases they are much larger. In both cases, however, these distances are very small compared to the macroscopic sizes. For example, in a gas like nitrogen or oxygen, under normal conditions, the mean distance between molecules is a few nanometers. We can also posit that a volume of 0.1 mm diameter of water contains along the order of 10^{17} molecules while one of air at ambient atmospheric pressure and temperature contains 10^{14} molecules. We see that we can consider both media to be continuous at the macroscopic level with a good approximation.

Consider the infinitesimal cubic element of a fluid shown in Fig. 1.1. Its volume is $dV = dx\, dy\, dz$ and its mass is $dm = \rho(x, y, z)dV$, where ρ is the fluid density, which is possibly a function of the position. The forces acting on the element can be divided into two categories:

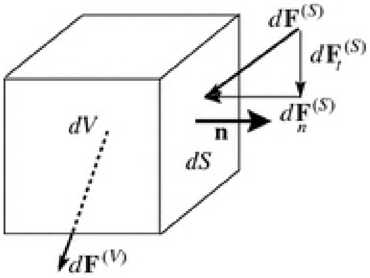


Fig. 1.1 Infinitesimal element in a fluid and forces acting on it

(1) the volume forces $d\,\mathbf{F}^{(V)}$ (see Fig. 1.1) that are proportional to the volume and the mass of the element. Such are the weight and, in a non-inertial frame, the apparent inertia forces. In general, we can say

$$d\mathbf{F}^{(V)} = \mathbf{G}(x,y,z)\,dm = \mathbf{G}(x,y,z)\rho(x,y,z)\,dV, \tag{1.1}$$

where **G** generically indicates the force per unit mass (it is the gravity acceleration **g** in the particular case of the weight).

(2) the surface forces $d\,\mathbf{F}^{(S)}$ (see Fig. 1.1) that act on each of the (geometrical) faces that delimit the element and are due to the fluid outside the element touching it on that face. As we shall soon see, it is convenient to divide the surface force into its components $d\mathbf{F}_n^{(S)}$ normal and $d\mathbf{F}_t^{(S)}$ tangent to the surface. The magnitude of the surface force is proportional to the area of the surface dS through which it acts. The force per unit area is called *stress* . The normal (to the surface) and shear stresses are

$$p = \frac{dF_n^{(S)}}{dS}\ (\text{normal stress})\,; \qquad \boldsymbol{\tau} = \frac{d\mathbf{F}_t^{(S)}}{dS}\ (\text{shear stress})\,. \tag{1.2}$$

The normal stress is more commonly called *pressure* . We shall immediately show that the normal stress at a given point in the fluid is independent of the direction in which it exerts. Pressure is a scalar, not a vector. As a matter of fact, we have anticipated this result and used the symbol p that already indicates a scalar. Usually, the normal stress on a surface is directed towards the inner part of the volume, being due to the external elements pushing on it. In exceptional situations, it may happen that the external elements attract the one under consideration. Under these conditions, the pressure is negative. We shall consider an example of that in Chap. 4, while in this one we assume the pressure to be always positive.

The shear stress has, constrastingly, a direction. This is, by definition, parallel to the surface on which it is exerted. If the fluid is in equilibrium, all its elements are by definition at rest. Considering that there is no constraint forbidding contiguous elements from shifting one over the other, we understand that, in a fluid at equilibrium, all the shear stresses must be zero. In a fluid, the equivalent of static friction, which opposes the relative motion of two surfaces of solid bodies in contact, does not exist. However, as we shall see, if the liquid is in motion, such as water in a duct, for example, shear stresses are present.

The physical dimensions of pressure and shear stress are those of a force divided by a surface. The unit in the SI is called pascal (Pa), in honor of Blaise Pascal (1623–1662). One pascal is one newton per square meter. This is a rather small pressure. To get an idea, imagine that the atmospheric pressure at sea level is about 10^5 Pa. A multiple of the pascal, the bar = 10^5 Pa, is often used, because it is close to the old "atmosphere" unit, even if it is not SI. We shall further discuss pressure measurements in Sect. 1.3.

We now show that, at every point in a fluid at rest, the pressure through all the surface elements through the point O is the same, independent of the orientation of the surface. This is an immediate consequence of the shear stresses being zero.

We take the coordinate axes with origin in O, as shown in Fig. 1.2. We then take a tetrahedron having three surfaces on the coordinate planes and the fourth one, ABC, inclined and near to O. Let $\mathbf{n}$ be the external normal to ABC. Both volume and surface forces act on the tetrahedron, which, by definition, is in equilibrium. We do not know the sufficient conditions for the equilibrium, because we are dealing with a fluid element. However, we know the necessary ones. We mentally solidify the element. The conditions necessary and sufficient for the equilibrium of the "solidified" element are certainly necessary for the equilibrium of the liquid element.

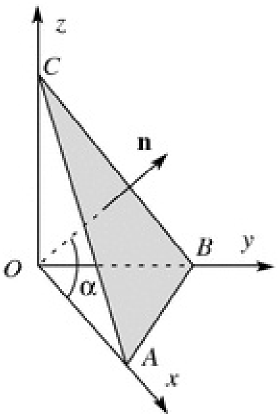


Fig. 1.2 Pressure is independent of direction

We first observe that the volume forces, the weight in particular, are infinitesimal of the third order of the linear dimensions, because they are proportional to the volume. The surface forces are infinitesimal of the second order. We can consequently neglect the volume forces. We conclude that, for equilibrium, the pressure forces must have zero resultant. Let us impose that, starting with one direction, *x*, for example. Let p_x be the pressure on the face normal to *x*, which is *OCB* and has the area dS_x. The corresponding force is in the positive direction of the *x*-axis. We call *α* the angle between the unit vector **n** and the *x*-axis and *dS* the area of *ABC*. The pressure *p* exerts a force on *ABC* that has the direction of −**n**. Its component on the *x*-axis is equal to $-pdS\cos\alpha$. Hence, the equilibrium condition is $p_x dS_x - p\cos\alpha dS = 0$.

On the other hand, geometry immediately gives $dS_x = \cos\alpha dS$, and we have $p_x = p$. Similar conditions hold for the other components and we have

$$p_x = p_y = p_z = p,$$

which is what we had to show.

The pressure in a fluid in equilibrium is independent of the direction in which it is exerted.

1.2 Fluid Statics

The density of a fluid can be different at its different points, namely it may be a function of the coordinates, *ρ*(*x*, *y*, *z*). Let us see how it varies. We start by choosing an inertial frame, as in Fig. 1.3, and mentally insulating an infinitesimal cubic element of the fluid.

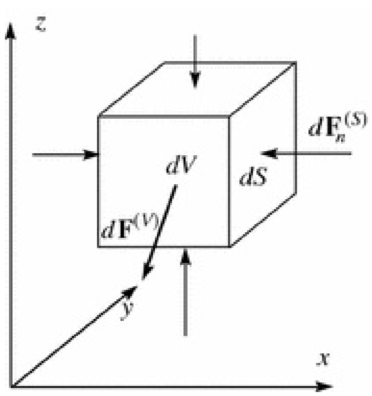


Fig. 1.3 Pressure and volume forces on an infinitesimal cubic fluid element

Let (x, y, z) be the coordinates of the element, $dV = dx\,dy\,dz$ its volume and dS the area of the faces. We indicate with $d\,\mathbf{F}^{(S)}$ the resultant of the (6 in number) surface forces and with $d\,\mathbf{F}^{(V)}$ the volume force. The latter is proportional to the mass of the element, namely

$$d\mathbf{F}^{(V)} = \mathbf{G}(x,y,z)\,dm = \mathbf{G}(x,y,z)\rho(x,y,z)\,dV.$$

A necessary equilibrium condition is the resultant force being zero

$$d\mathbf{F}^{(S)} + d\mathbf{F}^{(V)} = 0. \qquad (1.3)$$

Consider the components parallel to one axis, z, for example. The contributing surface forces are those that exert through the two planes normal to z. Adding them up, we have $p(x,y,z)\,dS - p(x,y,z+dz)\,dS = -\frac{\partial p}{\partial z}dz\,dS = -\frac{\partial p}{\partial z}dV$. The z component of the volume force is $d\mathbf{F}_z^{(V)} = \mathbf{G}_z(x,y,z)\rho(x,y,z)\,dV$, and Eq. (1.3) gives

$$\mathbf{G}_z(x,y,z)\rho(x,y,z)\,dV - \frac{\partial p}{\partial z}dV = 0,$$

or

$$\frac{\partial p}{\partial x} = \mathbf{G}_z(x,y,z)\rho(x,y,z).$$

Similar expressions hold for the other axes, and we can summarize

$$\frac{\partial p}{\partial x} = G_x\rho, \quad \frac{\partial p}{\partial y} = G_y\rho, \quad \frac{\partial p}{\partial z} = G_z\rho, \qquad (1.4)$$

or, in an equivalent, more synthetic notation

$$\mathbf{grad}\,p(x,y,z) = \rho(x,y,z)\,\mathbf{G}(x,y,z). \qquad (1.5)$$

A relevant example of the volume force is weight. Neglecting the effects of the earth's rotation, $\mathbf{G} = \mathbf{g} = (0, 0, -g)$, namely the gravity acceleration. Equation (1.4) becomes

$$\frac{\partial p}{\partial x} = 0, \quad \frac{\partial p}{\partial y} = 0, \quad \frac{\partial p}{\partial z} = -g\rho. \tag{1.6}$$

The pressure decreases with increasing height. Every volume element (see Fig. 1.4) is subject to its weight, a force that must be equilibrated by the pressure forces, acting on its horizontal faces. The pressure below must push more than the pressure above.

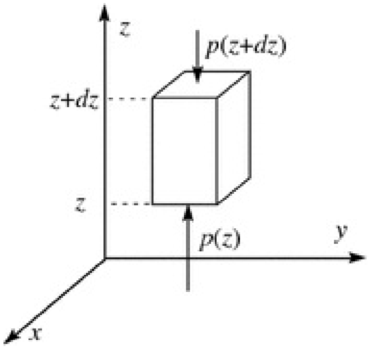


Fig. 1.4 A fluid element in the field of the weight

We also see that the pressure varies with height but not with the other coordinates, namely that it is constant on any horizontal plane. The surfaces of constant pressure are called *isobaric surfaces* . In the field of the weight force, the isobaric surfaces are the horizontal planes; such are also, it should be remembered, the equipotential surfaces . As a matter of fact, the equality of isobaric and equipotential surfaces is a general property when the volume force is conservative. If it is such, and we indicate its potential with ϕ, the force per unit mass is $\mathbf{G} = -\mathbf{grad}\phi$ and

$$\mathbf{grad}\, p(x,y,z) = -\rho(x,y,z)\,\mathbf{grad}\phi(x,y,z). \tag{1.7}$$

The gradients of the pressure and of the potential are parallel. The isobaric and equipotential surfaces, which are normal to them, are parallel too, and so they coincide.

A further important property is that the isobaric surfaces are also surfaces of constant density. We show that considering two isobaric, hence also equipotential, surfaces S and S', infinitely near to each other (see Fig. 1.5). Let $d\mathbf{n}$ be the vector normal to the surfaces from S to S', dp the pressure difference and $d\phi$ the potential difference.

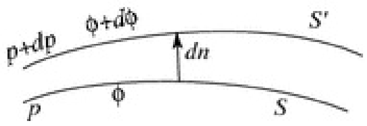


Fig. 1.5 Two infinitely close isobaric and equipotential surfaces

It follows from the properties of the gradient that *dp* is equal to the dot product of its gradient and of *d* **n**, and the same can be said for $d\phi$. Hence,

$$dp = \mathbf{grad}\, p \cdot d\mathbf{n}, \quad d\phi = \mathbf{grad}\phi \cdot d\mathbf{n}.$$

but, for Eq. (1.7), we can write $dp = -\rho(x,y,z)\,\mathbf{grad}\phi \cdot d\mathbf{n}$, and finally

$$\frac{dp}{d\phi} = -\rho(x,y,z).$$

Now, the left-hand side of this expression does not depend on the coordinates on the surfaces. Hence, the same must be valid for the density ρ.

In conclusion, in a fluid at rest in a field of conservative forces, the isobaric surfaces are also equipotential and constant density surfaces. We observe that when we look at the free surfaces of the liquids, which, on small scales, are horizontal planes. On larger extensions, comparable to those of the earth, such as on lakes and seas, free surfaces are spherical. The same arguments explain the spherical shape of the planets and of many celestial bodies. Several of them are gaseous, such as the sun and some of the planets. As for the solid planets, like earth, one should consider that they were fluid when they formed. To be precise, a fluid celestial body is not at rest in an inertial frame, due mainly to its rotation about its axis. We can extend the above arguments to include the inertial forces, in particular, the centrifugal force and its potential. The resulting equipotential surfaces are not spherical, but rather rotation oblate ellipsoids.

1.3 Fluids in the Weight Field

In this section, we shall consider the case of the weight force and take the *z*-axis vertical upwards. **G** is the gravity acceleration and we have

$$G_x = 0, \quad G_y = 0, \quad G_z = -g. \tag{1.8}$$

Equation (1.6) holds

$$\frac{\partial p}{\partial x} = 0, \quad \frac{\partial p}{\partial y} = 0, \quad \frac{\partial p}{\partial z} = -g\rho. \tag{1.9}$$

If we want to know how the pressure varies with height, we need to know something about density. We shall consider here two important cases; a liquid, like water, and a gas, like the atmosphere.

In the case of a liquid, the density can be considered constant, namely independent of pressure, in a good approximation. We can write the latter part of Eq. (1.9) as $dp = \rho g dz$. We integrate to find the pressure difference between two points, say A and B, at different heights (Fig. 1.6). We find $p(z_B) - p(z_A) = -\rho g(z_B - z_A)$. Calling $h = z_B - z_A$, we have

$$p(z_A) = p(z_B) + \rho gh, \tag{1.10}$$

which is known as the *Stevin law* , after Simon Stevin (1548–1620). The quantity ρgh is the *hydrostatic pressure* , which is the pressure exerted by a column of homogenous liquid of height h on its base. Indeed, consider a fluid cylinder of height h with base of area S. Its mass is $Sh\rho$ and its weight $Sh\rho g$. This is the force that acts on the basis. Dividing it by the area of the basis, we have the pressure, $Sh\rho g/S = \rho gh$.

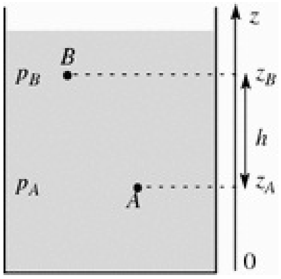


Fig. 1.6 Liquid cylinder and two points at different heights

The Stevin law tells us, in particular, how the pressure varies when we dive underwater. At every meter of depth, the pressure varies of $\rho g = 10^3$ kg/m^3 × 9.8 m s^{-2} ≈ 10^4 Pa. In other words, the pressure underwater increases by one atmosphere (about 10^5 Pa) every 10 m. This is why, diving down, one must “compensate” the pressure in the ears every few meters.

The atmospheric pressure on earth is the weight of the air column on its basis. In this case, however, we cannot consider the density to be constant. On the contrary, the pressure continuously decreases with increasing altitude. Already at 40 km, the air density is only a few per mille of the sea level value.

The first measurement of the atmospheric pressure and the correct interpretation of the experiment, in 1644, were carried out by two pupils of Galilei, respectively, Vincenzo Viviani (1622–1703) and Evangelista Torricelli (1608–1647). The instrument used by Viviani is known as the *Torricelli barometer* , and is shown in Fig. 1.7. To build it, one takes a glass tube about one meter long, closed at one extreme, turns it vertical and completely fills it with mercury. The open end is then closed with a finger (use a glove; mercury is toxic), the tube turned upside down and its extreme immersed in a basin of mercury. If we now open the extreme, we see the free surface of mercury in the tube lowering somewhat and then reaching the equilibrium level, when the column height is, say, *h*.

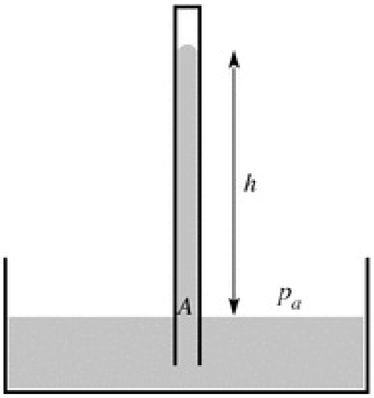


Fig. 1.7 The Torricelli barometer

Notice that if we move the tube vertically up or down, provided the lower extreme remains in the mercury and some space remains on top, the height *h* of the column over the free surface in the basin does not vary. In addition, if you repeat the experiment with tubes of different diameters and different shapes, including curved ones, you shall observe that the height of the column is always the same.

The interpretation of Torricelli was as follows. All the points at the level of the free surface in the basin, both on the surface itself and inside the tube, like *A* in the figure, are at the same pressure , which is the atmospheric pressure p_a; otherwise, it would not be in equilibrium. However, the air column does not act on the horizontal section *A* inside the tube. To what is the equilibrating force due? Torricelli stated that no air could have possibly entered into the tube above the mercury. That space had to be empty. Note that, at that time, the majority of scientists did not accept the existence of a vacuum . If this is the case, he continued, the equilibrium is due to the weight of the mercury column . Its pressure on the base in *A* is ρgh, where ρ is the density of mercury.

To be historically precise, the Viviani and Torricelli experiment had been anticipated by a similar and highly spectacular one made with water around 1641 by Gaspero Berti (ca 1600–1643). A reconstruction is shown in Fig. 1.8. G. Galilei had stated that water cannot be raised in a pipe more than “18 arms”, which is about 10.5 m. Berti fastened a vertical tube to the façade of his palace in Rome, somewhat longer, namely 12 m, to be safe, with a glass globe tightly fixed on the top. On the bottom, the tube ended in a jar and was closed with a tap. The upper globe had an opening in its top (*C* at the top of the figure). Globe and tube were completely filled with water through that opening, which was then closed with a cork. The tap near the road was opened. Water started to run out in the jar, and the free surface of water moved down in accord, for a while. But soon, a steady level was reached, which did not change until the following morning. When, at that time, the cork on the globe was removed, all the water poured out. Berti measured the height of the column, from the free surface in the jar. It was just 18 arms.

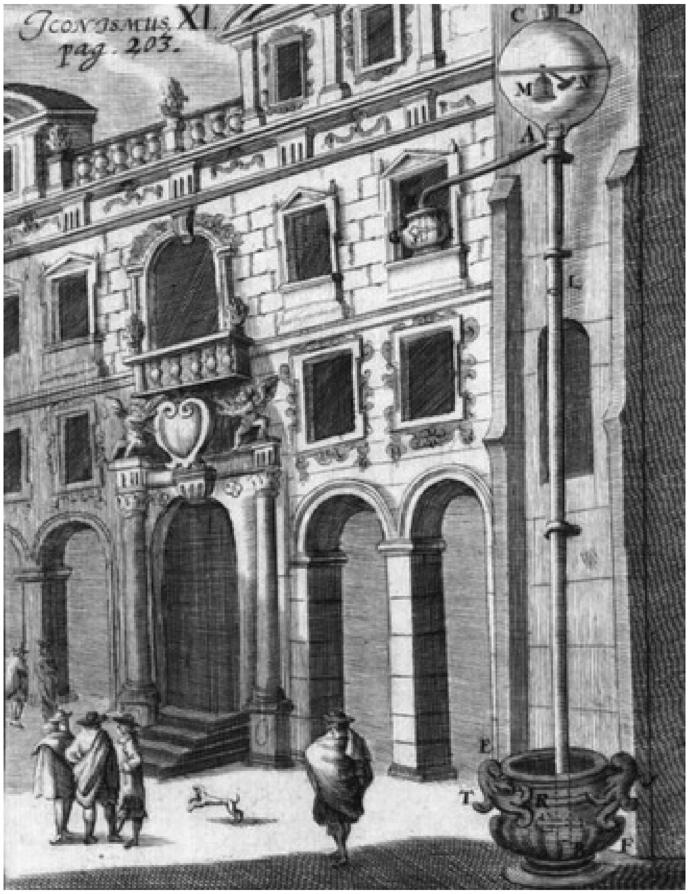


Fig. 1.8 The Berti's experiment. From "Technica curiosa, sive mirabilia artis" by C. Schott, 1687

The Torricelli barometer is still used in laboratories. With the known values of $\rho = 13.596 \times 10^3$ kg/m^3 and of *g*, we can determine the atmospheric pressure, by measuring the height *h*. Under normal conditions at sea level, it is $h = 760$ mm. The atmospheric pressure is then

$$p_a = \rho g h = 13.596 \times 9.81 \times 0.76 = 1.012 \times 10^5 \,\text{Pa}. \tag{1.11}$$

We have already used this value. Obviously, at a given point, the pressure varies in time depending on the weather conditions. The just-calculated value is an average, which is assumed by definition to be the standard atmospheric pressure . It is often used as measurement unit, called *atmosphere* (atm)

$$1\,\text{atm} = 1.012 \times 10^5\,\text{Pa}. \tag{1.12}$$

Another often used unit, non-SI, is the pressure of a 1 mm height mercury column, called a torr . Clearly, one torr is 1/760 of an atmosphere

$$1\,\text{torr} = 133.2\,\text{Pa} = 1.332\,\text{hPa}. \tag{1.13}$$

In practice, the atmospheric pressure is measured with barometers, which are handier to use. A common type is the aneroid barometer , which consists of a box partially exhausted of air. Its top is an elastic disk, which budges inwards under the pressure to measure. The deformation moves a pointer connected to the disk through suitable levers that amplify the displacement. The pointer moves on a scale, on which the pressure values can be directly read. While the Torricelli barometer gives absolute measurements, the aneroid barometers give only relative ones. Their scale must be calibrated on an absolute instrument.

We now consider how the pressure of a gas, subject to its weight, varies with the height, or the altitude in the atmosphere. To make things simpler, we assume the temperature to be uniform, namely independent of the height. Obviously, this is not the case in the atmosphere over large differences of altitude, but it is approximately true over moderate drops (a few hundred meters). In a gas at constant temperature, there is a well-defined relation between pressure and density; this is the gas law, which we shall study in the next chapter. We anticipate here that it simply states that the density is proportional to the pressure, or $p/\rho = \text{const}$.

We take the z-axis, as usual, to be vertical upward with its origin at sea level. We indicate with p_0 and ρ_0 the pressure and the density at this level. The gas law tells us that

$$\rho(z) = \frac{\rho_0}{p_0} p(z). \tag{1.14}$$

Using the Stevin law for a column of height dz, we have

$$dp = -\rho(z)\,g\,dz = -\frac{\rho_0}{p_0} g\,p(z)\,dz$$

and, spearing the variables

$$\frac{dp}{p(z)} = -\frac{\rho_0}{p_0} g\,dz.$$

This is a differential equation. The unknown is the function $p(z)$. As we have only one of the two variables on each side, we solve it by integration from the reference height $z = 0$ to the generic z

$$\int_0^z \frac{dp}{p(z)} = -\frac{\rho_0}{p_0} g \int_0^z dz \quad \Rightarrow \quad \ln \frac{p(z)}{p_0} = -\frac{\rho_0}{p_0} gz.$$

We take the exponential of both sides, obtaining

$$p(z) = p_0 e^{-z/\lambda} \quad \text{with} \quad \lambda = \frac{p_0}{\rho_0 g}. \tag{1.15}$$

The gas law immediately gives the density's corresponding dependence on the altitude

$$\rho(z) = \rho_0 e^{-z/\lambda}. \tag{1.16}$$

We see that both the density and the pressure decrease exponentially with increasing altitude. The constant λ defined in Eq. (1.15), which has the dimension of a length, is the drop over which the pressure and the density reduce their value by a factor $1/e$. In the atmosphere, λ is of the order of 7–8 km. Figure 1.9 graphically represents Eqs. (1.5) and (1.16).

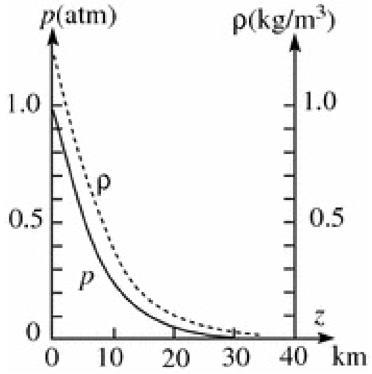


Fig. 1.9 Pressure (*left scale*) and density (*right scale*) in normal atmosphere versus altitude

Communicating vessels is the name given to a set of connected containers open on their upper surface. If a liquid is introduced into any of them, it will occupy all the connected vessels. At equilibrium, if the liquid is homogeneous, all the free surfaces are at the same level, whatever the shape of the vessels.

1.4 Archimedes Principle

The Archimedes principle , established by Archimedes of Syracuse (287 BC–212 BC), states that a body immersed in a fluid experiences an upward directed force equal in magnitude to the weight of the fluid it displaces. It is called buoyancy or upthrust . The buoyant force is present whether the body is completely or partially immersed in the fluid and is applied in the center of mass of the displaced fluid. To demonstrate the principle, let us start by thinking that the body is not present and consider the volume of fluid that will be occupied by the body when we put it back. This volume being at rest, the surface forces balance the volume force. The latter is the weight of the fluid (that will be displaced by the body). It is applied in the center of the fluid mass. If we now put the body back in its place, the surface forces do not change, the surface being the same. Their resultant is the weight of the displaced fluid mass.

If V_i is the immersed volume and ρ_l is the density of the fluid, the buoyancy is

$$\mathbf{F}_A = -V_i\rho_l\mathbf{g}. \tag{1.17}$$

Example E 4.1

An iceberg of density $\rho_i = 900$ kg/m^3 is immersed in seawater (density $\rho_w = 1030$ kg/m^3). Find the emerged fraction.

Calling V_i and V_e the immersed and emerged volume, respectively, the buoyancy is $F_A = \rho_w V_i g$ and the weight $F_w = \rho_i (V_i + V_e) g$. For equilibrium, it should be $\rho_i (V_i + V_e) g = \rho_w V_i g$, or

$$\frac{V_e}{V_i} = \frac{\rho_w - \rho_i}{\rho_i} = \frac{\rho_w}{\rho_i} - 1.$$

The emerged fraction is then

$$\frac{V_e}{V_i + V_e} = \frac{\rho_w - \rho_i}{\rho_i} = 0.14.\blacksquare$$

Consider now the case of a homogeneous body of density ρ_b completely immersed in the liquid. The volumes of the body and of the displaced liquid coincide. Call them V. Both forces are applied in the center of mass of the body. Their moment is zero. The resultant force is

$$\mathbf{F} = V_i(\rho_b - \rho_l)\mathbf{g}. \tag{1.18}$$

which is vertical downward if the body is denser than the liquid, and upward in the opposite case.

1.5 Fluid Equilibrium in the Centrifugal Field

Consider now a homogeneous liquid in a cylindrical container, which rotates about its axis at constant angular velocity ω. After an initial transient, the liquid assumes an equilibrium configuration in which the free surface is a rotation paraboloid, with its axis on the rotation axis.

The problem is conveniently studied in a frame rotating with the cylinder and z-axis on the geometrical axis. The x and y axes are horizontal and rotate with the cylinder. The reference is not inertial and we must include the apparent force, which is the centrifugal force . Both this force and the weight are proportional to the mass, both being volume forces.

The centrifugal force per unit mass is the centrifugal acceleration, which is proportional to the distance from the axis $r' = \left(x^2 + y^2\right)^{1/2}$ and is directed outwards (see Fig. 1.10). Calling $\mathbf{r}'$ the position vector taken from the axis, the centripetal acceleration is then $\mathbf{a}_c = \omega^2 \mathbf{r}'$. Consequently, the field in which the liquid is immersed is

$$\mathbf{G} = \mathbf{g} + \omega^2 \mathbf{r}' = \left(\omega^2 x, \omega^2 y, -g\right). \tag{1.19}$$

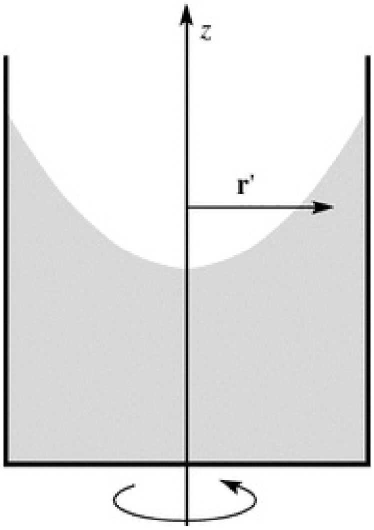


Fig. 1.10 A liquid in a rotating cylinder

The field is conservative and its potential, by integration, is

$$\phi(x,y,z) = -\frac{1}{2}\omega^2\left(x^2 + y^2\right) + gz + \text{const.} \tag{1.20}$$

Under equilibrium conditions, the free surface is an equipotential, ϕ = const, which we can solve for z and write

$$z = \frac{\omega^2}{2g}\left(x^2 + y^2\right) + z_0,$$

where z_0 is a constant to be determined. This is the equation of a round paraboloid; z_0 is the value of z for $x = y = 0$, and hence is the vertex of the paraboloid. It can be calculated knowing the volume of the liquid.

1.6 Viscosity

In Sect. 1.1, we saw that, in general, the surface forces acting on a fluid element have a component normal to the surface, which is the pressure, and one parallel to the surface, which is the shear stress. The latter are zero if the fluid is at rest or if it moves all together, as in the rotation in a stationary regime. Contrastingly, shear stresses are present if parts of the fluid, both liquid and gas, move relative to one another. We shall now begin to study these phenomena.

Consider two fluid elements flowing one over the other, as shown in Fig. 1.11. We call $\mathbf{v}_1$ and $\mathbf{v}_2$ the velocities of the two elements, which are parallel to the contact surface, which we call dS. We consider the case that $\upsilon_1 > \upsilon_2$. Like the friction, the shear stress always acts in opposition to the relative motion of the elements. In other words, the shear stress force direction is always opposite to the relative velocity. In the case we are considering, $\upsilon_1 > \upsilon_2$, the direction of the shear stress $d\mathbf{F}_{12}^{(S)}$ exerted by element 1 on element 2, which is slower, is such as to accelerate it, while the direction of the shear stress $d\mathbf{F}_{21}^{(S)}$ exerted by element 2 on element 1, which is faster, is such as to decelerate it. The two forces are an action-reaction pair and, as such, are equal and opposite.

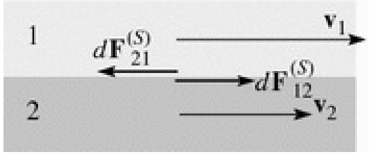


Fig. 1.11 Two fluid elements in relative motion

We already know that the surface forces are proportional to the surface on which they act. In addition, experiments show, as we shall see, that the shear stress is proportional to the derivative of the magnitude of the velocity in the direction **n** normal to the considered surface dS, namely that

$$dF^{(S)} = \eta \frac{\partial v}{\partial n} dS\,, \tag{1.21}$$

where η is an important quantity characteristic of the fluid, which is called *viscosity* and, more specifically, *dynamic viscosity* (to distinguish it from keninematic viscosity, as we shall see in Sect. 1.11). For a given fluid, viscosity depends on temperature and pressure. The physical dimensions of viscosity, such as one sees in Eq. (1.21), are

$$[\eta] = \left[\mathrm{M}\,\mathrm{L}^{-1}\,\mathrm{T}^{-1}\right] = [\mathrm{pT}]\,, \tag{1.22}$$

where p in the last member stands for pressure. The unit of viscosity is kg m^{-1} s^{-1} or, more frequently, Pa s. Notice that in the literature, one can still find an old unit called *poise* , which is 1 g cm^{-1} s^{-1}, hence equal to 0.1 Pa s.

Table 1.1 examples of the viscosity of several liquids and gases. In two relevant cases, water and air, the values at several temperatures are reported. Notice that the viscosity values span more than ten orders of magnitude. Notice also that the viscosity of the liquids usually decreases with increasing temperature, while that of gases increases.

Table 1.1 Viscosity of liquid and gases

Liquid	**Temp. (°C)**	***η*** **(μPa)**	**Gas**	**Temp. (°C)**	***η*** **(μPa)**
Water	0	1.787	Air	−32	15.39
	20	1.002		0	17.08
	40	0.653		18	18.27
	60	0.467		40	19.04
	80	0.355	Argon	20	22.17
	100	0.282	Helium	20	19.41
Methyl alcohol	20	0.597	Hydrogen	20.7	8.76
Glycerin	20	1.49	Neon	20	31.11
Olive oil	20	84	Nitrogen	27.4	17.81
Molasses	109	2.8×10^{6}	Oxygen	19.1	20.18
Pitch	20	10^{10}	Methane	20	10.87
Liquid tin	500	1.2	Water vapor	100	12.55
Liquid zinc	389	1.31		300	20.24

The shear stress can be obtained dividing Eq. (1.21) by the surface element *dS*, which gives

$$\tau = \frac{dF^{(S)}}{dS} = \eta \frac{\partial v}{\partial n}. \tag{1.23}$$

The physical meaning of viscosity can be understood considering the molecular structure of fluids. This will be done in Sect. 6.3.

We now go back to the liquid in a rotating cylinder of the previous section, where we considered its motion in a stationary regime. Suppose now the cylinder containing the liquid to be initially at rest. The surface of the liquid is a horizontal plane. Immediately after the rotation is started, the surface is still plane and the liquid is still at rest, while the surfaces of the cylinder rotate. The viscosity between the walls of the cylinder and the liquid elements accelerates layers of fluid, beginning with those closer to the walls, and gradually reaching the innermost ones. After a transient, the situation becomes stationary. The liquid now rotates as a solid body; there are no relative motions between layers of liquid, and the shear stresses are zero.

The validity of Eq. (1.21) can be verified experimentally with the device shown in Fig. 1.12. It is also used to measure viscosities and is called a *viscometer* . The internal cylinder C_1 of radius r_1 is rather massive and hangs on a torsion wire, bearing an index I to measure the rotation angles. The internal cylinder is contained in an external one, coaxial with it, with C_2 of radius r_2, a bit larger than r_1. Hence, we have $\Delta r \equiv r_2 - r_1 \ll r_1$. The external cylinder is fixed to an axis that can be put into rotation by an engine (not shown in the figure). The interspace between the cylinders is filled with the liquid under study up to the level AA. Let h be the height of the immersed part of the internal cylinder.

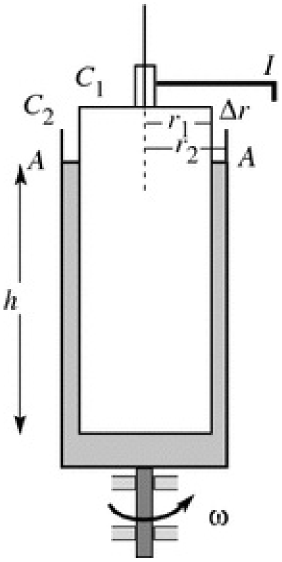


Fig. 1.12 A viscometer

We now put in rotation C_2 with a certain angular velocity ω, corresponding to the velocity $\upsilon = \omega r$ of its wall. We observe the index I moving to a new aquarium position, rotated at an angle α relative to, the position in absence of rotation. The device is a torsion balance; the angle α (and the elastic moment) is proportional to the moment M acting on the cylinder C_1. This is the moment due to viscosity.

We can think of the liquid as being composed of coaxial cylindrical layers. The outermost layer, which is in contact with cylinder C_2, is at rest relative to it, namely it moves with its velocity υ. Similarly, the innermost layer is in contact with cylinder C_1 and is still. The velocity of the intermediate layers varies gradually between these two values, maximum outside, minimum inside.

Equation (1.21) tells us that the viscous force is directly proportional to the area of the two surfaces in relative motion, to the difference of velocity (namely υ, because C_1 is at rest) and inversely to their distance. The proportionality coefficient is the viscosity that we want to measure.

Consider the liquid layer of height h and angular width $d\phi$ (see Fig. 1.13). The internal and external surfaces are, respectively, $dS_1 = hr_1 d\phi$ and $dS_2 = hr_2 d\phi$. They are not exactly equal, but they are close to being so, because the internal and external radiuses are, by construction, almost equal. We can then approximate each area with the mean of the two. If r is the mean of r_1 and r_2, this is $dS = hrd\phi$. The shear force on each of the two surfaces given by Eq. (1.21) is

$$dF = \eta \frac{v}{\Delta r} dS = \eta \frac{v}{\Delta r} rhd\phi.$$

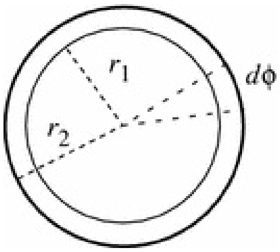


Fig. 1.13 Section of the viscometer of Fig. 1.12

The force is parallel to the relative velocity, and hence is perpendicular to the axis. Its moment about the axis is

$$dM = rdF = \eta \frac{v}{\Delta r} r^2 h d\phi.$$

Integrating we have

$$M = \eta h r^2 2\pi \frac{v}{\Delta r} = \eta r^2 S \frac{\omega}{\Delta r}, \tag{1.24}$$

where, we recall, $S = 2\pi rh$ is the area of the moving surfaces and $\omega = v/r$ is the angular velocity of C_1. We obtain the moment M from the measurement of the rotation angle α of the torsion balance.

We can now check the validity of Eq. (1.21). If we vary the angular velocity ω of the external cylinder, we observe that α varies in proportion. If we change the height h of the liquid, for example, using only half of it, and consequently halving the facing surfaces, we see that α varies in proportion too. We can verify the dependence on the distance Δr by changing the internal cylinder with one of smaller or larger radius. These tests show that Eq. (1.21) holds.

On the other hand, we can use the device to measure the viscosity of a liquid, measuring the moment M and the stationary state angular velocity ω and knowing the other quantities S, r and Δr by construction. This is given by

$$\eta = \frac{M \Delta r}{r^2 S\, \omega}. \tag{1.25}$$

1.7 Incompressible Flow

We now begin the study of *hydrodynamics* , or fluid dynamics. As we know from everyday experience, the motions of fluids can be very different, ranging from rather simple to very complicated. For example, the motion of the water in a river in a stationary regime is simple, while being much more complicated in a mountain creek or waterfall. The motion of air in the wake of an airplane or of a car is extremely complex.

We shall start from the simplest cases or even idealizations. We shall then move on to more realistic situations. In this chapter, we shall limit the discussion to situations in which the density of the fluid can be considered constant, independently, in particular, of pressure. In other words, we shall assume the fluid to be *incompressible* . Its motion is called incompressible flow . This approximation is good for liquids, as is obvious, but can also be applied to the gases in several dynamical problems. As a matter of fact, the volume changes during motion are generally very small, with the exception of instances in which the velocity is close to the speed of sound (340 m/s for air at ambient temperature).

We shall describe the motions in an inertial reference frame. We imagine the fluid divided into physically infinitesimal elements. Let us consider the situation in a certain instant t. In that instant, the velocity of the element in the position (x, y, z) is, say, $\mathbf{v}(x, y, z, t)$. The velocity is a vector function of coordinates and time. In the next immediate instant, we shall find a different fluid element at the same point, moving, in general, at a different speed. A vector function of the coordinates and, possibly, of time is called a *vector field* . If the vector is independent of time, the field is said to be *stationary* . The field we are considering is the *velocity field* . If the velocity field is stationary, and we fix our attention to any point in the fluid, all the different fluid elements that we see going through that point have the same velocity. For example, the water that goes through a section of a river in a stationary regime is always physically different, but its velocity is always the same.

In the first volume of the course, we studied force fields, which are vector fields exactly similar to the velocity field. In that instance, we found it useful to draw the field lines. This is also very useful now. The lines of the velocity field are called *streamlines* . A streamline is a line drawn in the fluid such that its tangent at each point is parallel to the local fluid velocity. The streamlines are infinite in number. At any point, there is only one streamline. The set of all the streamlines at a given instant constitutes the instantaneous *flow pattern*.

Let us examine the procedure for drawing the field lines, say at a certain instant t, to take into account non-stationary situations. We shall obtain a shot of the field at the considered instant. We start from point 1 in Fig. 1.14 and consider the velocity $\mathbf{v}_1$ of the fluid element passing at the considered instant. We make a small step $\delta\mathbf{s}$ in the direction of $\mathbf{v}_1$. We reach point 2. We consider the velocity $\mathbf{v}_2$ of the fluid element passing in 2 at the same instant t and make a small step in its direction, and so on. In this way, we draw a broken line, which, going to the limit of infinitesimal step length, becomes a curve. This is the flow line.

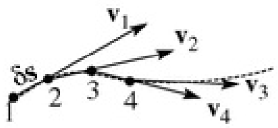


Fig. 1.14 Building a streamline

We can visualize the flow pattern in a liquid with the following artifice. We mix into the fluid a number of small particles, which we can see, and that can remain in suspension. We can use aluminum powder in water, for example. The particles make the fluid elements somewhat visible. We can take pictures and films. If we take a photo with an exposure time Δt, every particle will appear as an oriented segment $\mathbf{v}\,\Delta t$, which is the displacement of the particle in the exposure time. Hence, every oriented segment is proportional to the velocity at that point and has its direction. We obtain a snap shot of the velocity field at the considered instant.

In a non-stationary field, the streamlines change continuously, while in a stationary field, they do not. We shall consider only stationary conditions in the following, up to the point where it is advisable to do the contrary.

A second important concept is the *flow tube* , also called a *stream tube* . Consider any closed curve, like e Γ in Fig. 1.15. A flow tube is the set of the flow lines that pass along the points of this curve. A flow tube of infinitesimal section is called a *flow filament* or *stream filament* .

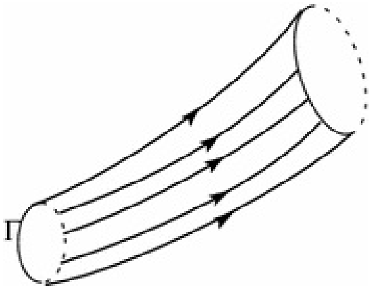


Fig. 1.15 A flow tube

It is often useful to look at the motion of a fluid from a different, complementary point of view. Let us fix our attention on a certain fluid element. We might paint it mentally in red, for example. We mark it by stating that it is the element that passes along the point (x_0, y_0, z_0) at time t_0. We look at its motion. Our element describes a trajectory, which we call the *path line* . At a subsequent instant t_1, we see another element passing at (x_0, y_0, z_0). We mentally paint it blue. We look at it and see its trajectory. In general, the red and blue trajectories may be different. However, if the velocity field is stationary, they are equal. It is evident that in a stationary field, the path lines and the streamlines coincide.

Let us now go back to the flow tube in a stationary regime. Indeed, it behaves exactly like a tube. No flux exits or enters from its lateral walls, even if these are ideal and not physical. Indeed, by definition, the velocity is tangent to its walls and consequently cannot have any normal component. This observation has the following important consequence.

Consider a flow tube having a section small enough to allow for considering the velocity to be equal at all the points of a given normal section (but obviously not in the different ones). Consider two normal sections S_1 and S_2, as in Fig. 1.16. Call $\mathbf{v}_1$ and $\mathbf{v}_2$ the velocities of the fluid in the two sections.

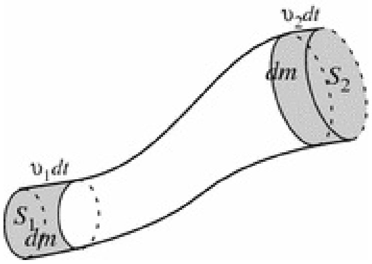


Fig. 1.16 A section of an infinitesimal flow tube

As the mass is conserved and as no mass can go through the lateral surface, the mass that crosses the section S_1 in any time interval dt must be equal to the mass crossing S_2 in the same time interval. Call it dm. The mass that crosses S_1 in dt is the mass contained in the volume having S_1 as the base and $\upsilon_1\, dt$ as height, hence, $dm = \rho\upsilon_1\, dt$. Similarly, through S_2, it is $dm = \rho\upsilon_2\, dt$. Hence,

$$\rho S_1 \upsilon_1 dt = \rho S_2 \upsilon_2 dt. \tag{1.26}$$

Considering that this equality holds for every pair of sections of the tube, the quantity $Q_m = \rho S\upsilon$ is constant on all the sections of the tube. Its physical meaning is to be the fluid mass crossing any section in one second and is called the *mass flow rate*

$$Q_m = \rho S \upsilon. \tag{1.27}$$

When, as we are assuming, the density is constant, the quantity

$$Q_V = S\upsilon. \tag{1.28}$$

is constant too. It is the fluid volume crossing the section per second, and is called the *volumetric flow rate* .

A consequence of the invariance of the flow rate along a flow tube is that, if the section of the tube shrinks, the velocity increases, as shown in Fig. 1.17. Graphically, this implies that the flow lines are denser where the velocity is larger.

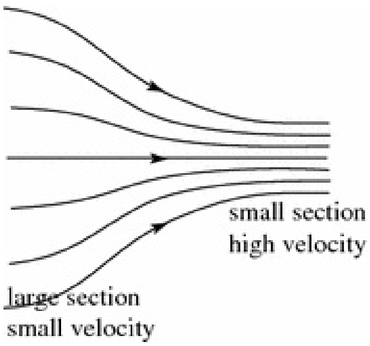


Fig. 1.17 Varying section flow tube

1.8 Bernoulli Theorem

In this section, and in the next one, we shall consider the dynamics of the *ideal fluid* . This is defined as having constant density and zero viscosity, in other words, being incompressible and inviscid. Obviously, ideal fluids do not exist. However, real fluids can often be considered as almost ideal ones.

We shall consider the motion of an ideal fluid in the field of weight force. In absence of viscosity, the forces are conservative, and we can use the mechanical energy conservation principle. We shall find an important result established in 1738 by Daniel Bernoulli (1700–1782), known as the *Bernoulli theorem* . We shall see a few examples in the next section.

The Bernoulli theorem is based on the assumptions that the fluid is ideal and the regime is stationary.

Figure 1.18 represents a portion of a flow filament between two sections *AA* and *BB*. The *z*-axis is vertical upward. The first section has area dS_1 (which is infinitesimal). Its height is z_1. The pressure at that point is p_1 and the fluid velocity is υ_1. Similarly, the second section has area dS_2; the height is z_2, the pressure is p_2 and the fluid velocity is υ_2. Let us consider the mass of fluid laying between the two sections at the instant *t* and call it Δ*m*. Soon after, at the instant *t* + *dt*, the mass Δ*m* has moved and is now between the two sections *A*′*A*′ and *B*′*B*′. The distance between *AA* and *A*′*A*′ is obviously $\upsilon_1\,dt$, and the distance between *BB* and *B*′*B*′ is $\upsilon_2\,dt$. The mass *dm* that crosses both sections in *dt* is the same, and consequently, as we already saw in the preceding section,

$$dm = \rho dS_1 \upsilon_1 dt = \rho dS_2 \upsilon_2 dt. \tag{1.29}$$

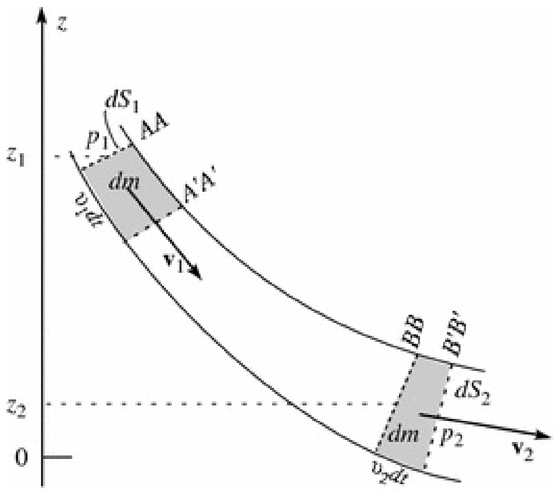


Fig. 1.18 Fluid motion in a flow filament

It will be useful to observe that the two volumes are also equal, given that the density is constant. Namely

$$dV = dS_1 v_1 dt = dS_2 v_2 dt. \tag{1.30}$$

We now apply the kinetic energy theorem to the motion of the mass Δm from the first section to the second. The mass moves under the action of the following forces:

- the weight, which is a conservative volume force
- the pressure forces on dS_1 and dS_2, which are normal to those surfaces and consequently parallel to the displacement
- the pressure forces on the lateral surface of the filament, which are normal to the displacement (because the fluid is non-viscous) and consequently do not do work.

The work, say dW_g, done by the weight is easily calculated observing that the mass of fluid between the sections *A′A′* and *BB* is the same before and after the displacement. All goes as if the mass *dm* had moved from position z_1 to the position z_2. Hence, $dW_g = g\,(z_1 - z_2)\,dm.$

The work of the pressure forces on the section dS_1 is equal to the magnitude of the force, $p_1\,dS_1$, times the displacement $\upsilon_1\,dt$. Force and displacement are in the same direction. This is similar for the work on section dS_2, taking into account that here force and displacement have opposite directions. Hence, we have

$$dW_p = p_1 dS_1 v_1 dt - p_2 dS_2 v_2 dt.$$

The sum of the works is equal to the variation of the kinetic energy. To find the latter, we can again imagine that everything goes as if the mass *dm* would have changed velocity from υ_1 to υ_2. Hence,

$$dU_{K2} - dU_{K1} = \frac{1}{2}v_2^2 dm - \frac{1}{2}v_1^2 dm.$$

We simplify the three expressions we just found, all of them having the volume *dV* occupied by $dm = \rho dV$ and using Eq. (1.30), obtaining

$$dW_g = \rho g z_1 dV - \rho g z_2 dV$$
$$dW_p = p_1 dV - p_2 dV$$
$$dU_{K2} - dU_{K1} = \frac{1}{2}\rho v_2^2 dV - \frac{1}{2}\rho v_1^2 dV.$$

The kinetic energy theorem states that

$$dU_{K2} - dU_{K1} = dW_g + dW_p.$$

Substituting the above expression and simplifying *dV* everywhere, we have

$$\frac{1}{2}\rho v_2^2 - \frac{1}{2}\rho v_1^2 = \rho g z_1 - \rho g z_2 + p_1 - p_2$$

and, then, moving the quantities relative to the same section to the same side

$$p_1 + \rho g z_1 + \frac{1}{2}\rho v_1^2 = p_2 + \rho g z_2 + \frac{1}{2}\rho v_2^2. \tag{1.31}$$

Finally, considering that the two sections in question are arbitrary, we can state that in all the sections of a fluid filament

$$p + \rho g z + \frac{1}{2}\rho v^2 = \text{constant}. \tag{1.32}$$

In other words, the Bernoulli theorem states that, in a stream filament of an ideal fluid in a stationary regime, the sum of the pressure, the potential energy per unit mass ($\rho g z$) and the kinetic energy per unit mass ($\rho\upsilon^2/2$), is constant. Notice, however, that this sum may be different in different stream filaments of the same fluid.

1.9 Applications of the Bernoulli Theorem

In this section, we shall study several applications of the Bernoulli theorem.

The Torricelli theorem . Figure 1.19 shows a tank of section *S* containing an ideal liquid. On the lower part of its wall, there is a hole *O*, whose section is small compared to *S*. The height of the free surface above the hole is *h*. The liquid exits from the hole with velocity υ, which is a function of *h*. The atmospheric pressure p_a acts on the free surfaces of the liquid, both on the horizontal one inside the tank and on the jet outside *O*. The velocity field can be considered stationary. The velocity of the fluid elements near the upper surface is small relative to υ. Consequently, the flow lines are sparse in the upper part, becoming denser close to the hole (see Fig. 1.19). We apply the Bernoulli theorem to any of these flow lines between any points on the upper free surface, like *A* in the figure, and *O*. We take the origin of the heights in *O* and write

$$p_a + \rho g h = p_a + \frac{1}{2}\rho \upsilon^2$$

or

$$\upsilon = \sqrt{2gh}. \tag{1.33}$$

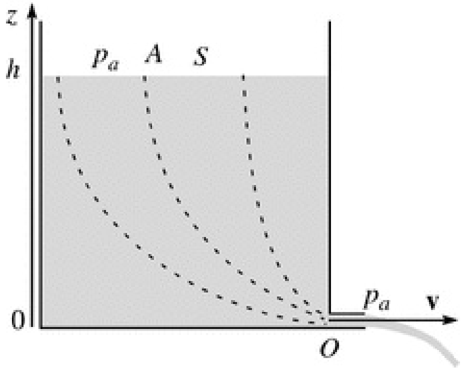


Fig. 1.19 The Torricelli theorem

Notice that the velocity of outflow is independent of the density of the liquid, being equal to the velocity of a body free falling along the drop *h* from rest. This is obviously a consequence of energy conservation in absence of dissipative forces, which we have neglected.

One might think to calculate the rate of liquid outflow simply as the product of the velocity in the jet and the area of the hole. This is not true, because the velocities of the fluid elements in the jet have a component inwards toward the axis of the stream. Consequently, the section of the jet decreases in the initial part. After that, the velocities become parallel and the jet section becomes constant. The distance at which the contraction ceases, and the ratio between the jet cross-section there and the area of the hole, known as the *contraction coefficient* , depends on the shape of the discharge tube. The product of the area of this section and the velocity gives the rate of outflow.

The Venturi effect was discovered by Giovanni Battista Venturi (1746–1822). The effect offers a practical method to measure the average speed of a gas or a liquid in a duct. It is based on the measurement of the pressure change due to a change in the diameter of the duct. To this aim, one inserts into the duct a segment, as shown in Fig. 1.20, called a Venturi tube. The standard section S_1 of the pipe is reduced to S_2 and then brought back to the initial value. The shape is designed so as not to alter the flow in the duct. A manometer measures the difference between the pressures p_1 and p_2 at the two sections.

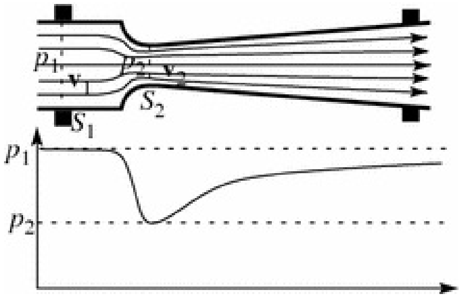


Fig. 1.20 The Venturi tube

Considering a horizontal tube, there is no contribution of the weight force, and the Bernoulli theorem gives

$$p_1 + \frac{1}{2}\rho v_1^2 = p_2 + \frac{1}{2}\rho v_2^2$$

or

$$p_1 - p_2 = \frac{\rho}{2}\left(v_2^2 - v_1^2\right).$$

The velocity increases when the section of the pipe shrinks from S_1 to S_2, because the flow is constant, as $v_2 = v_1 S_1/S_2$. We use this relation to eliminate υ_2, obtaining

$$v_1 = \sqrt{\frac{1}{1 - S_1^2/S_2^2}}\sqrt{\frac{2(p_1 - p_2)}{\rho}}. \tag{1.34}$$

This relation gives the velocity υ_1 of the unperturbed fluid from the measurement of the pressure drop p_1–p_2, being the sections and the fluid density known. To be precise, the velocity of the fluid is somewhat different at different points of the sections, and we want to measure the mean values of υ_1 and υ_2. Consequently, Eq. (1.34) is not exact and, in practice, the instrument must be calibrated.

As a matter of fact, the Venturi effect is more general. It is summarized in

Fig. 1.21. When the section of a duct decreases, the velocity of the fluid increases. This is a consequence of the conservation of the mass. The Bernoulli theorem allows us to add that, if the velocity increases, the pressure decreases. This is a consequence of energy conservation.

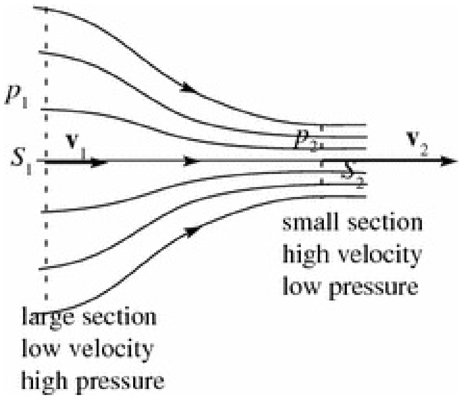


Fig. 1.21 The Venturi effect

The Pitot tube , invented by Henri Pitot (1695–1771), is used to measure the velocity of a fluid current, or of an object moving in a fluid, like an airplane. It is shown in Fig. 1.22. It consists of a thin tube parallel to the current flow.

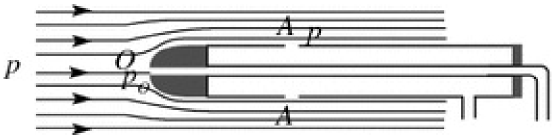


Fig. 1.22 The Pitot-Prandl tube

The flow lines, not altered upstream of the tube, open up when they meet its rounded head and then lap the lateral surface. In the center of the head, *O*, there is a hole that communicates through a narrow tube with one of the inputs of a differential manometer (an instrument, not shown in the figure, that measures the difference between the pressures at its two inputs). The flow line on the axis ends in *O*, where the velocity reduces to zero. Such a point is called a *stagnation point* .

In *A*, which is sufficiently far from *O* on the side of the tube, the fluid moves practically, at unaltered speed. This is consequently the velocity we want to measure; let us call it υ. The pressure is also the undisturbed one p. A few small holes are drilled into the sides of the tube. In this way, the pressure inside the large tube is also p. It is connected to the second input of the differential manometer.

We now apply the Bernoulli theorem to the fluid flow filament in *O*, between point *O* and any point upstream where the fluid is undisturbed. There, the pressure is *p* and the velocity *υ*. Let p_O be the pressure in *O*. The Bernoulli theorem then gives

$$p + \frac{1}{2}\rho v^2 = p_0.$$

As we said, we measure the pressure difference, that is

$$p_0 - p = \frac{1}{2}\rho v^2,$$

which is called the stop pressure, from which we find the velocity

$$v = \sqrt{\frac{2(p_O - p)}{\rho}}. \tag{1.35}$$

The device, as described, was the product of Ludwig Prandtl (1875–1953). It is commonly used to measure the velocity of fluid currents or of objects moving in a fluid. Notice that, for example, in the case of the plane, the measured velocity is relative to the air, not to the ground. The effects of viscosity are, in general, negligible.

Hydrodynamic paradox . We conclude with a final example, which, in contrast with the previous ones, is only a curiosity. Figure 1.23 shows the section of a device made of two parts. The upper part is a tube to which a perforated disk is attached. The tube is connected to a pump blowing air at speed **v**. The lower part is another disk, equal to the first one and kept parallel at a close distance from it. The disks are horizontal. One would think that the air blowing should push away the lower disk. However, just the opposite happens. The lower disk is attracted and reaches an equilibrium position at a certain distance from the upper one.

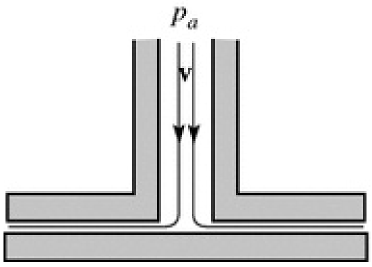


Fig. 1.23 Hydrodynamic paradox

The explanation is simple. The high-speed air jet that comes out of the tube expands between the two disks. The fluid elements follow radial trajectories towards the periphery. Their velocity decreases, becoming practically zero on the rim of the disks. Now, we apply the Bernoulli theorem to two points of a radial flow filament, one on the axis, one on the rim. Call p the pressure and υ the velocity at the first point. At the second point, the pressure is the atmospheric one p_a and the velocity is zero. We have

$$p + \frac{1}{2}\rho\upsilon^2 = p_a.$$

Hence, $p \ll p_a$. The pressure difference attracts the disk, and the corresponding force equilibrates the weight (provided this is not too large).

1.10 D’Alembert Paradox

Consider a solid sphere fully immersed in a perfect fluid flowing with horizontal velocity **v**. The diameter of the sphere is small compared to the extension of the current. The velocity field is stationary and, where not perturbed by the sphere, uniform. Figure 1.24 shows the flow pattern around the sphere. Suppose the weight to be balanced by the buoyancy , so that the sphere is in equilibrium in the vertical direction.

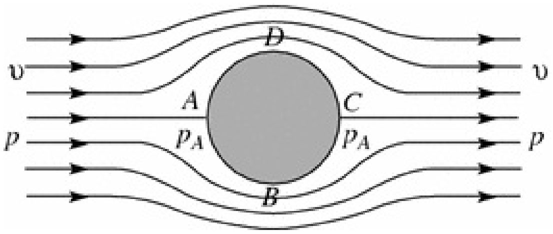


Fig. 1.24 Streamlines around a sphere in an inviscous and incompressible fluid

The streamlines that are parallel and equidistant upstream and downstream then open and close again symmetrically around the body. The middle streamline upstream terminates at the point A. The fluid stops at this point, which is the forward stagnation point . Let p_A be the pressure in A and p the pressure upstream in the undisturbed region along the same streamline . For the Bernoulli theorem, we have

$$p_A = p + \frac{1}{2}\rho\upsilon^2. \tag{1.36}$$

Namely, the pressure in A is larger than in the unperturbed flow by $\rho\upsilon^2/2$.

The overpressure produces a force pushing the body in the direction of the current. However, we must also look at the downstream side. At point *C*, which is symmetric to *A* and is the backward stagnation point, the pressure is, say, p_C. Downstream, in the unperturbed region along the same streamline, the pressure is *p*. The Bernoulli theorem applied along this streamline gives

$$p_C = p + \frac{1}{2}\rho v^2. \tag{1.37}$$

We see that p_C is equal to p_A. Namely, the pressure forces on the middle plane balance each other. The resultant is zero. We should also consider, however, the effects of the pressure forces all around the body. Let us fix our attention on one streamline. Being the flow stationary, this is also a path line, namely the trajectory of the fluid elements. Being the trajectory curved, the fluid element accelerates. The acceleration is due to a force exerted by the body. The force is normal to the surface because there is no viscosity. The fluid elements exert a force on the body that is equal and opposite, for the action and reaction law. In conclusion, a force normal to the surface acts at every point of the surface of the body. Both the magnitude of the force and whether it is directed towards or outside the surface depend on the flow pattern. Clearly, the resultant pressure force in the forward part of the body *BAD* is in the direction of the current, and the resultant of the backward part *BCD* is opposite to the current. In the present case, the flow pattern is perfectly symmetric, and consequently, the magnitudes of the two resultants are equal; they balance each other. The sphere does not move.

The argument we have just developed was developed in 1752 by Jean le Rond D’Alembert (1717–1783). To be sure, he proved the theorem for the equivalent case of a sphere moving in a fluid at rest. He proved that the drag force on a symmetric body moving at constant velocity in an inviscous and incompressible fluid is zero. The conclusion looks to be in conflict with evidence and is known as the D’Alembert paradox . We know that all solid bodies immersed in a current are subject to a drag force, which can be larger or smaller, but is never zero. This is due to viscosity, which is neglected in the D’Alembert theorem, but is always present. Viscosity has two consequences. First, the forces exerted by the fluid elements lapping the body have components tangent to the surface. The tangent components are in the direction of the current, both in the forward and in the backward parts of the body. Second, the flow pattern becomes forward backwards asymmetric, with the effect that the pressure forces do not balance any more. We shall study these phenomena in Sects. 1.13 and 1.14.

1.11 Laminar Viscous Flow

We shall now study fluid motions in the presence of viscosity. We start by considering a geometrically simple case. An incompressible fluid of viscosity η is included between two parallel horizontal plates, *AA* and *BB* in Fig. 1.25. We keep *BB* still and have *AA* moving with velocity $\mathbf{v}_0$ parallel to its plane. The fluid layer in immediate contact with *AA* remains adherent to the plate and moves with its constant velocity $\mathbf{v}_0$. Similarly, the layer in contact with *BB* has zero velocity. We can imagine the rest of the fluid divided into parallel layers, the velocities of which vary continuously from zero to $\mathbf{v}_0$. Under these conditions, the motion is said to be *laminar* . The situation is realized, for example, when two parallel metallic surfaces slide one over the other, being separated by a lubricant. Another example of laminar flow is in the viscometer we considered in Sect. 1.6. In this case, the fluid layers are concentric cylinders, which rotate about the axis.

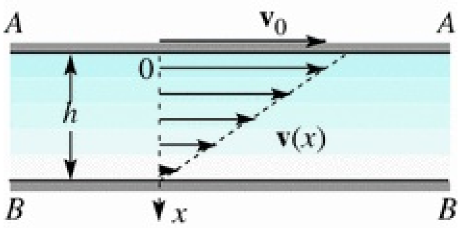


Fig. 1.25 Laminar flow between two plane surfaces

The geometry we have just considered is simple to describe but difficult to put into practice. The condition most usually met, and that which we shall now study, is the flow of a fluid within a cylindrical duct, of radius R. In this case, we can imagine the fluid divided in cylindrical layers, moving in the direction of the axis, each at a possibly different speed. In particular, the layer in contact with the wall is at rest. The velocity of the layers increases while moving inside, and is at a maximum at the axis. The flow we are considering is laminar.

As a matter of fact, the laminar flow is not the only possible one. It is the flow occurring for velocities below certain limits, which depend on the geometrical dimensions of the solid bodies immersed in or limiting the flow. As we shall see in the following sections, when the velocity becomes higher than these limits, the regular layers' stratification is destroyed by the formation of vortices, which mix up the fluid. This regime is called a *turbulent flow* . In this section, we shall study the laminar plane and cylindrical flows.

The plane laminar flow is the simplest geometrically. With reference to Fig. 1.25, we call x the distance from the plate AA. Consider two fluid elements, namely two infinitesimal portions of two adjacent layers of infinitesimal area dS. They exert one on the other equal and opposite forces directed to slow down the relative motion, as we discussed in Sect. 1.6. The force per unit surface, the shear stress , is, for Eq. (1.23)

$$\tau = \frac{dF^{(S)}}{dS} = \eta\frac{\partial \upsilon}{dx}. \qquad (1.38)$$

The symmetry of the problem requires the stress to be the same in all the separation surfaces between the layers, in other words, τ to be independent of x. Then, for Eq. (1.38), the gradient of the velocity should be independent of x too, namely $\partial \upsilon / dx = \text{constant}$. Through integration, we have $\upsilon = ax + b$, where a and b are the integration constants. We find them imposing the boundary conditions $\upsilon = 0$ for $x = 0$ and $\upsilon = \upsilon_0$ for $x = h$. In conclusion, we have

$$\upsilon = \frac{x}{h}\upsilon_0. \qquad (1.39)$$

The velocity varies linearly with the distance from the plate, as shown in Fig. 1.25.

We shall now consider the flow in a cylindrical duct. From the historical point of view, we observe that, even if the existence of two types of flow, laminar and turbulent, was known, the first precision experiments were performed by the German engineer Gotthilf H.L. Hagen (1797–1884), who published his results in 1839. The same results were independently obtained by the French physicist and physiologist Jean L.M. Poiseulle (1797–1869), who published in 1840.

Consider the motion of a fluid in a horizontal cylindrical duct, of circular section of radius R, in a laminar regime. Consider a length l of the tube and the pressures p_1 and p_2 at the two extremes. The fluid moves under the action of the pressure difference $\Delta p = p_1 - p_2$, called the *pressure loss* . The fluid element's velocity **v** being parallel to the axis of the tube, its magnitude υ is a function of the distance r from the axis. We call υ_m its mean value. In practice, one measures the volumetric flow rate Q_V, which is the volume of fluid going through a section in a second. The mean velocity is the volumetric flow rate divided by the section.

Hagen used copper tubes a few meters long and with diameters of a few millimeters. Poiseuille, who was interested in blood flow through the capillary veins, experimented with smaller diameters, of a few tenths of a millimeter (that is, the diameter of a capillary vein). Both authors established that the volumetric flow rate is directly proportional to the pressure drop and to the fourth power of the radius and inversely to the length of the tube. The Hagen-Poiseuille law is

$$Q_V = \frac{\Delta p}{l}\frac{\pi}{8\eta}R^4 \tag{1.40}$$

where η is the fluid viscosity . Intuitively, one might expect Q_V to be proportional to the tube section, namely to the second power of the radius. We shall understand the reason for the fourth power with the following analysis.

As we have already stated, we can consider the fluid to be divided in coaxial layers, moving with different velocities. As a consequence of the symmetry of the problem, the magnitude of the velocity is a function of the distance from the axis r alone. This is the function, call it $\upsilon(r)$, we now want to find. We can write Eq. (1.23) for the shear stress as

$$\tau = \frac{dF^{(S)}}{dS} = \eta\frac{\partial v(r)}{dr}. \tag{1.41}$$

Consider the fluid volume in the cylinder coaxial with the tube of radius r and length l. The force acting on its surface has magnitude equal to the shear stress τ, Eq. (1.41), times its surface, $2\pi rld$, and direction parallel and opposite to the velocity. As the regime is stationary, the velocity is constant and the resultant force must be zero. The equal and opposite force is due to the pressure difference Δp between the two faces of the cylinder. Its magnitude is Δp times the area of a face πr^2. We can write

$$\pi r^2\Delta p = -2\pi r l\eta\frac{\partial v(r)}{\partial r}$$

or

$$\frac{\partial v}{\partial r} = -\frac{\Delta p}{2l\eta}r.$$

Through integration, we get

$$v = -\frac{\Delta p}{4l\eta}r^2 + \text{const.}$$

We determine the integration constant from the boundary condition that the velocity is zero on the surface of the tube, namely $\upsilon(R) = 0$. Finally, we have

$$v(r) = \frac{\Delta p}{4l\eta}\left(R^2 - r^2\right). \tag{1.42}$$

We have thus found that the velocity field has a parabolic profile, as shown in Fig. 1.26. The velocity varies from zero on the surface to a maximum value on the axis, equal to $\frac{\Delta p}{4l\eta}R^2$.

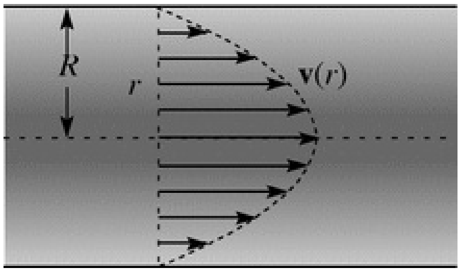


Fig. 1.26 The velocity field for a laminar flow in a cylindrical duct

Let us now determine the volumetric flow rate Q_V. We must take into account the dependence of the velocity on the distance from the axis, which we have just found. Consider the circular zone with its center at the axis and radiuses r and $r + dr$. Its area is $2\pi r dr$. The volume of fluid crossing this area in one second is the volume of a cylindrical annulus with that area as the basis and height equal to the velocity of the fluid at r, namely $dV = \upsilon(r)2\pi r dr$. Using Eq. (1.42) for $\upsilon(r)$, we have

$$dV = \frac{\pi\Delta p}{2l\eta}\left(R^2 - r^2\right) r dr.$$

The volumetric flow rate is the integral of this expression between 0 and R, namely

$$Q_V = \frac{\pi\Delta p}{2l\eta}\int_0^R\left(R^2 - r^2\right) r dr = \frac{\pi\Delta p}{2l\eta}\int_0^R\left(R^2 - r^2\right) dr^2.$$

The integral on the right hand side is immediately done, giving Eq. (1.40), the Hagen-Poiseuille law . We now understand that the reason why the volumetric flow rate is proportional to the fourth power of the radius of the duct is the parabolic profile of the velocity field. Increasing the radius, the flow rate increases faster than the section, because the central portion, in which the velocities are larger, becomes a larger fraction of the total.

We now express the mean velocity υ_m. The flow rate being the product of υ_m and the area of section πR^2, we have

$$\upsilon_m = \frac{\Delta p}{l}\frac{1}{8\eta}R^2, \tag{1.43}$$

which tells us, in particular, that the mean fluid velocity is proportional to the pressure gradient $\Delta p / l$.

A simple demonstration of the pressure drop in a fluid moving through a horizontal tube is shown in Fig. 1.27. If we take into account that Q_V is the same at all the points of the tube, the Hagen-Poiseuille law tells us that the pressure decreases along the tube proportionally to the distance. The device is a glass tank full of water (colored to be easily seen) up to the height h. A horizontal tube is connected to the tank near its bottom. A few vertical tubes at equal distances allow for visual evaluation of the pressure along the horizontal tube, from the heights of the water columns (h_1, h_2, h_3). The pressure at the bottom of the tank is $p_a + \rho g h$ (p_a is the atmospheric pressure) and p_a at the end of the tube. One observes that the pressure decreases linearly along the tube. To be precise, there is always a small pressure drop between the bottom of the tank and the beginning of the tube. For this reason, the line joining the heights of the columns extrapolates a bit below the free surface in the tank.

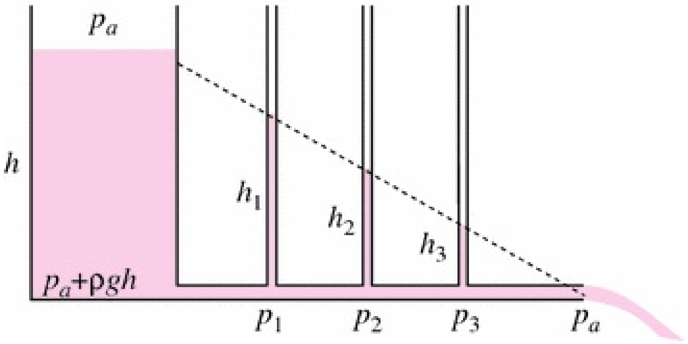


Fig. 1.27 Demonstration of the linear pressure drop foreseen by the Hagen-Poiseuille law

It is sometimes interesting to know the mass crossing a section of the duct per unit time, called mass flow rate Q_m. This is simply the volumetric flow rate multiplied by the density ρ, because this is constant

$$Q_m = \frac{\Delta p}{l}\rho\frac{\pi}{8\eta}R^4. \tag{1.44}$$

This expression and the Hagen-Poiseuille law contain quantities depending on the geometry of the tube and on the applied pressure drop, which are the same for both expressions, a quantity that depends on the fluid, which is the viscosity η for Q_V, the ratio η/ρ, for Q_m. The latter enters into several fluid dynamics expressions and is called kinematic viscosity or relative viscosity

$$\nu = \frac{\eta}{\rho}. \tag{1.45}$$

To avoid confusion when necessary, η is called dynamic viscosity . The units of the kinematic viscosity are the m^{2}s. In terms of kinematic viscosity, Eq. (1.45) is obviously

$$Q_m = \frac{\Delta p}{l}\frac{\pi}{8\nu}R^4 \tag{1.46}$$

1.12 Turbulent Flow. Reynolds Number

The flow of a fluid in a pipe is laminar, as we have considered so far, only if the fluid velocity is small enough. If the velocity exceeds certain limits, which we are going to discuss, the flow becomes turbulent. We begin by discussing a simple experiment. Figure 1.28 shows a tank containing a liquid, say water, connected at its lower portion to a horizontal tube, through which we have the water running. A second vessel, at a higher position, also contains water, colored to distinguish it. Its bottom is connected to a tube of small cross-section, which is used to inject an axial flow into the principal water current from the bigger tank. A tap R can be used to reduce or increase the flow rate. All the parts are made of glass to allow for easy observation.

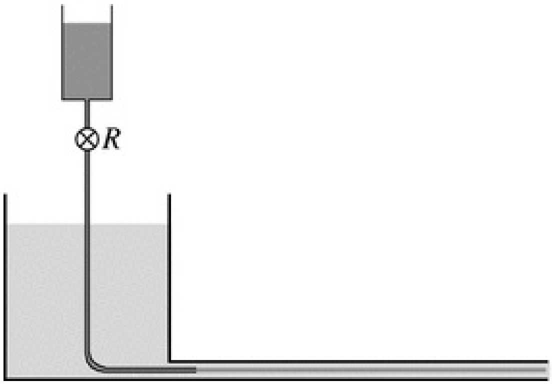


Fig. 1.28 Experiment to observe the laminar to turbulent flow transition

Initially, R is closed. If we open it a bit, we observe a colored water filament flowing at the axis of the horizontal tube. It keeps its identity, without mixing with the main current, even if the tube is rather long. If we gradually open the tap further, the velocity of the colored filament continually increases, becoming larger and larger than the speed of the main current, with which it is in contact. Still, the filament maintains its identity. The flow is laminar. The situation is stable under these conditions; if, for example, we give a shock to the device, creating spurious motions in the two fluids and causing them to mix, we observe that, after a while, the flow readjusts automatically to the previous condition; the two flows separate.

However, when the relative velocity of the two fluids reaches a well-determined value (we can control the stability of this value with repeated experiments), the regime changes. The first change is that the colored filament is no longer straight; it develops oscillations, which are more or less periodic and advance with the filament. The regime is no longer stationary. If the relative velocity increases further, the oscillations become faster and more chaotic, then vortices develop and the two fluids get mixed. The regime becomes turbulent . In the turbulent flow, the velocity of the fluid elements varies in an irregular and chaotic way. In the description of this motion, we shall use a mean velocity, mediated over periods long enough to smooth the chaotic fluctuations.

As opposed to that of the laminar flow, the description of the turbulent flow presents enormous mathematical difficulties, which cannot be handled with analytical methods, even in the simplest cases. The flow patterns in several relevant situations can be found with numerical computations using very powerful computers. Even so, the methods based on the physical dimensions of the parameters are important and we shall exploit them in the following.

Consider once more the flow through a horizontal tube. Figure 1.26 shows the velocity field in the laminar regime. To characterize the transition to the turbulent flow, we shall develop simple arguments based on the physical dimensions of the quantities of the problem. Clearly, the flow regime depends on the characteristics of the fluid and the duct, and on the relative velocity.

The fluids have two physical properties, the density ρ and the viscosity η. The relevant quantity of the tube is its diameter D, while the relative velocity is the mean velocity υ of the fluid (its mean value, as stated above). Hence, we must deal with four physical quantities with the dimensions

$$[\rho] = \mathrm{kg\,m^{-3}}, \quad [\eta] = \mathrm{kg\,m^{-1}\,s^{-1}}, \quad [\upsilon] = \mathrm{m\,s^{-1}}, \quad [D] = \mathrm{m}. \tag{1.47}$$

We observe that the quantities that characterize the regime, for example, the transition from a linear to a turbulent flow, are complicated functions of our four quantities. Consequently, their arguments must be pure numbers. Let us thus search a dimensionless combination of the four quantities. First, we observe that the dimension of the mass can be eliminated only by taking the ratio between viscosity and density, namely the kinematic viscosity $\nu = \eta/\rho$. Its dimensions are $[\nu] = \mathrm{m^2\,s^{-1}}$. Once more, we have only one way to eliminate the dimension of time, namely dividing υ by ν: $[\upsilon/\nu] = \mathrm{m^{-1}}$. Finally, we eliminate the length, multiplying by the diameter *D*. We have thus found the unique dimensionless combination of the four quantities

$$Re = \frac{vD}{\nu} = \frac{\rho v D}{\eta}, \tag{1.48}$$

which is called the *Reynolds number* , after Osborne Reynolds (1842–1912).

Our simple arguments cannot tell us the values of the Reynolds number at which the flow regimes change. These values must be determined experimentally, as we shall soon see. The above arguments, however, tell us that, if we consider different fluids, in cylindrical ducts of different sections and with different velocities, the flow regime is the same if the Reynolds number has the same value. Consequently, for example, the velocity of a regime change is inversely proportional to the radius of the tube; it is twice as large in a tube twice as thin. That velocity is, in addition, proportional to the viscosity; thicker fluids maintain the laminar flow up to larger velocities.

As we have seen, in the laminar flow of a fluid in a tube, the velocity profile is parabolic, as in Fig. 1.26. The layer in contact with the wall adheres to it and is at rest; the velocity gradually increases, moving inwards and becoming maximized at the axis. In the turbulent flow, the layer in contact with the wall is still at rest, but the mixing due to the vortices and chaotic motions is a much more efficient mechanism for increasing the mean velocity receding from the walls. Consequently, the mean velocity profile is as shown in Fig. 1.29. The mean velocity is independent of the distance from the axis in the largest section of the tube. It goes to zero in a thin layer, called the boundary layer , which has a thickness δ and is represented in a shaded tint in Fig. 1.29. The thickness of the boundary layer is always much smaller than the tube radius and, as we shall see in Sect. 1.14, decreases with an increasing Reynolds number.

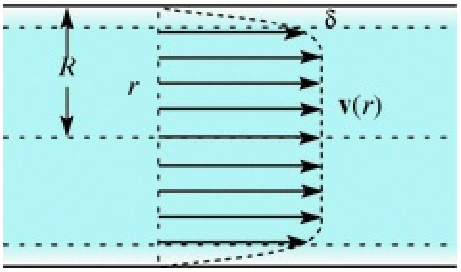


Fig. 1.29 The mean velocity field for a turbulent flow in a cylindrical duct. Boundary layer is the *shaded area*

An important consequence of what we have just stated is that the relation between pressure gradient $\Delta p/l$ and the mean velocity of the fluid is independent of viscosity, as opposed to the laminar flow. Once more, the argument, which follows, is based on the dimensions of the physical quantities.

The physical dimensions of the pressure gradient are $[\Delta p/l] = \mathrm{kg\,m^{-2}\,s^{-2}}$. We must find a combination of the other physical quantities having the same dimensions. We can do that with the density, mean velocity and diameter of the tube. We do not need the viscosity. Their unique combination with the right dimensions is $\rho v^2/D$. It is standard to divide it by two, and we shall use $\frac{1}{2}\rho v^2/D$ (that is, the kinetic energy per unit mass divided by the tube diameter). We can conclude that

$$\frac{\Delta p}{l} = f(Re)\frac{\rho v^2}{2D} \tag{1.49}$$

where *f* is a dimensionless coefficient, called the *Darcy friction factor* after Henry Darcy (1803–1858). The friction factor is a function of the unique dimensionless quantity of the problem, namely the Reynolds number.

Written explicitly, the Darcy friction factor defined by Eq. (1.49) is

$$f(Re) = \frac{2D\Delta p}{l\rho v^2}. \tag{1.50}$$

We notice here that, even if the viscosity does not have an effect on the relation between the pressure gradient and flow velocity, its effects are relevant within the boundary layer. Within this layer, in fact, the velocity varies very rapidly, and consequently, the shear stresses are sizeable.

In the laminar flow, when the Hagen-Poiseuille law holds, the friction factor is inversely proportional to the Reynolds number. If we substitute Eq. (1.42) for the flow velocity and Eq. (1.48) for the Reynolds number in Eq. (1.50) we have

$$f(Re) = \frac{64}{Re} \tag{1.51}$$

a very simple expression indeed.

We shall now discuss the main characteristics of the friction factor as a function of the Reynolds number. These are the result of a series of experiments starting with those conducted by G. Hagen between 1839 and 1869.

Figure 1.30 shows the behavior of the friction factor in cylindrical tubes with two types of internal surface. The continuous curve is for a smooth surface, the conditions under which the above discussion applies, the dotted curve is for rough surfaces, like in a concrete duct, with a particular value of roughness, taken as an example. Notice that both scales are logarithmic.

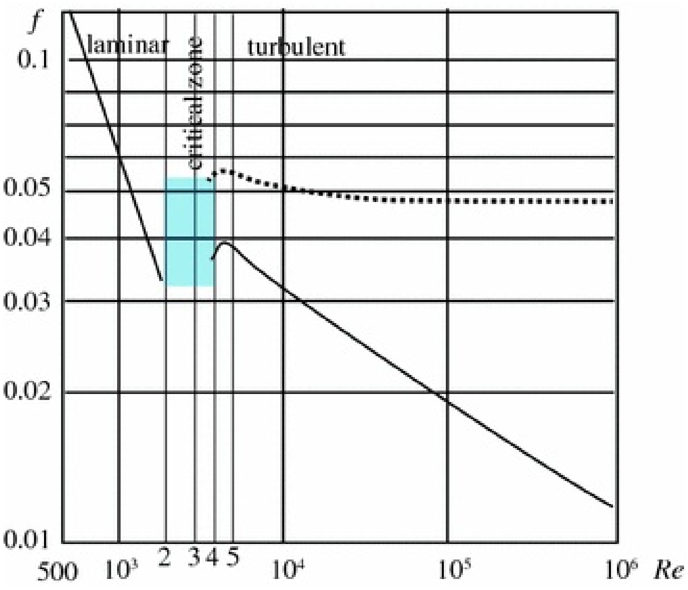


Fig. 1.30 Darcy friction factor versus Reynolds number. *Continuous curve* Smooth tube. *Dotted curve* Rough tube with $\varepsilon/D = 0.02$

Firstly, notice that a constant friction factor, which, graphically, is a horizontal line in the diagram, would mean a pressure drop proportional to the square of the flow velocity. When the Hagen-Poiseuille law holds, the friction coefficient is inversely proportional to the Reynolds number, Eq. (1.51). This function is represented in the log-log diagram by a straight line of slope −1. Experiments show that this is really the case for $Re < 2000$ (approximately). Under these conditions, the flow is laminar . If the Reynolds increases, the flow enters into a chaotic and unstable regime, called the critical zone or transition region . This is defined as the interval $2000 < Re < 4000$. In this region, it is not even possible to find a single curve representing $f(Re)$, because the friction factor does not only depend on the Reynolds number, but also on the tube diameter and fluid viscosity. For still higher Reynolds numbers, $Re > 4000$ (approximately), the flow becomes turbulent . The friction factor can again be represented by a universal curve, provided the surface is smooth.

If the surface of the tube is not smooth, a further quantity with the physical dimensions of a length enters the game, the roughness ε, which is defined as the mean size of the surface irregularities. The corresponding dimensionless parameter is the ratio ε/D. The roughness has important influence on the boundary layer and, as a consequence, on the friction factor. The case of $\varepsilon/D = 0.02$ is shown, as an example, as the dotted curve in Fig. 1.30. We see that in the turbulent region, f is roughly constant, at the value of 0.05. The pressure drop is proportional to the square of the flow velocity.

1.13 Drag at Small Reynolds Numbers

We shall now consider the uniform motion of a solid body in a fluid at rest in an inertial frame. We might consider the same problem in a frame moving with the body, in which the fluid moves at constant speed. For the relativity principle, the drag force acting on a body moving in a fluid or on a body in a flowing fluid should be the same. Notice, however, that this conclusion holds as long as the influence of the walls containing the fluid can be neglected, namely if they are far enough apart. Indeed, relative to the walls, the body that moves in one case is at rest in the other.

In our discussion, we shall always consider a body completely immersed in a homogeneous fluid to have a sufficiently large extension. We assume that the vertical forces, buoyancy and weight, balance each other. In the case of the D'Alembert paradox discussed in Sect. 1.10, the acting forces were the pressure forces. In addition, we now have the viscous force.

The problem is similar to that of the flow in a duct and we shall treat it similarly, using dimensional arguments. The physical quantities of the game are, once more, the density ρ and the viscosity η of the fluid, the velocity υ and one geometric dimension of the body. For the simple shape we shall consider, which will be a sphere, we shall take the diameter D. Clearly, the only dimensionless combination is, again,

$$Re = \frac{\upsilon D}{\nu} = \frac{\rho \upsilon D}{\eta}, \tag{1.52}$$

which is also called the Reynolds number .

Consider this question: under which conditions are the flow patterns about two geometrically similar bodies geometrically similar? Consider, for example, two spheres of different radiuses immersed in two different fluids, say, a gas and a liquid. The question is: under which conditions do the streamlines in the two cases have the same shape? When these conditions are satisfied, we speak of *dynamic similarity* . The answer, which we give without demonstration, is that in any pairs of geometrically similar points in the two cases, the ratio of the two acting forces, the pressure drag (that is normal to the surface of the element) and the viscous drag (that is the shear stress tangent to the surface), should be the same. This, indeed, happens for equal values of the Reynolds number, for completely immersed bodies. This is an important conclusion that has relevant practical applications. For example, we can determine the behavior of an airplane wing without having to build the airplane. Rather, we can test a model of reduced dimensions at a velocity that gives the same Reynolds number. For this purpose, wind tunnels are used to test small-scale airplanes and cars. The method works as long as the compressibility of the fluid can be neglected.

Let us now consider situations in which the viscous drag is comparable or equal to the pressure drag. This happens for Reynolds numbers less than or roughly equal to one (we shall be more precise below). These conditions can be satisfied in various ways; the density of the fluid is large enough (a body moving in honey or molasses, for example), or the motion is very slow, or the size of the body is very small (for example, the fog droplets moving in air). Under these conditions, the body must, so to speak, open up its way, deforming the fluid elements. The medium appears to be more solid than the fluid, with a tendency to maintain its shape. The resistance to motion is mainly due to the forces necessary to deform the fluid elements. The resulting drag force is proportional to the velocity. We observe, in addition, that for the small values of the Reynolds number, we are considering, the stress field that develops in the medium extends to large distances from the body. Contrastingly, for large Reynolds numbers, as we shall see, the deformation of the fluid is mainly limited to the boundary layer , very close to the object. As a consequence, for small Reynolds numbers, the influence of the surrounding walls can be neglected only if they are at rather large distances.

The fluid dynamics differential equation, which is called the Navier-Stokes equation after those who discovered it, cannot, in general, be solved analytically. However, the solution is known in the case of the sphere. Consider a spherical rigid body of radius r and of perfectly smooth surface, moving in a fluid with constant velocity υ. The expression of the viscous drag in the laminar regime was theoretically determined by George Gabriel Stokes (1819–1903) in 1851. It is called the Stokes law . Its expression is very simple, namely

$$R = 6\pi r \upsilon \eta . \tag{1.53}$$

We shall now discuss its limits in regard to validity and its experimental verifications.

We preliminarily state that Stokes derived his law according to the following assumptions:

1. The medium is homogeneous; in practice, the non-homogeneities, if present, should be much smaller than the diameter of the sphere. While this condition is satisfied in liquids, it is not necessarily so in gases. In a gas, the mean distance travelled by a molecule between two collisions with another one is called the mean free path . For the atmospheric gases at normal temperature and pressure (to which it is inversely proportional), the mean free path is

about $l = 70$ nm. Robert Millikan experimentally determined in 1913 that the corrections to the Stokes law for spheres of radius r are of the order of r/l. This is, for example, <1 % for radiuses >5 µm.

2. The medium is unlimited; in practice, the surrounding walls must be far enough apart to have negligible effects.

3. The sphere is rigid and its surface is smooth.

4. The fluid does not slip, but remains adherent to the sphere surface.

5. The velocity of the sphere is small, such that the resistance to motion is due only to viscosity.

Under these hypotheses, the Stokes theory predicts Eq. (1.53) to be rigorously valid. An accurate experimental verification only came about more than half a century later, in 1910, when Harold De Forest Arnold (USA, 1883–1933) undertook a series of measurements on the fall velocity of small spheres in water and in alcohol. The forces acting on the sphere are the weight directed vertically down $(4/3)\pi r^3 \rho' g$, where ρ' is the density of the sphere, the buoyancy vertical upwards $(4/3)\pi r^3 \rho g$, and the viscous drag, opposite to motion, hence vertically upwards. During their fall, the spheres soon reach the regime velocity, in which the resultant force is zero, namely

$$6\pi r v \eta = \frac{4}{3}\pi\left(\rho' - \rho\right) g r^3 . \qquad (1.54)$$

We can verify the Stokes law by measuring the velocity, the two densities and the radius.

Arnold found an ingenious method for producing small, perfectly spherical, droplets. He used the so-called Rose alloy , a metal that melts at 82 °C and is consequently liquid at the boiling water temperature.

Arnold put some of this metal into a vertical glass tube terminating at its lower end with a long capillary tip. He suspended the capillary segment in a second vertical tube about 70 cm long and 3 cm in diameter. He filled the larger tube with water and heated the water in such a way that its upper surface was maintained at 100 °C, with the lower part being about 60 °C. Using compressed air, he produced pressure in the small tube, pushing the melted metal down so as to exit through the capillary tip in the hot water. He obtained droplets that cooled down enough to become solid before reaching the bottom of the larger tube. Observed with a microscope, they had perfectly spherical shapes and smooth surfaces. The success of the Arnold method is due to the rather slow motion of the droplets and to the gradual decrease of the temperature of the water through which they fall. The uniform cooling tends to produce a homogeneous structure, while the small velocity allows the drops to keep the spherical shape they have taken when liquid. The measured radiuses were between 20 µm and 1 mm.

Arnold measured the velocities of a number of falling spheres of different diameters and compared the results with the predictions of the Stokes law, with a few per mille accuracy.

Are the five assumptions in Stokes satisfied in Arnold’s experiments? This is certainly the case for the first, third and fourth hypotheses. The second assumption is satisfied for the smaller spheres, but only approximately for the larger ones. For these, however, Arnold used a correction formula that had been experimentally established by Landenburg. In conclusion, the first four conditions are satisfied. Through his experiments, Arnold verified that the Stokes law, Eq. (1.55), is exactly verified when the fifth condition is also satisfied, namely for a small enough Reynolds number, up to $Re < 1.2$.

1.14 General Expression of Drag

In the vast majority of situations, the Reynolds number is not as small as that considered in the previous section. In general, the pressure forces are larger or much larger than the viscous force. Consider that the kinematic viscosity of the air at normal temperature and pressure is about $\nu \cong 1.5 \times 10^{-5}\,\mathrm{m}^2\,\mathrm{s}$. For example, for a sphere of 5 cm radius moving in the air at 1 m/s, the Reynolds number is about 7000. Under these conditions, the pressure drag is much larger than the viscous drag.

One should not think, however, that the effects of viscosity are negligible. On the contrary, they are crucial. Indeed, in the absence of viscosity, the fluid elements lap the surface of the body and freely move relative to it. Contrastingly, in the presence of viscosity, the shear stresses make the fluid elements adhere to the surface of the body. This is always the case, even for extremely small values of viscosity. This means that the fluid particles in contact with the body are at rest; the farther and farther the particles are from the body, the more their velocities increase, up to the point of equalling the velocity of the unperturbed fluid. The change takes place, for a Reynolds number above a few units, in a thin layer near to the surface, which is the *boundary layer* . One sees that the boundary conditions of the velocity fields are completely different for a real fluid, compared to an ideal fluid. As a consequence, the flow pattern is different from the one in Sect. 1.10 and the pressure forces are different as well. In particular, their resultant is not zero and is, indeed, the resistance to motion. The inviscous flow is not the limit of the viscous one for viscosity tending to zero.

The boundary layer has a fundamental role in the behavior of the flow. Even if its thickness δ cannot be precisely defined on general grounds, it is, however, found that it is a decreasing function of the Reynolds number. It can be shown with dimensional arguments that it is approximately

$$\Delta \approx a/\sqrt{Re} \tag{1.55}$$

where *a* is a linear dimension of the body. For example, for a boundary layer of a sphere of diameter $a = 10$ cm immersed in an airflow of $\upsilon = 30$ m/s, the Reynolds number is $Re = 2 \times 10^5$ and the thickness of the boundary layer is $\delta \approx 0.2$ mm.

As stated above, in the boundary layer, the velocity changes from zero to the undisturbed value. The velocity gradient in the boundary layer is very large and the viscous forces, which are proportional to the velocity gradient, are intense.

Equation (1.55) also tells us that, for Reynolds numbers of the order of one, the thickness of the layer becomes comparable with the size of the body. The very concept of the boundary layer loses its meaning, because the changes of velocity gradually take place in a wide volume, as we have seen in the previous section.

For Reynolds numbers larger than one, the effects of viscosity are twofold.

(a) the development in the boundary layer of forces parallel to the surface of the body opposite to the direction of motion, directly due to the friction between fluid elements. Their resultant is the viscous drag ;

(b) a modification of the geometry of the streamlines, which, in turn, changes the pressure forces and so contributes to the total drag with a force that we shall call the pressure drag .

The resistance of the fluid, the total drag, is the sum of the two contributions; the ratio between the second and the first contribution increases with the Reynolds number.

We shall now discuss the total drag as a function of the Reynolds number. We start, once more, with a dimensional argument. The physical quantities of the problem are

$$[\rho] = \mathrm{kg\,m^{-3}}, \quad [\eta] = \mathrm{kg\,m^{-1}\,s^{-1}}, \quad [v] = \mathrm{m\,s^{-1}}, \quad [A] = \mathrm{m^2} \tag{1.56}$$

where A is the transverse cross-section of the body. The dimensions of the drag force are $[R] = \mathrm{kg\,m^{-1}\,s^{-2}}$. Again, there is one combination with these dimensions, $\rho v^2 A$. The drag must be expressed as this quantity multiplied by a function of null physical dimensions. The latter must be a function of the sole dimensionless quantity, which is the Reynolds number, and we can state that the general expression of the drag is

$$R(v) = C_D(Re)\frac{\rho v^2}{2}A \tag{1.57}$$

where we introduced the 1/2 factor to follow the costumes. The function C_D is the *drag coefficient* . It must be determined experimentally. The drag coefficient depends on the shape of the body and on the status of its surface, smooth or with different degrees of roughness.

We shall limit the discussion to a smooth surface sphere. Figure 1.31 shows the drag coefficient as a function of the Reynolds number. Notice that both scales are logarithmic.

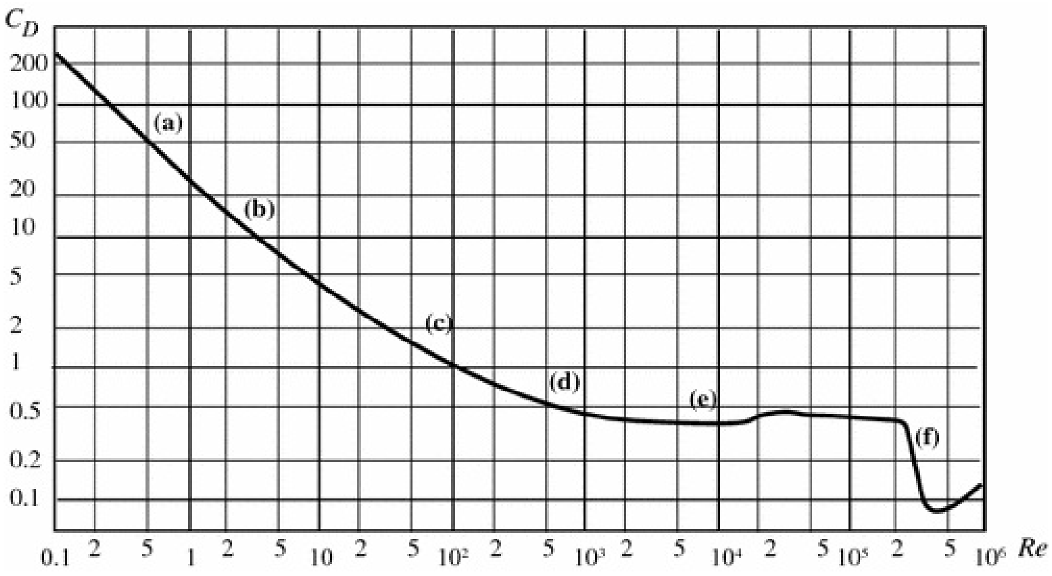


Fig. 1.31 The drag coefficient for a smooth sphere versus a Reynolds number. Letters mark the different flow regimes with reference to Fig. 1.32

We shall now examine, in a qualitative way, how the flow pattern around the sphere changes as the Reynolds number increases. These changes can be subdivided into several flow regimes . The changes from one regime to the next are gradual and cannot be sharply defined. Figure 1.32 shows a series of cartoon flow patterns with increasing *Re*. The corresponding positions on the drag coefficient curve are marked with the same letters in Fig. 1.31.

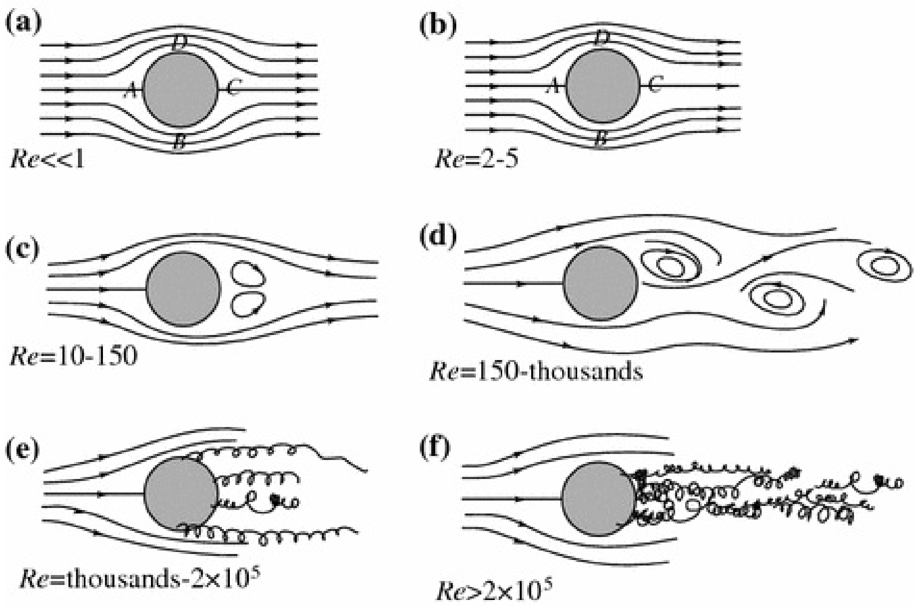


Fig. 1.32 Cartoon showing the flow patterns around a sphere in different regimes

Figure 1.32a shows the flow pattern for *Re* < 1. The drag coefficient curve in Fig. 1.31 is a straight line with a slope equal to −1. In a log-log diagram, this means that the ordinate is inversely proportional to the abscissa. This is just what we expect, because the drag is proportional to the velocity when the Stokes law holds (for shapes other than a sphere, the drag force in this regime is proportional to the velocity and to the linear dimensions of the body anyway). Indeed, if we substitute the Stokes Eq. (1.52) for the drag in Eq. (1.57) and $A = \pi D^2/4$ for the cross-section of the sphere of diameter D, we obtain

$$C_D(Re) = 24/Re. \qquad (1.58)$$

Under these conditions, as we have seen, the drag force is almost completely a viscous drag. The streamlines are qualitatively similar to the inviscous flow of Fig. 1.24. Although not shown in the figure, the velocity increases only gradually away from the surface of the sphere. There is no well-defined boundary layer at these very small Reynolds numbers.

Figure 1.32b shows the flow pattern for *Re* = 2–5. We shall use as an example here and for the following regimes the velocity of a 1 cm diameter ball in the air at normal conditions with $\nu \cong 1.5 \times 10^{-5}\,\mathrm{m}^2\,\mathrm{s}$. Its velocity is between 3 and 7.5 mm/s in this range of Reynolds numbers. The boundary layer has developed; the pressure drag is a few times larger than the viscous drag. The flow is laminar, boundary layer included. The streamlines are denser near the points *B* and *D* on the transverse section. Here, the velocity is larger and the pressure smaller than in the undisturbed fluid. On the backside of the sphere, the streamlines rarefy, the velocity decreases and the pressure increases. In other words, when a fluid element passes from the forward side to the section *BD*, it is pushed forward by a pressure difference that increases its kinetic energy. When it goes to the back region, it moves against an increasing pressure, losing kinetic energy. In the absence of dissipative forces, the kinetic energy gained in the first phase would be exactly what is needed to overcome the pressure increase in the second phase. In practice, the shear stresses, due to viscosity, reduce the kinetic energy during all the phases of the motion. This happens mainly in the boundary layer in which the shear stresses are large. Consequently, the flow velocities at points in front of the sphere are always larger than at the symmetric points on the back. The streamlines are more sparse and straighter at the backside of the sphere. The pressure forces are smaller at the back than at the front. The resulting difference is the pressure drag. This increases with the Reynolds number faster than foreseen by the Stokes law.

As the Reynolds number increases, the velocity gradient increases for two reasons: because the difference of the velocity of the fluid in contact with the surface and in the main stream increases and because the thickness of the boundary layer, in which the change happens, diminishes. The shear stresses, which are proportional to the velocity gradient, consequently increase in the boundary layer. These stresses act on the fluid elements, slowing them down considerably, especially near to the sphere. As *Re* increases, the fluid elements near the sphere slow down so much as to become at rest or even invert their velocity. Immediately after the sphere, the fluid ascends, relative to the main flow, separates from it and, reached a certain distance, turns back in the forward direction. Eddies, or vortices, start to form around $Re = 25$, causing the separation of the streamlines of the main flow.

Figure 1.32c shows the flow pattern between $Re \approx 10$ and $Re \approx 150$. The velocity of our example ball in the air is between 1.5 and 20 cm/s. Two quite regular vortices are present, which remain stably attached to the rear surface of the sphere (if the sphere moves, the vortices move with it), up to $Re \approx 100$. The streamlines outside the boundary layer go around the vortex region and join back together further downstream. The point of flow separation is close to the rear of the sphere. Notwithstanding the vortices, as these are stationary, the flow is still laminar.

As *Re* further increases, the vortices become unstable and begin to oscillate. The point of separation moves to the side of the sphere. For $100 < Re < 150$, the vortices stay close to the sphere, but for $Re > 150$, they detach, alternatively on one side and then the other, and move downstream. A vortex trail is formed, which extends considerably far downstream, symmetrically populated of eddies, which decay as another form. This is called a Kármán vortex street after Theodore von Kármán (1881–1963). Figure 1.32d shows a typical pattern for *Re* between 150 and a few thousand. The flow in the wake is no longer stationary and becomes turbulent. However, outside the trail, upstream and downstream of the sphere, the flow is laminar and stationary, namely the streamline pattern does not vary in time.

Figure 1.31 shows that the drag coefficient is roughly constant, at a value of about 0.5, in this range of Reynolds numbers and up to $Re = 2 \times 10^5$ (corresponding to 7.5 m/s for our example ball). This means that the drag is roughly proportional to the velocity squared. We talk of wake drag . The turbulent wake forbids the main current streamlines from joining back together at the rear of the sphere. The detachment point is near the diameter at 90°. The pressure forces are completely unbalanced. In the forward region, the streamline pattern is similar to the one for the inviscous fluid, with the corresponding pressure forces on the sphere. At the rear part, contrastingly, the streamlines practically do not touch the surface. The pressure force in the downstream region is close to zero.

The situation changes once more when *Re* reaches values of a few thousand (Fig. 1.32e). The boundary layer is still laminar. The entire wake is filled with turbulent eddies. The larger vortices contain smaller vortices and these, in turn, contain even smaller ones. The drag coefficient remains substantially constant at about 0.5. The drag force is proportional to the square velocity. Recall that in the D'Alambert paradox, the pressure at the front stagnation point is $\rho v^2/2$. It can be shown that the pressure at the front is, in any case, proportional to $\rho v^2/2$. The pressure force is then proportional to $A\rho v^2/2$, where *A* is the cross-section of the sphere. In conclusion, the pressure-dominated drag can be expressed as

$$R(v) = \text{const}\, A\frac{1}{2}\rho v^2 = \text{const}\, r^2\frac{1}{2}\rho v^2. \qquad (1.59)$$

At very high Reynolds numbers, around *Re* = 300,000, the drag coefficient suddenly drops, as shown in Fig. 1.31. Figure 1.32f shows the new flow pattern; the wake appears contracted compared to Fig. 1.32e. The consequence is a decrease in the wake cross-section and, see Eq. (1.59), a decrease in the pressure drag. The reason for the phenomenon, sometimes called the *drag crisis* and discovered by L. Ludwig Prandtl (1875–1953) in 1914, is as follows. At the values of Reynolds number we are considering, turbulence starts to develop in the boundary layer. This turbulence extends somewhat at the rear part of the boundary layer. The flow turbulent separation takes place downstream on the rear of the surface of the sphere at a position of 120°–130° from the front stagnation point.

This interpretation has been checked with a vertical cylinder, for which the situation is very similar to the sphere, by placing two thin sheets on the generators in the plane normal to the motion, forcing the streamline to detach from them and, consequently, the section of the wake to be independent of the Reynolds number. Under these conditions, the drop of the drag coefficient is not observed.

In the discussion of this section, we have assumed the surface of the sphere (or, more generally, of any object moving in a fluid) to be perfectly smooth. Indeed, the roughness and the structure of the surface at the sub-millimetric scale have strong effects on the boundary layer, and consequently on the drag. We have already seen similar effects for the flow in a pipe, in Sect. 1.12. As a matter of fact, a surface looking smooth to the naked eye might reveal consequential roughness at the microscopic level.

Golf balls are a good example of (visible) roughness. Their surfaces are covered with small dimples. A ball hit by a skilled golfer can leave the tee at a typical speed of $\upsilon \cong 80$ m/s. The corresponding Reynolds number ($D = 41.1$ mm in the UK, 42.7 mm in the USA) is $Re = 2.2 \times 10^5$, which is below, but not too far from, the drag crisis for a sphere of smooth surface. The dimples help to trigger the transition from a laminar to a turbulent boundary layer, bringing it down to $Re \approx 6\text{–}8 \times 10^4$, well within the range of a good golfer. The resulting drag reduction doubles the distance flown by the ball over what can be achieved with a smooth ball. Another example is the skin of fast-swimming sharks. It exhibits riblet structures aligned in the direction of flow that are able to reduce skin friction drag in the turbulent-flow regime by up to 10 %. Certain types of Olympic swimsuits are made of a material that mimics shark skin. The opposite effect has been measured with tennis balls. No drag crisis is observed in experiments in wind tunnels. This is due to the fuzz elements that cover their surfaces. If the ball is shaved, taking away the fuzz, the drag crisis appears. These examples show how the drag, for a given Reynolds number, can be very different for different types of surface. Indeed, the roughness and the structure of the surface at the sub-millimetric scale have strong effects on the boundary layer, and consequently on the drag.

Notice also that, especially at high Reynolds numbers, the drag force strongly depends on the shape of the body, especially its rear part. The narrower the wake is, the smaller the drag. One then searches for aerodynamic profiles, to minimize the drag. Typically, as shown in Fig. 1.33, the forward part should be rounded, while the rear is smoothly tapered. The streamlines then follow the shape of the body without strong variations of their density and direction. In this way, the pressure drop at the back part is strongly reduced. The production of vortices is also reduced. The detachment of the streamline, which is at the origin of vortex formation, takes place only in the neighboring parts of the tail of the body and the section of the wake is strongly reduced.

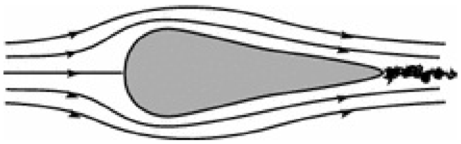


Fig. 1.33 Aerodynamic profile

We notice, in conclusion, that even when we have talked of high velocities, we have always meant them to be small compared to the speed of sound, in order that it might be possible to consider the fluid, even if it is a gas, to be incompressible.

Problems

1.1. We want to weigh 10 g of water with a precision balance with an error smaller than 1 mg. Do we need to correct for the buoyancy?

1.2. Two iron spheres of different dimensions are simultaneously dropped into the sea from the surface. Which one reaches the bottom first?

1.3. Consider the force due to the atmospheric pressure on the plane of a table of 1 m^2 area. We lay on the table a weight equal to that force. How much is its mass (in order of magnitude)?

1.4. A composite pendulum is made of a metal bar pivoted about an axis outside the center of mass. We know its mass m, its moment of inertia I and its period T. If we sink it completely in water, how does the period vary?

1.5. A wooden ball floats on the surface of a body of water. If we sink it to a depth equal to its radius and we abandon it, it will oscillate up and down. Are the oscillations harmonic?

1.6. In a Pitot tube immersed in a flow of density $\rho = 1.1 \times 10^3 \text{ kg m}^{-3}$, we measure the pressure difference $\Delta p = 4.95 \times 10^3 \text{ Pa}$. Find the fluid velocity.

1.7. A thin tube folded as in Fig. 1.34 is immersed in a water current, with its opening on the side of the flow. The water rises in the vertical part of $h = 150$ mm. Find the velocity of the current.

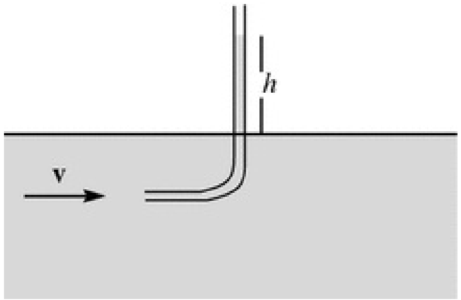


Fig. 1.34 The device of problem 1.7

1.8. A container is full of water to the height $h = 50$ cm. The water viscosity is $\eta = 1.3 \times 10^{-3}\,\mathrm{Pa\,s}$. A horizontal tube is connected to the lateral wall of the tank at the height of its base. The tube has a radius $r = 1$ mm and length $l = 1$ m. A tap at the beginning of the tube is initially closed. Then, the tap is opened and the water starts flowing. Assuming the pressure at the beginning of the tube to be equal to that at the bottom of the tank, determine if the initial flow (before the level in the tank is diminished appreciably) is laminar or turbulent and then determine the volumetric flow rate.

1.9. An air current, above a hot ground area, flows vertically up with velocity $\upsilon_a = 0.2$ m/s. The airflow transports powder particles, which we consider to be spherical with radius r. The particle upwards velocity is constant and smaller than that of the air, $\upsilon_p = 4$ cm/s. The density of the air is $\rho_a = 1.3\,\mathrm{kg\,m^{-3}}$ and that of the particles $\rho_p = 5 \times 10^3\,\mathrm{kg\,m^{-3}}$. The air viscosity is $\eta = 1.7 \times 10^{-5}\,\mathrm{Pa\,s}$. (a) Assuming the flow to be laminar, calculate the radius of the particles. (b) Verify if the assumption is correct.

1.10. The initial velocity of a tennis ball hit by a champion is $\upsilon = 60$ m/s, its diameter being $D = 6.5$ cm; the velocity of a football/soccer ball after the kick of a top player is $\upsilon = 30$ m/s, its diameter being $D = 22$ cm. How much are the Reynolds numbers?

Alessandro Bettini, A Course in Classical Physics 2—Fluids and Thermodynamics, Undergraduate Lecture Notes in Physics, DOI 10.1007/978-3-319-30686-5_2

2. First Law of Thermodynamics

Alessandro Bettini[1] ✉

(1) Dipartimento di Fisica e Astronomia, Università di Padova, Padova, Italy

✉ **Alessandro Bettini**
Email: alessandro.bettini@pd.infn.it

Thermodynamics developed historically after mechanics, mainly in the XIX century. Development was motivated by two main needs, both outside of pure physics. The first motivation was the search to understand how the leaving creature “produces” energy; the second was the desire to develop engines capable of transforming heat, generated, for example, by the burning of coal, to produce mechanical work which otherwise had to be done by humans or animals. However, thermodynamics is not important for engineering and biology alone, but represents one of the fundamental chapters in physics.

In the first volume of this course, dealing with mechanics, we saw that, for an isolated system, the total energy, or, more precisely, the total mechanical energy, namely the sum of potential and kinetic energy, is conserved, i.e., remains constant, only if all acting forces are conservative. If dissipative forces are present, the energy appears not to have been conserved. However, this non-conservation is only apparent due to the fact that other forms of energy exist beyond the mechanical energy that we did not include in the balance. Thermodynamics shows us that all physical bodies contain energy, called internal or thermal energy, which does not depend on their velocity (like kinetic energy) or position (like potential energy) but on other variables, like temperature and pressure, which are called thermodynamic coordinates. Thermal energy can be exchanged between systems in two ways: one is work, which we already know, and the other is heat, which we shall learn about in this Chapter.

Figure 2.1 shows the life spans of the major contributors to thermodynamics, starting with the theory of gases.

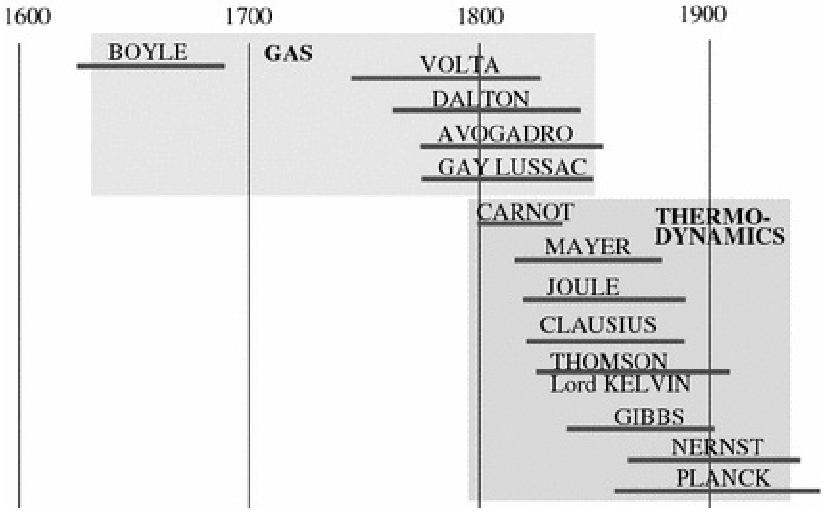


Fig. 2.1 Life spans of the fathers of thermodynamics

In 1824, the French engineer N.L. Sadi Carnot published the brief but fundamental article “Reflections on the motive force of heat”, in which he made completely clear the limits within which heat can be transformed into work, or, more precisely, the limit on the efficiency of any heat engine. Carnot developed his theory, which became the second law of thermodynamics, when heat was still believed to be a fluid, called “caloric”. Notwithstanding that, his arguments are completely correct. Forty years after his death in 1832 from the plague, his notes were published. Reading these notes, one can understand how the young genius had already understood the equivalence of heat and work.

Twenty-one years after Carnot’s article, in 1845, R.J. Mayer published the paper in which he fully established the equivalence between heat and work. This is the first law of thermodynamics, which is the law of energy conservation. In the very same year, J.P. Joule published his simple and ingenious experiment, which we shall discuss in Sect. 2.7.

Thermodynamics is closely linked to mechanics. Indeed, all thermodynamic phenomena can be interpreted through statistical mechanics, as we shall see in Chaps. 5 and 6. All bodies are made of molecules, which are matter particles whose motion follows the laws of mechanics. However, the number of molecules is so huge that it is practically impossible to describe the motion of each of them in detail. But even if it were possible, such a description would be useless. As we have seen in the previous chapter, several of the motions of fluids have already proven incapable of being analytically treated with mechanics equations. Similarly, statistical mechanics considers suitable mean values of the kinematic quantities (velocity, kinetic energy, etc.) Historically, the development of statistical mechanics is mainly due to Maxwell and Boltzmann in the second part of the XIX century.

It is important to realize that thermodynamic and statistical mechanics points of view are different and complementary. The fundamental laws of thermodynamics are established by inference starting from the experiments. They are then assumed to be axioms; their consequences are logically deduced and experimentally controlled. The method is powerful because it allows for obtaining very precise results, while statistical mechanics is sometimes forced to introduce simplifying assumptions in order to be able to proceed. In addition, thermodynamic laws are general. They also hold for systems not composed of molecules, like electromagnetic radiation in a metal box. As a matter of fact, the discovery of quantum mechanics by Planck happened during the study of the thermodynamics of the electromagnetic field. On the other hand, statistical mechanics unifies thermodynamics and mechanics by showing which elementary mechanical processes are at the basis of heat exchanges and, more generally, of all thermodynamic processes.

In this chapter, we shall start by introducing the concept of the thermodynamic system and the main thermodynamic variables, or coordinates, the pressure, the temperature and the volume. We shall define the thermodynamic state and discuss the different types of processes from one state to another and the equations that govern them. We shall then discuss the experiments and arguments that led to the establishment of the first law of thermodynamics and discuss its consequences.

2.1 The Thermodynamic State

Consider a certain amount of gas contained in a box. The constituent particles, i.e., the molecules, are free to move about inside the box and may have any velocity. There are no constraints, as there are for rigid bodies, limiting the degrees of freedom of the system. If we then want to know the mechanical state of the system, we need to know 6 N parameters, the three coordinates and the three components of the velocity of each molecule. The number N of molecules being huge, this is impossible in practice.

In thermodynamics, the concept of the state of a system is different from that of the same system in mechanics. The thermodynamic state is defined by a small number of variables , which are different from the mechanical ones. We shall start with a few cases.

Homogeneous fluid of only one chemical species at rest contained in a box.

Think, for example, of a bottle of nitrogen under pressure, of a balloon full of helium, of a pot of water, etc. The quantities that we can easily measure are: the mass of the fluid m, its volume V, its pressure p and its temperature θ. We have considered the volume to be small enough so that the pressure and temperature may be the same at all points of the system. If the system were, for example, the earth's atmosphere, this would not be true. We have already defined the pressure; we shall subsequently define the temperature in the next section. For the moment, just consider what you would measure with a common thermometer.

We might think that the volume would not be enough to characterize the geometrical properties of the system. We might also think that we need to know its shape. The container of the water might be, for example, spherical or cubic, or taller than it is wide, etc. Why did we not include, for example, the area of the surface amongst the variables? Only the experiment can give the answer. As a matter of fact, we have experimentally found that the largest fraction of the thermodynamic properties of a fluid is independent of its shape. However, when the surface to volume ratio is large, as in fog droplets or soap bubbles, the surface must also be considered. We shall do that in Chap. 4.

Homogeneous solid body of only one chemical species.

In this, case we must consider, beyond the variables considered for a fluid, the shear stresses, and the pressures and tensions that can be present in the bulk of the body. Think, for example, of a metal parallelepiped subject to external tension along one axis and pressure on the perpendicular faces. Such situations quickly become complicated, and we shall not discuss them.

Rubber band .

The thermodynamic variables are the length and the temperature of the band.

System composed of one chemical species in different states (*or aggregation phases*).

Consider, for example, a system composed of liquid water and ice in a container at °C, or an alcohol and its vapor. To specify the thermodynamic state, we now need, beyond the already mentioned variables, the fraction of ice and water, or of alcohol liquid and vapor, etc., for each of the phases. This variable is called concentration .

System composed of more than one chemical species .

The thermodynamic state is defined by the variables: total mass, volume, pressure, temperature and concentrations of the different chemical species.

More complex to describe is the non-homogeneous system, in which some of the variables, for example, the temperature or the pressure, vary from point to point. To study them, one must divide the system into parts that are small enough to be able to be considered homogeneous. We shall not deal with any of these problems.

We shall always assume, as we have already implicitly done, that all of the system's parts are at rest, or moving so slowly that their kinetic energies can be neglected. Notice that the thermodynamic state of a system does not change if it moves all together. For example, the thermodynamic state of a pot of water at a certain temperature is the same whether it is on the ground or on a train moving at 300 km/h.

From what we have established, it is clear that knowledge of the thermodynamic state of a system gives very little to no information on the mechanical state of its molecules. Consider a gas contained in a box. The mass, volume, pressure and temperature are constant; its thermodynamic state does not vary. But its molecules move continuously, and their coordinates and velocities change. The mechanical state varies, while the thermodynamic one is constant.

Particularly important amongst the thermodynamic states are the states of thermodynamic equilibrium . These are the states that remain unaltered as long as the external conditions do not change. A state is of thermodynamic equilibrium when the following conditions are satisfied:

(a) *mechanical equilibrium* .

If, as is usually the case, pressure is one of the thermodynamic variables, it must be equal in all the parts of the system. Otherwise, movements would happen inside the system. Consider, for example, a gas inside a cylinder closed by a piston of surface S, which is movable without friction, as in Fig. 2.2. If p is the pressure of the gas, the force exerted by the gas on the piston is pS, directed vertically up. Let us call the external force on the piston directed down F_a. This is the resultant of the force due to the atmospheric pressure, the weight of the piston and, possibly, of another weight that we laid on the piston. The mechanical equilibrium is reached when the gas pressure force is equal to F_a. If there is no weight on the piston and the weight of the piston itself is negligible, the internal pressure is equal to the external pressure at mechanical equilibrium. If the gas is enclosed in a rigid bottle, its pressure can be completely different from that which is external.

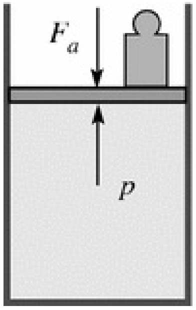


Fig. 2.2 A gas in a cylinder closed by a movable piston

Going back to Fig. 2.2, suppose now that there is friction between the piston and the cylinder, as is always the case in practice. In this case, the mechanical equilibrium can exist, even if the pressure of the gas is larger or smaller than that which is external, as long as the force resulting from the pressure difference is smaller than the maximum static friction force.

As a second example, consider a rubber band fixed at one extreme and hanging vertically down, as in Fig. 2.3. If we apply a force F_a to the other extreme, the rubber band will stretch up to the deformation at which the elastic force T is equal to F_a.

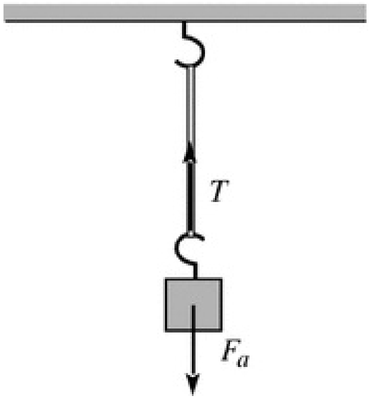


Fig. 2.3 Rubber band in equilibrium

(b) *thermal equilibrium* .

The first necessary condition is that all the parts of the system have the same temperature. Otherwise, temperatures tend to become equal and we do not have equilibrium. The second condition is that temperature should not vary due to the external environment. Suppose that, as is usually the case, the environment has a definite constant temperature. Clearly, if the temperatures of the system and the environment are equal, the system is in thermal equilibrium. However, this is not a necessary condition. If the walls surrounding the system are thermally insulating, the temperature of the system does not vary even if different from that which is external. Walls having this property are said to be adiabatic , from the Greek words *a* (not), *dia* (through) and *bainein* (to go). Even if perfectly adiabatic conditions cannot be realized in practice, for example polystyrenes boxes or dewars make for good approximations.

(c) *chemical equilibrium* .

If there are more chemical species with thermodynamic equilibrium, the concentrations of the different species must be constant in time. The same is true if there is only one species, in different phases (liquid and solid, liquid and vapor, etc.)

2.2 Temperature

The concept of temperature is linked in the common sense of the world to the feeling of hot or cold. In physics, the concept must be precisely defined. We shall define temperature operationally, namely as the set of operations needed to measure it. We state immediately that we shall proceed by approximations, gradually increasing the precision of the definition. The reason for this is that the most precise definitions of temperature require thermodynamic concepts, which, in turn, require some knowledge of temperature. There is no risk of circular arguments, as we shall always rely on experiments.

Our simpler definition stems from the following considerations. We start from our own perception of “temperature” as a state in which an object is perceived as being either colder or hotter. We take two bodies of different temperature and put them in contact. We feel both temperatures and determine that they vary with time but, after a while, both become stationary. We conclude that the two bodies are in thermal equilibrium. Can we state that they have the same temperature? Not yet. We must pay attention to the fact that equality enjoys the transitive property; if A is equal to B and B is equal to C, then A should be equal to C. We need to check if the property is satisfied through experiment; we cannot state it through logic. Experimentally, we find that, if body *A* is in thermal equilibrium with body *B* and, separately, body *B* is in equilibrium with body *C*, then, if we put *A* and *C* in contact, their temperatures do not vary; they are in thermal equilibrium as well.

We can now define temperature as follows. We take two bodies, one, say *A*, is the thermometer, the other, *B*, is the body the temperature of which we want to measure. We put them in contact and wait for thermal equilibrium. If we now have a third body, *C*, which is in thermal equilibrium with *B*, the transitory property we have found insures that the thermometer *A* will measure the same temperature when put in contact with *C* as it did with *B*.

The traditional thermometers consist of a glass bulb containing a liquid connected to a capillary tube several centimeters long. When in contact with a warmer body, the liquid expands; the higher the temperature, the higher it rises in the capillary. Mercury was in standard use as thermometer liquid until the 1990s, when it was judged to be too risky to handle, being poisonous and thus potentially dangerous in cases of the glass accidentally breaking. Mercury was subsequently replaced as a thermometric liquid by colored alcohols. These thermometers, in turn, were soon replaced by so-called electronic thermometers. We shall come back to their working principles at the end of the section, and base our discussion on liquid-in-glass thermometers, which are conceptually simpler.

For quantitative measurement, we need a scale. To have a scale, we must fix two values, the zero and the step or degree of the scale, namely the unit. For that, we need two systems having a well-defined temperature. We observe that a mixture of pure water and ice always has the same temperature provided it is at the same pressure. We can check with our thermometer, even if we still lack a scale, by observing that it always sets at the same level when in contact with the mixture. The same is true for a mixture of water and its vapor in equilibrium at the same pressure. On the Celsius scale , zero (0 °C) is defined as the water-ice equilibrium temperature at the standard atmospheric pressure. It is named after Anders Celsius (Sweden, 1701–1744). The standard atmospheric pressure is defined as $p = 1.013 \times 10^5$ Pa. A temperature of one hundred degrees (100 °C) is defined as the water-vapor equilibrium temperature at the standard atmospheric pressure.

We can proceed as follows. We emerge our thermometer in the ice-water mixture at atmospheric pressure and mark a line on the capillary at the level of the thermometer liquid, writing a 0 on the tube. We do the same with the water-vapor mixture, and mark 100. We still need the degree. The best we can do is to divide the length between the two marks into one hundred equal parts. We can continue with the same step above 100 and below 0.

Pay attention, however. In dividing the length into equal parts, we have implicitly assumed that the length of the liquid in the capillary does vary linearly with temperature. Is that true? One way to check is to build, following the above procedure, a number of thermometers, say several made of different glasses, with different diameters, containing different liquids, etc. Then, we measure the temperature, which should be different from 0 and 100, of a reference body with all of them. We find that the readings of the thermometers are equal only upon first approximation. We can find differences of a few tenths of a degree between thermometers with the same liquid and up to a few degrees if the liquid is different. As a matter of fact, the thermal dilatation of the bodies is not exactly a linear function of temperature. In addition, the temperature reading depends on the glass as well and, even if only weakly, on the age of the glass.

In conclusion, the liquid-based thermometers are very simple, cheap and easy to use. They are particularly useful for somewhat imprecise measurements. However, in physics, temperature is a very important quantity, and must be defined as accurately as possible (as usual, infinite accuracy does not exist). The method is the ideal gas thermometer .

We preliminarily observe that the gas thermometer measures the *absolute temperature* . As a matter of fact, the Celsius scale (and the Fahrenheit scale as well) is arbitrary; it is not based on a physical law. However, a temperature exists that must necessarily be considered to be zero on a physical basis. This is called *absolute zero* . The most direct experimental evidence comes from the laws experimentally established by Alessandro Volta (Italy, 1745–1827), Joseph Louis Gay-Lussac (France, 1778–1850) and others. These heuristic laws were later included in the gas law that we shall discuss in the next section. We anticipate here that these authors found both the pressure of a gas at constant volume and its volume at constant pressure to be linear functions of the temperature, which was measured on the Celsius scale. An important observation was that both the pressure and the volume tend towards zero when the temperature tends towards the same well-defined value, which is −273.15 °C. The same laws predict both pressure and volume to be negative below that temperature. This fact being meaningless, the temperature of −273.15 °C is the absolute zero. It is physically impossible to reach temperatures lower than that. The zero of the absolute temperature scale is the absolute zero.

We now need the unit of temperature, which is called the kelvin [1] (K), after Lord William Thomson , Baron of Kelvin (UK, 1824–1907). The definition of kelvin is, as for the other units, the responsibility of the Bureau International des Poids et Mesures (BIPM, for short). The BIPM has changed the definition over time to make it as precise as possible, taking advantage of technological progress. The kelvin is defined by establishing the temperature of a fixed point. A mixture of the three phases of a substance, water in particular, is in equilibrium at a certain temperature and a certain pressure only. This is called the triple point . The water triple point temperature (water, ice and vapor in equilibrium) is by definition 273.16 K. The number has been chosen to have one kelvin be almost equal to the pre-existing one degree Celsius.

As we have already stated, the gas thermometer is a precision instrument. As such, it is not of simple use. Accurate procedures are required to reduce systematic errors as much as possible. We shall not enter into such issues, being interested here in the operation principles.

The thermometer, schematically shown in Fig. 2.4, consists of a bulb (made of quartz or metal) containing a gas, which is connected through thin tubing to a mercury manometer, so as to measure the gas pressure. All the pressure measurements are done *at constant volume*. We have also prepared the water, ice, and vapor mixture in equilibrium at the triple point. We are going to measure the temperature of a reference system, say a mixture of water and vapor in equilibrium at atmospheric pressure.

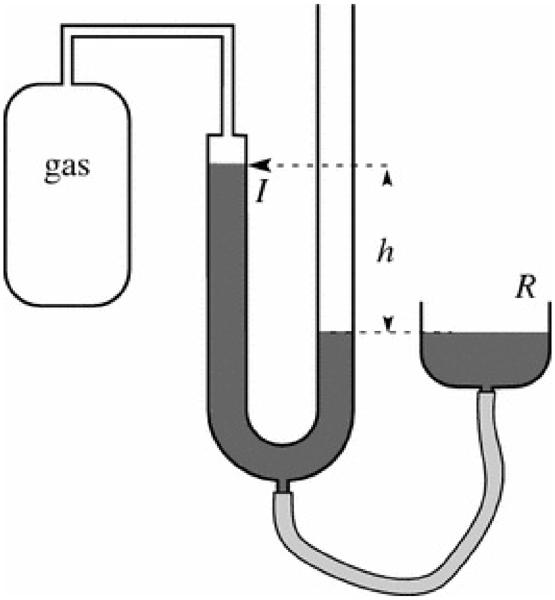


Fig. 2.4 The gas thermometer

We put into the bulb a certain quantity, say m_1, of a gas, say nitrogen. We then put the bulb in contact with the triple point bath. The auxiliary container *R*, which is connected to the manometer by a flexible rubber tube, is used, lifting or lowering it according to need, to ensure that the level of the mercury in the branch of the manometer on the side of the gas is always at the same level, for every measurement. The position is marked by the index *I*. In such a way, we guarantee that the volume of the gas will always be the same. The measurement of the height *h* gives the difference between atmospheric and gas pressures, and the latter when the atmospheric pressure is known. Let us call it p_{tr}. It is convenient to choose a mass m_1 small enough to have a rather low p_{tr} compared to the atmospheric pressure.

We now put the bulb in contact with the water vapor mixture. We measure the pressure as before and we call it *p*. We now preliminarily define the temperature of the mixture, assuming it to be proportional to the pressure, namely as

$$T(p) = \frac{p}{p_{tr}} 273.16\,\mathrm{K} \tag{2.1}$$

where we anticipated in the notation that the so-defined temperature might depend on the pressure *p*.

The definition can be accepted only if we find the same result using another gas. In practice, this is not so. We perform three measurements with bulbs full of, respectively, nitrogen, oxygen and helium. We regulate the quantities of the different gases to obtain the same pressure at the triple point. It is equal to 80 kPa in this case. We find that the three pressures measured with the bulb in contact with the water-vapor mixture are a bit different. Consequently, the temperatures given by Eq. (2.1) are also a bit different. They are shown in Fig. 2.5 at the abscissa of 80 kPa. The differences are of a few tenths of a degree. They are small, but nonetheless tell us that the definition Eq. (2.1) is not sufficiently accurate.

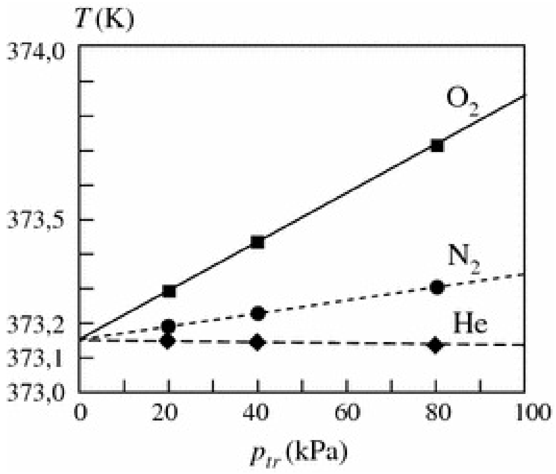


Fig. 2.5 Temperature, as defined by Eq. (2.1) measured with different gases and different pressures

We proceed, introducing into the bulb smaller quantities of the gases. The pressure measured at the triple point is now $p_{tr} = 40$ kPa for all the gases. We repeat the operations and find the three temperatures in Fig. 2.5 at the abscissa of 40 kPa. The values are now closer to one another. We perform a third set of measurements, again halving the pressures to $p_{tr} = 20$ kPa. We find that the measured temperatures get still closer. But Fig. 2.4 shows us more. If we linearly extrapolate the measurement made with each gas to zero p_{tr}, all of them lead to the same value! The extrapolated value does not depend on the gas we use. The behavior of the gases tends to be the same at zero pressure. The gas having this limit behavior is called an *ideal gas* .

In conclusion, we define ideal gas temperature , which is also the absolute temperature, as

$$T(p) = \left(\lim_{p_{tr}\to 0} \frac{p}{p_{tr}}\right) 273.16\,\mathrm{K}. \tag{2.2}$$

In practice, the procedure just discussed is very delicate and requires weeks of work. The measurements are done in metrological laboratories . The temperatures of the triple points and, at a definite pressure, the fusion and boiling points of a number of pure substances are accurately measured. These points are then employed to calibrate secondary thermometers that are simpler to use.

The gas thermometer does not allow for measuring very low temperatures, because all gasses liquefy at low enough temperatures. The lowest liquefaction temperature is for He, at 4 K at atmospheric pressure. In practice, the lowest measurable temperature with the gas thermometer is about 1 K. Below that, the ideal gas temperature is not defined. We shall see in Sect. 3.5 how a temperature scale based on thermodynamic arguments can be defined. This is called thermodynamic temperature and is meaningful at all temperatures. In the interval of definition of both, the two scales coincide.

Many types of thermometers are commercially available. They are based on a number of temperature-dependent physical processes, work in a certain temperature range, and have different accuracies, depending on the use for which they are meant (medical, room, laboratory, etc.). Here, we recall only the already mentioned and very common electronic thermometers , also called digital because they display the temperature on a digital screen. The majority of them are based on the thermoelectric effect ; the electric resistance of some metal varies almost linearly with temperature. The resistance of the metal sensor is measured by means of an electronic circuit and displayed. The thermometers are calibrated by the production company. The accuracy is typically a few tenths of a degree (±0.1°–0.2° for medical models). Better accuracy can be obtained with thermistors, which employ the temperature dependence of the electric resistance of semiconductors.

We add a final consideration. In any temperature measurement, whatever the thermometer may be, the temperature of the thermometer is what is measured. Consequently, we must always ascertain that the thermometer is in thermal equilibrium with the object whose temperature we are measuring. In practice, the thermal contact between the two must be good, and one must wait for the equilibrium to be reached. In this process, there is always some heat transfer from body to thermometer, or vice versa. As a consequence, both temperatures vary. However, if the mass (or, even better, the heat capacity, which we shall define) of the thermometer is much smaller than that of the body, the temperature change of the latter is negligible.

Lastly, we notice that the water-ice equilibrium temperature at normal atmospheric pressure, namely 0 °C, is equal to 273.15 K.

2.3 State Equation

Consider a homogeneous fluid made of a single substance. As we said, the thermodynamic coordinates of the system are the occupied volume V, the pressure p and the absolute temperature T, which we take in the kelvin scale. In order for the coordinates to have definite values, the system must be in thermodynamic equilibrium .

In our study, we shall consider only closed systems . Consequently, in its processes, the mass of the fluid does not vary. The state variables are then three: p, V and T. A system described by these variables is called a hydrostatic system . It is experimentally found that the three variables are not independent; only two of them are as such. For example, if a gas is enclosed in a given volume and we exert a certain pressure, then it assumes a well-defined temperature. Similarly, if we take a defined volume of a gas at a certain temperature, its pressure assumes a definite value. The relationship amongst the three values is called a *state equation* and can be expressed as

$$f(p, V, T) = 0. \tag{2.3}$$

Every hydrostatic system is characterized by its own state equation. No real system has a state equation that can be expressed analytically. It is, however, always possible, and extremely useful, to determine the equation experimentally. To do that, we put the system in a sufficiently large number of different states and measure the thermodynamic coordinates in each of them. We have limited our considerations to hydrostatic systems for simplicity. Notice, however, that all the thermodynamic systems are ruled by a state equation that connects its state variables. We shall see an example at the end of the section.

As a consequence, two thermodynamic variables are sufficient to define any equilibrium state of a hydrostatic system of given mass. The most often used pair is volume and pressure. The equilibrium states are represented as points on the p, V Cartesian plane, with V as abscissa and p as ordinate. Clearly, only the positive part of the V axis is meaningful. The pressure is almost always positive too, but it may exceptionally be negative.

Several scientists in the XVII and XVIII centuries were responsible for developing the physics of gases and, in particular, their thermodynamics. We have graphically represented the life spans of the main contributors in Fig. 2.1. They are: Edme Mariotte (France, 1620–1684), Robert Boyle (UK, 1627–1691), Alessando Volta (Italy, 1745–1827), John Dalton (UK, 1766–1844), Amedeo Avogadro (Italy, 1776–1856) and Joseph Louis Gay-Lussac (France, 1778–1850).

The results of their work can be summarized as follows. If we change the volume of a gas, keeping its temperature fixed, the pressure varies in almost inverse proportion to the volume. In other words, the product of pressure and volume remains nearly constant at constant temperature. If the temperature varies, the product of pressure and volume vary proportionally to it, provided it is the absolute temperature. The proportionality constant is, in turn, proportional to the mass of the gas. The state equation of the gas we are considering is thus

$$pV = rmT$$

where r is the proportionality constant, which is different for different gases. The equation is very simple and, as we shall now see, contains a lot of information. But there is more to it.

A quantity of a well-defined substance can be measured in two units: the kilogram, which measures its mass, and the *mole* , which measures its number of molecules, which are of one species only because the substance is defined. The symbol of the mole is mol . A mole of a substance is an Avogadro number of molecules of that substance. When the mole was defined, and originally called a gram molecule, the unit of mass was the gram. The molecular masses were defined in grams as well. A mole is the number of grams of the considered substance that contains an Avogadro number of molecules. For example, a mole of hydrogen has a mass of one gram. The Avogadro number is one of the fundamental constants of physics. Its first digits should be remembered by heart. It is very large, namely

$$N_A = (6.02214129 \pm 0.00000027) \times 10^{23}\,\mathrm{mol}^{-1}. \qquad (2.4)$$

Coming back to the gas equation, let us express the quantity of gas in number of moles, $n = m \times 10^{-3}/M$, where M is the molar mass of the substance and the 10^{-3} factor is due to the fact that m is in kilograms. The state equation becomes

$$PV = nRT \qquad (2.5)$$

where R is a new constant ($R = rm/n$). Up to now, we have simply re-defined the proportionality constant. However, what matters is that experiments show that R is a universal constant, namely that it is the same for all gases. This is called the *gas constant* . Its value is

$$R = 8.3144598 \pm 0.0000048\,\mathrm{J\,mol^{-1}\,K^{-1}}. \qquad (2.6)$$

Rigorously spiking, no gas follows Eq. (2.5) exactly. However, the most common gases, such as the gases in the atmosphere, behave approximately according to that equation in a large interval of pressures and temperatures. The approximation is better the higher the temperature and the lower the pressure. The *ideal gas* is defined as a gas that rigorously obeys Eq. (2.5). This is called the *ideal gas equation* , but also the *Boyle law* and *Boyle-Mariotte law* . We can see now that the definition of absolute temperature given in the preceding section was based on this law.

A fundamental implication of the gas law was established in 1811 by Amedeo Avogadro . The Avogadro law states that *equal volumes of different (ideal) gases with the same conditions of temperature and pressure contain the same number of molecules*. Today, we see that this is an immediate consequence of the universality of the R constant. A useful quantity is the *molar volume* , which is the volume occupied by a mole of (ideal) gas at, as it is called, STP , meaning standard temperature and pressure . These are defined as $T = 273$ K (about 0 °C) and $p = 10^5$ Pa (about one atmosphere). The molar volume of ideal gases is

$$V_m = 0.0224\,\mathrm{m}^3 = 22.4\,\mathrm{L}. \qquad (2.7)$$

The gas equation Eq. (2.5) contains several heuristic laws that were discovered by different researchers, as mentioned in the previous section. Very importantly, it contains the law discovered in 1801 by John Dalton , called the *Dalton law* after him , and also the *law of partial pressures* . The law states that: "*In a mixture of different gases the partial pressure of each of them is the pressure it would exert if it occupied all the volume alone*." It is as if each gas was kept unto itself without any interaction with the other ones. The Dalton law is rigorously valid for the ideal gases.

We close this section considering an example of a non-hydrostatic thermodynamical system; the rubber band represented in Fig. 2.3. The thermodynamical variables are the length l and the tension τ, which depends at equilibrium on the applied force F_a, and the temperature T. The elastic constant of rubber depends on the temperature, increasing as the temperature increases. The rubber becomes "harder" if heated. In other words, for a fixed length, the tension depends on temperature. We can also find that stretching the band causes its temperature to increase, while relaxing the band causes it to diminish. Even in this case, there is a state equation, linking the three thermodynamical variables. Only two of them are independent. However, this equation of state cannot be expressed analytically.

2.4 Processes

A thermodynamic process , or thermodynamic transformation (the two terms are synonyms), happens when a thermodynamic system changes from an initial to a final state. Generally speaking, each of them can be an equilibrium state or not. However, we shall only consider the former case. The states of the system during transformation can never be rigorously of equilibrium, because an equilibrium state is stationary. We shall now define different types of processes.

Quasi-static processes . A process is said to be quasi-static if the states taken by the system differ from stationary states by infinitesimal quantities. In practice, the coordinates of the system should vary very slowly, allowing it the time to adjust to the changed conditions before the coordinates change again. The initial and final states of a quasi-static process are necessarily equilibrium states.

Consider, for example, a gas in a cylinder closed by a piston. If we want to increase its pressure in a quasi-static manner, we must move the piston slowly. At each small displacement of the piston, the pressure of the gas increases in a layer immediately under the piston. The other parts of the gas still have the initial pressure. Soon, the pressure increase propagates gradually throughout the volume. The piston must move slowly enough to allow the pressure throughout the gas volume to reach (close to) the same value. If we lay a heavy weight on the piston and abandon it, the piston will move down quickly and the process will not be quasi static. Contrastingly, if we gradually pour some sand on the piston, the process will be quasi-static.

Reversible processes . A process is reversible if it is quasi-static and if dissipative forces are negligible. Consider a generic state P in a certain instant of a quasi-static process. P is almost an equilibrium state. Immediately before that instant, an infinitesimal variation of the external conditions, the system was in a state infinitely near to P. If the process is reversible, when we invert the infinitesimal variation of the external conditions with the system in P, the system goes back to the state it was in immediately before P.

Let us again consider the gas contained in a cylinder, as in Fig. 2.6. Suppose friction to be present between the piston and the cylinder. The force F_a is applied externally to the piston, normally downward. To have a quasi-static expansion process, we slowly decrease the applied force. Assume a friction force F_r to be present. Its direction is opposite to that of the motion. The gas exerts a force pS on the piston directed vertically upward. To have the piston moving up, it should be $F_a < pS - F_r$, as in Fig. 2.6a. If we want the motion to be slow, the inequality should be just satisfied.

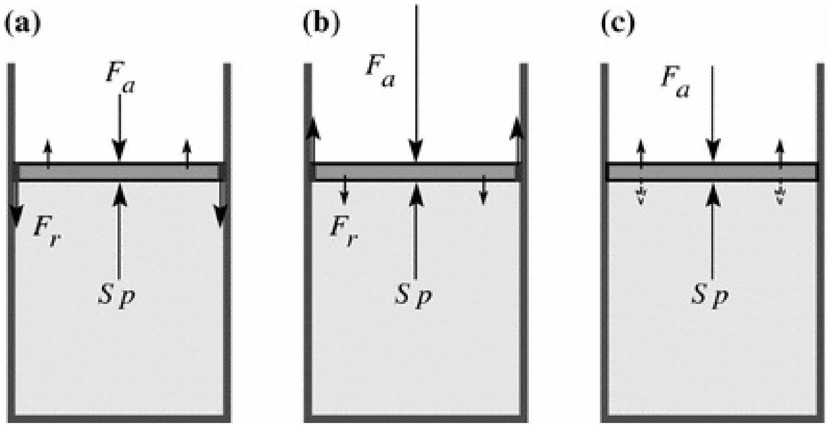


Fig. 2.6 Forces on the piston. **a** Expansion, with friction, **b** compression, with friction, **c** reversible, no friction

The process is quasi-static but not reversible. Indeed, if we want to invert the process, namely to compress the gas and have the piston moving down, we must increase F_a, but not only by an infinitesimal quantity. This is because the friction force, which always opposes the motion, changes signs and is finite, not infinitesimal. The condition becomes $F_a > pS + F_r$, as in Fig. 2.6b.

We see that, in order to pass from expansion to compression, changing the external conditions by an infinitesimal quantity (the force F_a in the example), the friction must be zero, as in Fig. 2.6c.

Rigorously speaking, quasi-static processes do not exist, and even less so the reversible ones. However, we can often operate in conditions that are close enough.

Irreversible processes . All the processes that are not reversible (all the natural ones), namely when dissipative effects are present, or the system goes through states that do not only differ by infinitesimals from equilibrium states (or both), are irreversible.

Cyclic processes . A process is cyclic if initial and final states coincide. Clearly, a reversible process can be quasi-static, reversible or irreversible.

We now define the *adiabatic wall* . Consider two systems A and B at different temperatures, T_A and T_B, respectively. If we put them in contact, both temperatures vary until they become equal. We make a series of experiments with plates of different materials introduced between the two systems, as shown in Fig. 2.7. We find that the time necessary to reach thermal equilibrium, starting from the same temperatures, are different for different materials. If the plate is metallic and thin, the time is very short; if it is foamed polystyrene and thick, the time is very long. Other materials give intermediate results. The materials of the first type are called *good thermal conductors* , those of the second, *good thermal insulators* .

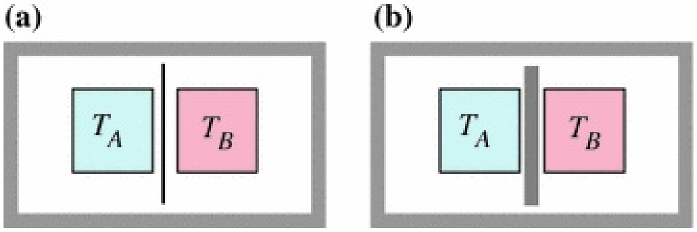


Fig. 2.7 Thermodynamic systems at different temperatures separated by (**a**) diathermic wall, (**b**) adiabatic wall

Two idealized cases are considered. We call walls perfectly permeable to heat *diathermic walls* and walls that completely block the heat transfer *adiabatic walls*. When two systems are separated by a diathermic wall, the temperature equilibrium is reached exactly as if the wall was not present; if the wall is adiabatic, each system keeps its temperature indefinitely.

Figure 2.8 represents two thermodynamic systems closed in a single container. The walls of the container are adiabatic; the two systems can interact with one another, but not with the environment.

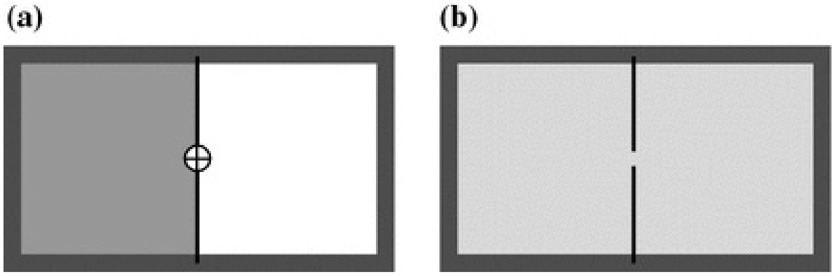


Fig. 2.8 A gas in an adiabatic and rigid box. **a** All the gas on the *left-hand side*, tap closed, **b** gas in all the box, tap open

A system is *closed* if it does not exchange matter with the surrounding environment. A pot of water boiling on the fire is not a closed system (even with a cover), because vapor exits; water in a pressure cooker, before it whistles, is a closed system.

A thermodynamic system is *isolated* if it can exchange neither mass nor energy with the environment. A necessary condition is for it to be completely enclosed in adiabatic walls. The condition forbids energy exchanges as heat (see discussion in the following section), but not as work. Consequently, the condition is not sufficient.

An isolated system in an equilibrium state remains indefinitely in that state. To move it out of that state, it is necessary to add energy to it or subtract energy from it, which is not possible by definition. If, contrastingly, the system is initially in a non-equilibrium state, it spontaneously performs a process that brings it to an equilibrium state. The processes of this type are called *spontaneous processes* . Consider, for example, the two parts of a system, isolated from the external environment, in thermal contact at different temperatures in Fig. 2.7a. They are not in an equilibrium state. The system spontaneously evolves towards the equilibrium state in which the temperatures are equal.

Consider a container with adiabatic and rigid walls. A diaphragm with a tap divides the container into two halves. The half on the left-hand side contains a gas; the half on the right-hand side is empty. This is an equilibrium state, but ceases to be one when we open the tap. The system then spontaneously evolves until the point when the gas occupies the entire volume with uniform pressure. The system is now in equilibrium and its state no longer changes.

Obviously, all spontaneous processes are irreversible.

As we already mentioned in the previous section, we can represent any equilibrium state of a hydrostatic system with a point on a *Vp* plane. In Fig. 2.9, points *A* and *B* represent equilibrium states. The non-equilibrium states cannot be represented by points, because at least some of their thermodynamic coordinates are not defined. In this case, only V is defined, because the system is in a rigid container, while the pressure and the temperature are not defined. A quasi-static process is represented on the Vp plane by a curve; the points of the curve are the equilibrium states crossed by the system. The curve Γ in the figure is an example of a quasi-static process from A to B. If the process is not quasi-static, even if it joins two equilibrium states, it cannot be represented by a curve. In Fig. 2.9, we have schematically drafted such a process with a grey area, to indicate the lack of definition of the intermediate states.

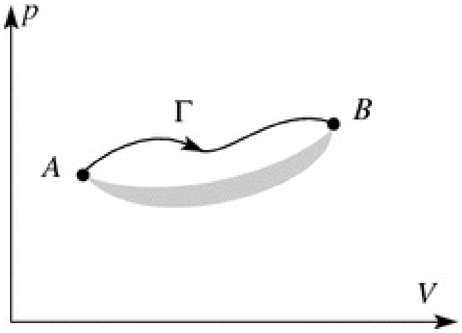


Fig. 2.9 A and B are equilibrium states of a hydrostatic system. Of the two processes, Γ is reversible, the *grey* area is not

---

2.5 Work

Any hydrostatic system has a definite volume bounded by surfaces that may be the surrounding walls of a container, the free surface of a liquid, etc. The system can interact with the surrounding environment by exchanging energy through its surfaces. External forces may be present and exert work, as defined in mechanics, on the system. The following convection is adopted in thermodynamics: work is positive if exerted on the system, negative if exerted by the system. The underlying reason for this is the fact that thermodynamics began historically as the science of engines.

The forces with which the system exchanges work with the surroundings act on its surfaces. They can do work only if their application points move, meaning that at least part of the surface must move. Movement of the surface is, however, not sufficient. To have work, the volume of the system must also change, as we shall now see.

Consider the simple system of Fig. 2.10, which is a gas contained in a cylinder with mobile piston of area S. Let p be the pressure and F_a the external force, perpendicular to the piston. Suppose that the friction between piston and cylinder is negligible. Consider a reversible expansion. The piston moves very slowly. The forces on the piston are F_a and the pressure of the gas Sp. The latter is larger than the former by a mere infinitesimal quantity. The work done by the pressure forces of the gas for an elementary displacement dx of the piston is

$$\delta W_g = pS\,dx = pdV \tag{2.8}$$

where, at the last member, we have taken into account that $S\,dx$ is the volume variation dV. The work of the external force F_a is

$$\delta W_e = F_a dx. \tag{2.9}$$

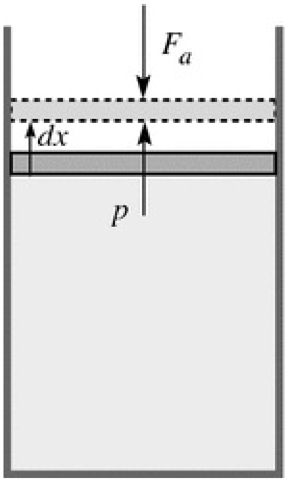


Fig. 2.10 Forces acting on the piston, in absence of friction

Notice that in both expressions, we have used the symbol δW rather than dW, because these are infinitesimal quantities but are not, in general, the differentials of any function, namely they are not exact differentials .

The two works we have expressed are equal in magnitude, because the two forces are equal. Notice, however, that if the friction is not negligible, and the process is quasi-static but not reversible, the two works are different. In this case, the work of the pressure force is larger, because it acts against the sum of F_a and the friction force (see Fig. 2.6). If the process is not even quasi-static, the gas goes through non-equilibrium states, in which the pressure is not even defined. Equation (2.8) loses its meaning. However, we can calculate the external work made by the gas on the surroundings even if the process is irreversible. Indeed, the motion of the piston very often takes place to produce a useful work (raising a weight, rotating the axis of an engine, etc.). In all these cases, an external force, F_a, acts against the motion and we can use Eq. (2.9) to calculate the work done. If friction is present, the work against the friction force must be added.

Within the mentioned limitations, the expression of the elementary work of the pressure force for a variation of the volume can be generalized to surfaces of any shape. Consider a hydrostatic system with pressure *p* enclosed in a surface *Σ* of arbitrary shape (see Fig. 2.11). Consider an infinitesimal quasi-static process in which the surface changes to Σ_1. The process being quasi-static, the pressure remains definite and independent of the position. Let *dΣ* be an infinitesimal surface element and *dn* the magnitude of its displacement, measured normally to the surface. The pressure force of the gas on the surface element is *p dΣ*. Its work is *p dΣ dn*. We obtain the total work by integration on the surface. *p* being constant, we have $\delta W = p\int_\Sigma dn\, d\Sigma$. As we see in the figure, the quantity $\int_\Sigma dn\, d\Sigma$ is the volume variation *dV*, and we thus get

$$\delta W_g = p dV. \tag{2.10}$$

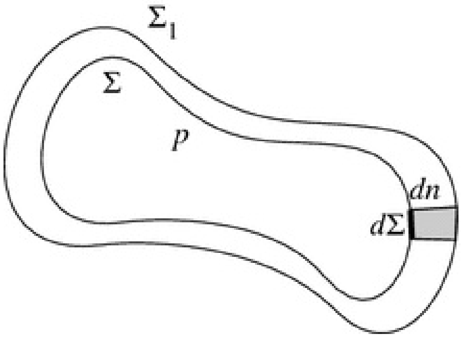


Fig. 2.11 Infinitesimal expansion of a hydrostatic system

The work done by the gas in a quasi-static process *Γ*, from state *A* to *B*, is obtained by integration

$$W_{AB;\Gamma} = \int_{A;\Gamma}^{B} p dV. \tag{2.11}$$

In general, the work depends not only on the initial and final states, but also on the particular process joining them.

Consider the quasi-static process *Γ* of a hydrostatic system from the initial state *A* to the final state *B*. In the plane *Vp*, the process is represented by an oriented curve, which we also call *Γ*, as in Fig. 2.12. In the process, the system goes through states having definite values of pressure and volume. We can then consider the pressure to be a certain function *p*(*V*) of the corresponding volume. The curve in the *Vp* plane represents this function. The work in the process is

$$W_{AB;\Gamma} = \int_{V_A;\Gamma}^{V_B} p(V)\, dV \tag{2.12}$$

where V_A and V_B are the volumes of the initial and final states. Graphically, the work is the area under the curve, which is grey in the figure. The work is positive if the final volume is larger than the initial one, and negative in the opposite case. We easily understand how, if the process between the same states is changed, the work changes too, in general. The work for a different process, as with Γ_1 in the figure, is the area under this curve, which is different from the area under Γ.

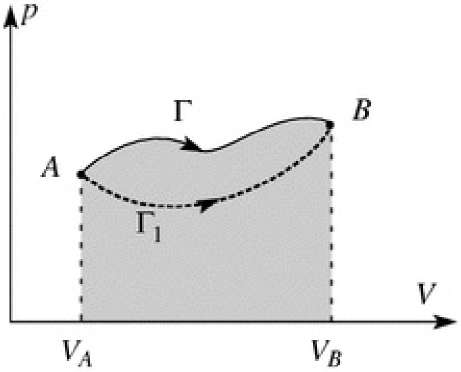


Fig. 2.12 Two quasi-static processes between the same initial and final states

Up to now, we have implicitly assumed the process to be such that the system has a certain state only once. Only in this case, the function $p(V)$ is single-valued. If this is not the case, it is always possible to divide the process into parts, for which p is the single valued function of V. Let us look at an important example.

In a cyclic process , the system starts from a state, goes through the process and then returns to its initial state. A cyclic quasi-static process of a hydrostatic system is represented by a closed curve in the pV plane. Figure 2.13 shows such a process starting from the initial state I and eventually coming back to it.

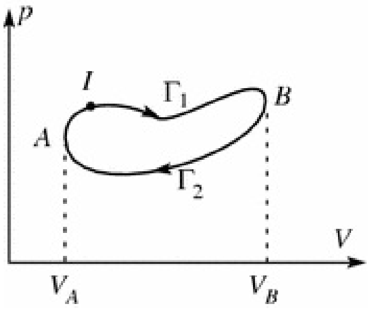


Fig. 2.13 Cyclic quasi-static process of a hydrostatic system

The work generated by the system in the cycle is, geometrically, the area enclosed within the cycle taken with the positive sign if the direction is clockwise, and negative if it is anti-clockwise. Let A and B be the states of minimum and maximum volume reached through the process, respectively. Let Γ_1 be the section of the process from A to B and Γ_2 the section from B to A. The work done in section Γ_1 is the area under this curve and is positive. The work done in section Γ_2 is the area under it and is negative. The total work is the difference between the two areas, corresponding to what we have stated. It is immediately understood that the work would have been negative if the sense was anti-clockwise. Notice that work is a definite quantity when the cycle is given; it does not depend on the initial and final state I.

Isochoric process . A process is said to be isochoric if the corresponding external work is zero. When the work can be expressed by Eq. (2.10), we must have $dV = 0$ for all the elements of the process, namely the isochoric processes must take place at constant volume. As a matter of fact, that is the meaning of the name (in Greek, *iso* means equal, and *choros* means space). However, the definition is more general. A process may also be isochoric when Eq. (2.10) does not hold.

Isobaric process. This is a process at a constant pressure.

Let us calculate the work of a hydrostatic system in an isobaric process. This is immediate. The pressure being constant, we can take it out of the integral and write

$$W_{AB} = \int_A^B p\,dV = p\,(V_B - V_A)\,. \tag{2.13}$$

Isothermal process . This is a process at a constant temperature.

The isothermal curves for a gas in the Vp plane are hyperbolas having asymptotes on the axes. Let us calculate the work of n moles of an ideal gas in an isothermal process from $A = (p_A, V_A, T)$ to $B = (p_B, V_B, T)$, represented in Fig. 2.14. Using the gas state equation Eq. (2.5), we have

$$W_{AB} = \int_A^B p\,dV = \int_A^B \frac{nRT}{V}dV = nRT\int_A^B \frac{dV}{V} = nRT\,(\ln V_B - \ln V_A)\,.$$

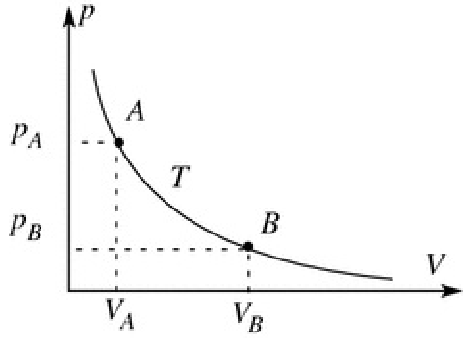


Fig. 2.14 Isothermal process in a gas

The last member looks strange, because the arguments of the mathematical functions should always be dimensionless. But it is just a matter of writing it. Indeed, the difference between two logarithms is the logarithm of the ratio of their arguments. It is better for us to write the expression in the form

$$W_{AB} = nRT\frac{V_B}{V_A}. \tag{2.14}$$

2.6 Heat

We have mentioned several times that if two systems at different temperatures are brought into thermal contact, their temperatures will vary until they become equal. In this type of process, the two systems exchange energy. The form of exchange is not work, because nothing is moved mechanically; rather, the form of exchange is heat . Heat is an energy exchange between two systems at different temperatures, as we shall now discuss.

We warn the reader that the common linguistic use of the word “heat” might generate confusion in this case. Indeed, the common language often uses the words “heat” and “temperature” interchangeably. In physics, however, they have very different meanings. We have already discussed temperature; we shall now discuss heat.

The physical dimensions of heat are the same as those of energy. However, heat is NOT energy. Similar to work, heat is an ENERGY EXCHANGE. Heat, as work, cannot be “owned” by a system. If one moves a book that is on the floor up onto a shelf, one does work on the book. Nobody would think that the book now “has” that work. It has potential energy, relative to the floor, which is equal to that work. In the common language, we say that friction “produces” heat. For example, brakes become hot when they act. But no heat has been produced. Even more to the point, no heat has been exchanged. Instead, the friction forces have done work, the kinetic energy of the car has decreased and the temperature of the brakes has increased. As we shall see, the temperature increase corresponds to an increase in the internal energy of the brakes.

Another example of confusion is sentences like: “I have been in Texas and the heat was unbearable”. What was unbearable was the temperature, not the heat. The very word “hot” tends to induce confusion, as it comes from heat. However, a body is hot not when it “has” a lot of heat, but when its temperature is high. In physics, we repeat, heat is always an exchange. To talk of heat, at least two systems must be present, not only one.

Let us now precisely define heat. We will give an operational definition, namely we define the set of operations needed to measure heat. The instruments used to measure heat are calorimeters . By definition, heat is the physical quantity measured with a calorimeter. We must thus define the calorimeter.

In thermodynamics, by convention, heat is positive if absorbed, negative if released by the system. The reader will note that the conventions on the sign for heat and work are opposite. This is a consequence of the fact that thermodynamics was developed to build engines able to absorb heat and produce work.

The calorimeter is itself a thermodynamic system. It is made of a substance, called the calorimetric substance , sealed inside a container. The substance is often pure water or ice or a mixture of the two. The masses of the substance are known, having been measured. To measure a heat exchange, we need a second system, a body that gives out or absorbs the heat to be measured. The exchange must take place with the calorimeter alone. Consequently, the calorimeter and the body must be enclosed in an adiabatic container, to forbid, as much as possible, any heat exchange with the surroundings.

We know that heat exchanges produce temperature changes. The simplest, preliminary choice is assuming the temperature change to be proportional to the heat exchange, at least for variations of the former that are not too large. We must then guarantee that no other thermodynamic variable actually varies but temperature. First, we shall work at constant volume. Consequently, we shall use a liquid or a solid as the calorimetric substance, whose volume does not vary much with temperature, and we shall avoid gases. Second, we shall work at constant pressure. Third, we shall avoid chemical reactions, by using pure substances.

Let us now build our calorimeter. As an adiabatic container, we shall use a vacuum flask, technically known as a Dewar flask . This type of vessel, called a thermos, greatly lengthens the time over which its contents remain hotter or cooler than its surroundings. The technique was invented by James Dewar (UK, 1842–1923) in 1892. It consists of two flasks, one inside the other, joined at the neck. The walls of the flasks are made of thin glass and are aluminized like mirrors. The gap between the flasks is evacuated. In this way, the different modes of heat transmission, convection, conduction and radiation (which we shall study in Chap. 6) are strongly reduced. The flask has an opening that can be closed with an insulating plug.

We introduce a certain quantity of pure water into the flask, of which we have measured the mass, m, and the temperature, T_i. The system that exchanges heat with the calorimeter can be, for example, a piece of metal that we heat at a temperature higher than T_i. We introduce the body into the calorimeter and close the plug. In the calorimeter, we also have a thermometer. We see the measured temperature increasing and finally becoming constant at a certain value T_f. The heat given out by the body and the heat absorbed by the water are equal because the system is isolated. We can say that this heat Q absorbed by the water is proportional to $T_f - T_i$.

As one can imagine, the temperature rise $T_f - T_i$ depends, for the same absorbed heat, on the mass of water. We find experimentally that the product of the water mass and the temperature difference $m(T_f - T_i)$ is proportional to the absorbed heat. We finally state that

$$Q = cm\left(T_f - T_i\right). \tag{2.15}$$

The proportionality constant c depends on the substance. If we had used, for example, an oil or an alcohol in place of water, we would have found different values. The constant is called the *specific heat* of the substance. The measurement unit for heat, which is the *kilocalorie* , is defined by fixing the specific heat of the water, to be exact, at a certain temperature.

The definition is as follows: *The kilocalorie is the heat quantity that must be given to a kilogram of water in order to increase its temperature from 14.5 to 15.5 °C, at the constant pressure of one atmosphere*.

The temperature of the exchange must be specified because the specific heat varies, even if not by much, with the temperature.

We note here that the calorimetric measurements are always very delicate. For example, the thermal insulation, even if good, is never perfect, and corrections must be applied to the measurements to take into account heat leaks. What is interesting here is the principle of operation of the calorimeter.

We finally note that, heat being the physical dimension of energy, its unit in the SI is the joule . We need the provisional calorimetric unit of the kilocalorie, for the following discussion.

2.7 Equivalence of Heat and Work

The equivalence between heat and work was established for the first time by Julius Robert von Mayer (Germany, 1814–1878) in a paper published in 1845. We shall specify here what “equivalence” means.

We have seen that both work and heat are forms of energy exchange. The two quantities are measured in completely different manners and a priori look very different. However, they are strongly connected. Indeed, engines absorb heat and produce work. Let us look at two examples.

In the first example (Fig. 2.15a), which is a system we have already considered, the gas is contained in a cylinder closed by a piston, on which there is a weight, producing the force F_a. Suppose now that we give heat to the system using a flame through the bottom of the cylinder, which is diathermic. We observe that the gas expands and the piston rises, elevating the weight with it. Work is produced. The process is the constant pressure expansion we considered in Chap. 5.

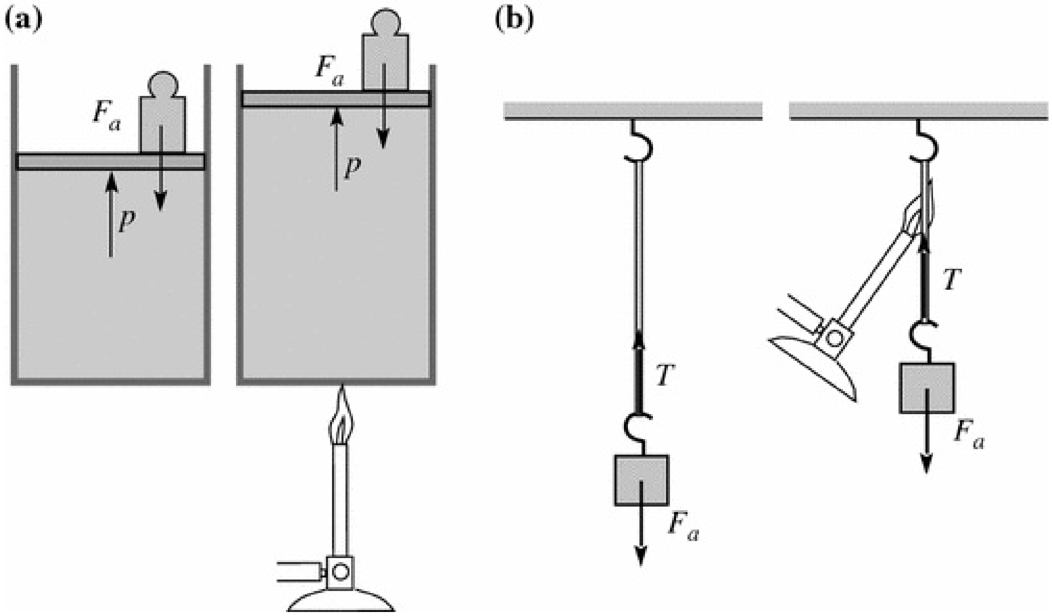


Fig. 2.15 **a** A gas in a cylinder, **b** a rubber band

In the second example, which we have also already considered, we have a weight hanging on a rubber band (Fig. 2.15b). If we gently heat the band with a flame, it contracts, because the elastic constant increases with the temperature. The weight rises. Again, the system absorbs heat and produces work, which is the product of the tension (that is constant) and the displacement.

We cannot conclude from these examples that the work produced by a system is equal to the heat it absorbs. As a matter of fact, both exchanges also vary the energy of the system. Here, we mean energy related to the internal, thermodynamic state of the system, which is different from the mechanical kinetic and potential energies. It is the *internal energy* of the system that we shall define in the next section. Before doing that, we shall prove its fundamental property: that the variation of the internal energy of a system is always equal to the algebraic sum of the heat received and the work done. This is the first law of thermodynamics.

Indeed, the first law of thermodynamics is the energy conservation law . In mechanics, the energy of an isolated system is conserved only if all the acting forces are conservative. The decrease of *mechanical* energy observed in the presence of dissipative forces does not really mean that energy is not conserved. The point is that the mechanical energy is not the only form of energy in the game. As a matter of fact, the energy that seems to be lost actually did transform into internal energy.

Clearly, what we have just stated needs to be experimentally proven. We shall now describe the elegant and fundamental experiment conducted by James Prescott Joule (UK, 1818–1889) in 1845. The experiment establishes what is known as equivalence between heat and work. More importantly, it establishes the existence of internal energy.

Figure 2.16 shows a schematic of the experiment, which is performed in two phases, (a) and (b) in the figure. A thermally-insulated vessel contains water (or another liquid). A vertical axis with a number of horizontal vanes is placed inside of the vessel with enough room so that it can rotate under the action of two external weights linked to a system of pulleys. Other vanes soldered onto the walls are interleaved with the mobile ones to hinder the common mode motion of the water. The water is forced to move slowly between the vanes, to maximize the mechanical energy dissipation in the liquid.

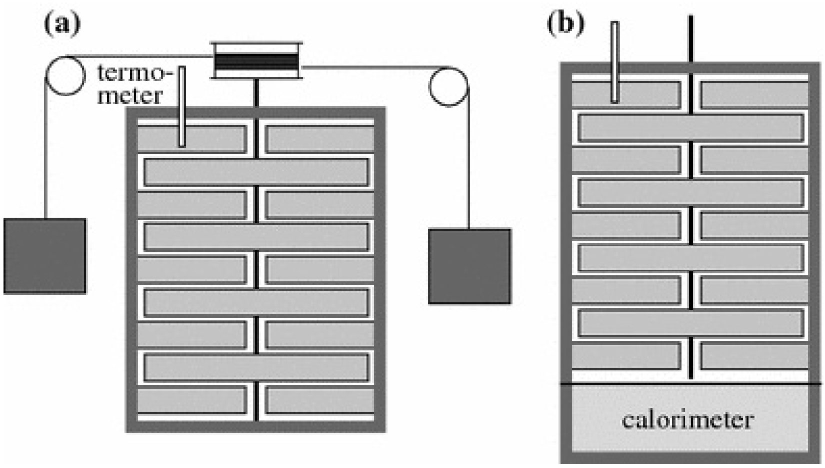


Fig. 2.16 The Joule experiment; (**a**) the first process; (**b**) the second process

We determine the initial state of the system by measuring its mass M and its temperature T_i. The mass of each weight is m.

In the first process, we perform the mechanical work W_1 on the system with no heat exchange ($Q_1 = 0$). We take the weights in their higher position and let them descend. They move very slowly and we can neglect their kinetic energy. We measure the drop h. The work done on the system (remember the sign convention) is $W_1 = -2mgh$.

In the second process, we bring the system back into its initial state. We do that extracting heat without any work being done ($W_2 = 0$). We take the thermal insulation off of the bottom of the vessel and we lay it on a calorimeter, as shown schematically in Fig. 2.16b. While the system gives away heat to the calorimeter, we measure its temperature, and we stop the process when it is back to the initial value T_i. Let Q_2 be the heat measured by the calorimeter, which is also the heat given away by the system. Overall, the process is a cycle.

The total work in the cyclic process is $W = W_1 + W_2 = W_1 + 0 = W_1$ and the heat exchanged is $Q = Q_1 + Q_2 = 0 + Q_2 = Q_2$. We take the ratio between the two quantities we measured and we find

$$\frac{W}{Q} = 4186\ \mathrm{J/kcal}. \tag{2.16}$$

Up to now, we have not proven anything. Indeed, the ratio of two quantities must have a value. However, if we repeat the experiment with different quantities of water, different liquids, different quantities of work, etc., provided that the process is cyclic, we find that the ratio of total work and total heat *always* has the value of Eq. (2.17), within the experimental uncertainties. The symbol for the ratio is J and we write that, for *every cyclic process* ,

$$J \equiv \frac{W}{Q} = 4186\,\mathrm{J/kcal}. \tag{2.17}$$

J is called the *mechanical equivalent of heat* . We can also state, in an equivalent manner, that experiments show that. In any thermodynamic cyclic process, the equation

$$JQ - W = 0 \tag{2.18}$$

holds, where W is the sum of all the works taken as positive if done by the system, negative if done on the system, and Q is the sum of all the heat exchanges, taken as positive if absorbed by the system, negative if released by it. Notice that we established the equation using irreversible processes. The equations we found, Eq. (2.19) in particular, hold for any cyclic process.

Having shown that heat and work are homogenous physical quantities, we can measure both of them in the same measurement unit. From now on, we shall measure the heat in joules. Equation (2.18) becomes

$$Q - W = 0. \tag{2.19}$$

2.8 First Law of Thermodynamics

In the previous section, we experimentally established an extremely important law of physics, the law of energy conservation .

Let us demonstrate that. Consider two different equilibrium states A and B of any thermodynamic system and two arbitrary processes 1 and 2 from A to B, as well as a third one 3, from B to A, as shown Fig. 2.17. The processes 1 + 3 and 2 + 3 are cyclic. We then apply Eq. (2.19) to both of them, obtaining $(Q-W)_1 + (Q-W)_3 = 0$ and $(Q-W)_2 + (Q-W)_3 = 0$. Subtracting the two relations, we have $(Q-W)_1 - (Q-W)_2 = 0$, or

$$(Q-W)_1 = (Q-W)_2\,.$$

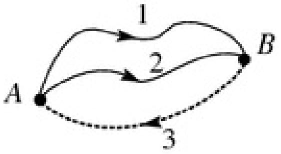


Fig. 2.17 Processes from A to B and from B to A

Given the arbitrary nature of processes 1 and 2, we can conclude that the quantity $Q - W$ depends only on the origin and the end of the process and not on the particular transformation. This quantity can be written as the difference between the values in the final and initial states of a state function which we call U:

$$\Delta U = U(B) - U(A) = Q - W. \tag{2.20}$$

State function means that U is a function of the thermodynamic coordinates (temperature, pressure, volume, chemical species concentration, tension for a rubber band, etc.). With $U(A)$, we mean the value of the function for the coordinates of the state A. U is the *internal energy* of the system. Equation (2.20) defines the internal energy a part of an additive constant, as is the case with energies.

Internal energy is a thermodynamic quantity. As such, it regards the macroscopic state of the system. Indeed, thermodynamics does not deal with the underlying microscopic physics. We shall come to that in Chap. 5 when we address statistical mechanics. We anticipate a few hints here. Consider a system made of only one chemical species. It is composed of a very large number of identical molecules, on the order of the Avogadro number. The molecules move at different speeds, which increase, on average, with increasing temperature. As a matter of fact, their mean kinetic energy is proportional to the absolute temperature. In addition, the molecules have interactions between them. The internal energy of the system is the sum of the kinetic energy and potential energies of its molecules (namely their mechanical energies). If the system is an ideal gas, the molecules are considered non-interacting and their energy is only kinetic. The molecules of a monoatomic gas can be considered point-like and their kinetic energies are those of their center of mass. If the gas is polyatomic, the kinetic energy about the centers of mass must be considered too. If more chemical species are present, the internal energy can vary when chemical reactions take place.

We now come back to thermodynamics. Equation (2.20) is the mathematical expression of the *first law of thermodynamic*s . In particular, if the system is isolated, both work and heat exchanges are zero and the internal energy is constant for any process.

Equation (2.20) has a completely general validity, for whatever process, provided the initial and final states are equilibrium states. Otherwise, we cannot define any function of the thermodynamic coordinates, because these coordinates are not defined.

Consider now the particular case of *quasi-static* , in which all the intermediate states, not only those that are extreme, are equilibrium states. Consequently, the internal energy is defined in the intermediate states as well, and Eq. (2.20) also holds for any elementary part of the processes, in the form

$$dU = \delta Q - \delta W. \tag{2.21}$$

Pay attention to the fact that, from the mathematical point of view, dU is the differential of a function (U), namely an exact differential, while δW and δQ are not. They are infinitesimal quantities, but there is no function for which they would be differentials. For this reason, we use the symbol δ rather than d.

We now consider the changes in internal energy for some relevant processes.

Isochoric processes . The work is zero by definition. The variation in internal energy is equal to the absorbed heat:

$$\Delta U = Q. \tag{2.22}$$

This was the case for the second process of the Joule experiment of the previous section. In general, we can state that the internal energy of a system increases through heating (namely giving it heat) at constant volume, and decreases through cooling at constant volume.

Adiabatic processes . The exchanged heat is zero. The variation in internal energy is the opposite of the work done by the system:

$$\Delta U = -W. \tag{2.23}$$

This was the case for the first process of the Joule experiment. In general, when the system does work on the surroundings $W > 0$, without heat exchange, as in an adiabatic expansion, its internal energy diminishes. Contrastingly, the internal energy increases in any adiabatic compression.

Example E 2.1

A body of mass m falls on the floor from a height h_1 and, after the collision, bounces back to the height h_2. Assuming that the body absorbs all the dissipated mechanical energy, how much does its internal energy vary?

The dissipated mechanical energy is $mg(h_2 - h_1)$. This energy is lost in the collision and corresponds to a mechanical work done on the body. Under the thermodynamic conventions, it is negative. Indeed, the heat exchanges during the brief instant of the collision can be neglected. Hence, for the first law, the variation of internal energy is $\Delta U = mg(h_1 - h_2)$.

As we shall see, the internal energy is an increasing function of temperature. Consequently, the final temperature of the body is higher than the initial one. If we wish, we can bring back the body to its initial state by extracting a quantity of heat equal to ΔU, that is, to the lost mechanical energy. One can find written instances stating that the process is a transformation of work into heat. This statement is wrong. The work has transformed into internal energy. □

Example E 2.2

We heat a room having volume V from temperature T_1 to temperature T_2. Knowing that the internal energy of a gas is $U = nc_VT$ (see Sect. 2.12), where c_V is a constant, how much does the internal energy vary? We must pay attention, because the quantity of air at the end of the process is different from that at the beginning. Indeed, even if the room is closed, some leakage under the windows and the doors will always be present. When the temperature of the air increases, the air expands and becomes less dense, the pressure remaining constant. We are dealing with a constant pressure process in an open system. Let n_1 and n_2 be the numbers of moles in the initial and final states, respectively. We write the gas equation as $pV = nRT$ and notice that p and V are constant during the process. Consequently, $n_1T_1 = n_2T_2$ and also $U_2 = U_1$. □

2.9 Specific Heats

If a body absorbs the heat δQ, its temperature increases by dT. The ratio of the two quantities is called the *heat capacity* of the body.

$$C = \frac{\delta Q}{dT}. \tag{2.24}$$

As already mentioned in Sect. 2.6, it has been experimentally found that the heat capacity of a given substance, under the same thermodynamic conditions, is proportional to its mass. We thus define it as *specific heat*, namely the heat capacity per unit mass. Thus, if the temperature of a body of mass m of that substance increases by dT when it absorbs the heat δQ, the specific heat of the substance is

$$c = \frac{C}{m} = \frac{1}{m}\frac{\delta Q}{dT}. \tag{2.25}$$

Note that heat capacity is a characteristic of a body, while specific heat is a characteristic of a substance.

The definitions we just gave are not sufficiently precise, because the heat absorbed by a body for a certain increase in its temperature depends on the process in which the absorption takes place. Consequently, a substance does not have just one, but indeed many, or even infinite, specific heats.

Consider the case of the hydrostatic system. Figure 2.18 represents two isothermal transformations of such a system, one at temperature T, and the other at $T + dT$. Starting from a state at temperature T, namely a point on that isothermal curve, we can pass on the second curve, namely increase temperature by *dT*, in an infinite number of different ways (see Fig. 2.18). For each of them, the heat exchange *δQ* is different. The system has infinite specific heats.

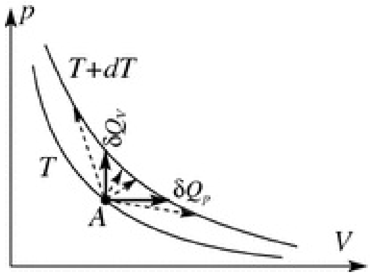


Fig. 2.18 Elementary heat exchanges at constant volume and at constant pressure

Particularly important are the specific heats at constant volume and constant pressure. In the particular case of a hydrostatic system, whose states are represented by points along the *pV* plane, all the other specific heats are linear combinations of those two (see Fig. 2.18). If the temperature of a mass *m* of a substance increases by *dT* when it absorbs the heat δQ_p or δQ_V, respectively, at constant pressure and constant volume, the specific heats of the substance are

$$c_p = \frac{1}{m}\frac{\delta Q_p}{dT}, \quad c_V = \frac{1}{m}\frac{\delta Q_V}{dT}. \tag{2.26}$$

Other useful quantities are the *molar heats*, which are the heat capacities of a mole of the substance. Namely, if the temperature of *n* moles of a substance increases by *dT* when it absorbs the heat δQ_p or δQ_V, the molar heats of the substance are

$$C_p = \frac{1}{n}\frac{\delta Q_p}{dT}, \quad C_V = \frac{1}{n}\frac{\delta Q_V}{dT}. \tag{2.27}$$

Consider now a hydrostatic system and suppose that the entire heat exchange takes place in quasi-static processes. The first law tells us that

$$dU = \delta Q - \delta W = \delta Q - p\,dV. \tag{2.28}$$

If the process is at a constant volume, $dV = 0$, and consequently $\delta Q_V = dU$. All the absorbed heat goes towards increasing the internal energy, and we can write for both the specific and the molar heat

$$c_V = \frac{1}{m}\left(\frac{dU}{dT}\right)_V, \quad C_V = \frac{1}{n}\left(\frac{dU}{dT}\right)_V \tag{2.29}$$

where the subscripts indicate that the derivatives are taken at constant volume *V*. This notation, often used in thermodynamics, is necessary because the internal energy (and other state functions) does not depend solely on the volume, but also on the other thermodynamic coordinates.

We now heat (i.e., give heat to) the system at constant pressure. The heat now goes in part towards increasing the internal energy and in part towards producing work, because the volume of the system varies. The first law is now

$$\delta Q_p = dU + p dV = d(U + pV) \tag{2.30}$$

where we could write the last member because the pressure is constant. We see that the heat is now equal to the differential of the function

$$H = U + pV. \tag{2.31}$$

This is another state function of the system, because U, p and V are such. This is called the *enthalpy* of the system. We can write Eq. (2.30) as

$$dH = \delta Q_p. \tag{2.32}$$

The expression tells us that the heat transferred to the system at constant pressure determines an equal variation of the enthalpy.

In practice, we often operate at atmospheric pressure and, consequently, the heat exchanges are at constant pressure. Some examples are the chemical reactions, which usually take place in open containers. In these cases, the reaction heat is the enthalpy variation from the initial to the final state.

In conclusion, the specific and molar heats at constant pressure can be expressed as derivatives of enthalpy:

$$c_p = \frac{1}{m}\left(\frac{dH}{dT}\right)_p, \quad C_p = \frac{1}{n}\left(\frac{dH}{dT}\right)_p. \tag{2.33}$$

We observe that the specific and molar heats at constant pressure are always larger than those, of the same substance, at constant volume:

$$C_p > C_V. \tag{2.34}$$

One might believe this difference to be due simply to the fact that, in the exchange at constant volume, all the heat goes towards increasing the internal energy, while at constant pressure, part of it goes towards external work, because the system expands. However, this is not so. Indeed, we also see that $C_p > C_V$ in the (very few) cases in which the substance contracts, rather than expands, through heating. This is the case for water between 0 and 4 °C. The property is a consequence of a general law of thermodynamics, the *Le Châtelier principle* , which we shall discuss in the next section.

2.10 Le Chãtelier's Principle

The *Le Châtelier principle,* or law of mobile equilibrium , was established by Henry Louis Le Châtelier (France, 1850–1936) in 1885. It allows us to foresee, in several cases and without any calculation, the direction in which a thermodynamic system will change in response to a change in the external conditions. The law states that, if the conditions external to a thermodynamic system in equilibrium change, the equilibrium of the system will change in the direction to oppose to the variation.

The law is useful in thermodynamics and chemistry. Suppose we have a container with different chemical species (*A*, *B*, *C*, *D*) that react with the reaction

$A + B \leftrightarrow C + D.$

The reaction can proceed in both directions. The chemical equilibrium is reached when the concentrations of the four substances remain constant. Suppose the reaction from left to right to be endothermic. In this case, if the temperature is increased, the equilibrium moves to the right side to increase the concentration of the species *C* and *D*. Indeed, this implies the absorption of heat, opposing the external change in this way (the temperature increase). The opposite happens for exothermic reactions.

Let us show that a hypothetical system that does not obey the Le Châtelier principle cannot have equilibrium states. Consider, for example, an exothermic reaction favored by an increase in the external temperature. In the presence of an increase in the temperature, even if very small, the reaction of the system would be an increase in released heat. This would cause, in turn, a further increase in the temperature and a further increase in the reaction process. The process would not stop until all the substances that were able to react have completely disappeared. Such behavior does not contradict any thermodynamic principle, but it does stand in contrast with the existence of equilibrium states. This argument should make clear the nature of the Le Châtelier principle. It is not a consequence of the laws of thermodynamics and, additionally, does not have the same importance. It is, however, useful, because it characterizes the states of stable equilibrium.

We now use the Le Châtelier principle to show that the specific heat at constant pressure is always larger than that at constant volume. Let us transfer a certain quantity of heat to the system at constant volume. Let dT_V be the corresponding temperature increase. As a consequence, the pressure of the system varies too, thus altering the equilibrium conditions. Notice that the pressure has increased if the system expands upon heating and has decreased if it contracts. The Le Châtelier principle tells us that the equilibrium of the system moves in opposition to the change in conditions. Namely, it has to return to the initial pressure. Its pressure must decrease in the former case, and increase in the latter. The process must imply a certain release of heat (it must oppose). This means that the change in temperature at constant pressure dT_p is less than that at constant volume dT_V, for the same heat quantity δQ. It follows that the specific heat at constant pressure is larger than that at constant volume.

We shall make further use of the principle in Sect. 4.3.

2.11 Solid Body

In this section, we consider a thermodynamic system consisting of a solid body. Its shape and volume are practically invariable. Approximately, we can neglect thermal dilatation. Thus, the only thermodynamic variable is the temperature. The internal energy is consequently a function of the temperature alone, $U(T)$. Consider a state A with temperature T_A and a generic state P with temperature T. For whatever process leading from A to P, the first law states that

$$Q - W = U(T) - U(T_A)\,. \tag{2.35}$$

The work of $p\,dV$ type is zero because the volume does not vary. In principle, other types of work might be done on the body, for example, by hammering or wiping. But we do not have any work of this type in the process we are considering. Then, the absorbed heat is equal to the variation in internal energy, and we have

$$Q = U(T) - U(T_A)\,. \tag{2.36}$$

If the process is quasi-static, we can write the same for all its infinitesimal elements:

$$\delta Q = dU. \tag{2.37}$$

If m is the mass and c the specific heat of the body, we have

$$\delta Q = mc\,dT\,. \tag{2.38}$$

Note that, as we have assumed the volume to be invariable, the body has only one specific heat, namely at constant pressure.

We can then write for the derivative of the energy with respect to the temperature

$$\frac{dU}{dT} = mc(T) \tag{2.39}$$

where we have explicitly written that the specific heat might be a function of temperature. We now obtain the internal energy by integration:

$$U(T) - U(T_A) = m\int_{T_A}^{T} c(T)\,dT\,. \tag{2.40}$$

In practice, the temperature dependence of the specific heat of several substances is modest, provided one considers temperature intervals of several degrees. Under these conditions, considering the specific heat *constant*, we have

$$U(T) - U(T_A) = mc(T - T_A). \tag{2.41}$$

We shall come back to the specific heats of solids in Sect. 5.3.

Let us now go back to the calorimeter we have considered in Sect. 2.6. It can be used to measure the specific heat of solid substances. Suppose, for example, we want to measure the specific heat c of copper. We proceed as follows. We heat a small block of pure copper of mass M at the temperature T_C, higher than that of the water in the calorimeter. We measure the water temperature T_i. We introduce the block into the water and close the plug. We read the temperature of the water and wait for it to stabilize at, say, T_f. Block and water now have the same temperature. Let us write down that the heat released by the block is equal to that absorbed by the water, namely

$$Q = m\left(T_f - T_i\right) = cM\left(T_C - T_f\right) \tag{2.42}$$

where we have remembered that the specific heat of water is equal to 1. From this equation, we have c. In practice, a number of measures that we did not mention are necessary. The most important is taking into account the heat absorbed by the walls of the calorimeter, by the thermometer and by any other auxiliary equipment that might be present in the calorimeter, referred to as the calorimeter water equivalent.

Table 2.1 reports the specific heats of several substances at 25 °C temperatures for several substances in J kg^{-1} K^{-1}.

Table 2.1 Specific heats of several materials at 25 °C in J kg^{-1} K^{-1}

Substance	**Spec. heat (kJ kg^{-1} K^{-1})**	**Substance**	**Spec. heat (kJ kg^{-1} K^{-1})**
Acrylic	1.4–1.5	Hydrogen	14.27
Aluminum	0.90	Ice (0 °C)	2.05
Argon	0.52	Iron	0.44
Beryllium	1.83	Lead	0.13
Bricks	0.85	Mercury	0.14
Calcium	0.65	Neon	1.03
Cesium	0.24	Nitrogen	1.04
Copper	0.39	Oxygen	0.92
Diamond	0.52	Platinum	0.13
Germanium	0.32	Silicon	0.70
Glass (crown)	0.67	Silver	0.24

Glass (flint)	0.50	Sulfur (yellow)	0.73
Gold	0.13	Tungsten	0.13
Graphite	0.71	Zinc	0.39
Helium	5.19	Wood	1.67

2.12 Internal Energy of the Ideal Gas

Consider a hydrostatic system. As we know, for a given mass, only two of the three thermodynamic variables, p, V and T, are independent. Consequently, its internal energy is, in general, a function of two variables. Even if their choice is arbitrary, it is usually convenient to choose volume and temperature. Indeed, at the microscopic level, the kinetic energy of the molecules is proportional to the temperature, as we have already mentioned. In addition, the potential energy of their interactions depends on the distances between molecules, whose mean value obviously depends on the volume. In conclusion, we shall write the internal energy as $U(V, T)$.

J. Joule conducted an experiment that gave important information on the internal energy of gases, known as the *free expansion experiment* . Figure 2.19 shows the scheme of the experiment. The two vessels A and B have metallic walls. Consequently, we can consider their volumes to be invariable. The vessels are connected by a tube that can be closed or opened with the tap R. Joule introduced a gas at high pressure into vessel A (up to 200 kPa) and pumped the air out of vessel B (R being closed, obviously). Opening R, the gas expands to fill both vessels. Notice that no external work is done in the expansion, as the gas expands in a space that was prepared to accept it. The expansion does not push anything. The external work is zero, because the total volume in the rigid surrounding walls does not vary. This process is called free expansion or expansion without external work.

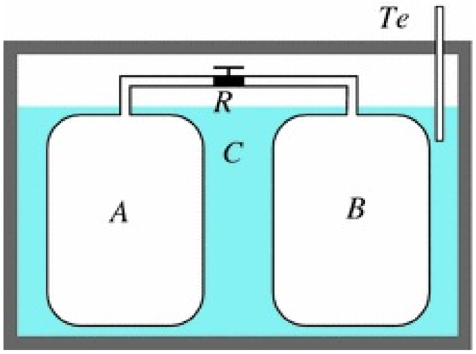


Fig. 2.19 The Joule free expansion experiment

The two vessels, whose walls, being metallic, were diathermic, had been lodged in a calorimeter. Joule did not measure any temperature change during the expansion. The following two conclusions can be extracted from the result: (1) the global heat exchange during the free expansion is zero; (2) the temperatures of the gas before and after the expansion are the same, considering that the gas is in thermal equilibrium with the calorimeter before and after the process.

We note here that the sensitivity of the experiment is rather limited, because the heat capacity of the gas is small. Joule worked with high initial pressures to partially reduce this limitation. However, in doing so, he moved away from the ideal gas behavior. The above conclusions should be regarded as valid only in a first approximation. We shall see in Sect. 4.4 that they are valid only for the ideal gas, as later established by Joule himself in more sensitive experiments performed in collaboration with W. Thomson.

Let us now go back to the consequences of the above observations. We can talk of infernal energy in the initial and in the final states because they are equilibrium states (the intermediate states are not so, but this does not matter). In the process leading from the former to the latter, both work and heat exchanges were zero. Consequently, $U\left(V_f,T\right) - U(V_i,T) = Q - W = 0$. Here V_i is the initial volume of the gas (one vessel), V_f is its final volume (two vessels) and T is the temperature that does not vary. The internal energy is the same in the two states, while their volumes are different. Consequently, the internal energy does not depend on the volume. It can depend only on temperature. For an ideal gas, we can write

$$U = U(T). \tag{2.43}$$

We can reach the same conclusion with a different argument. We start from the expression of the specific heat at constant volume Eq. (2.29)

$$C_V = \frac{1}{n}\left(\frac{dU}{dT}\right)_V \tag{2.44}$$

and integrate it between a reference state A and the generic state P, obtaining

$$U(T) - U(T_A) = n\int_A^P C_V dT. \tag{2.45}$$

Here, we need experimental input to know how the molar heat depends on temperature. Experiments tell us that it is almost constant (exactly constant for ideal gases) and we have

$$U(T) - U(T_A) = nC_V(T - T_A) = mc_V(T - T_A) \qquad (2.46)$$

where, to be complete, we have also included the expression in terms of the specific heat.

Let us now consider the other state function we have met, enthalpy. For the ideal gas, it also depends on temperature alone. It is

$$H = U + pV = U + nRT. \qquad (2.47)$$

Taking the derivative with respect to temperature and dividing it by the number of moles *n*, we have

$$\frac{1}{n}\frac{dH}{dT} = \frac{1}{n}\frac{dU}{dT} + R. \qquad (2.48)$$

There has not been any need to specify whether the derivatives are made at constant pressure or at constant volume, because neither enthalpy nor internal energy depends on these variables. We then recognize that the left-hand side is the molar heat at constant pressure and the first term in the right-hand side is the molar heat at constant volume, and can write

$$C_p - C_V = R \qquad (2.49)$$

This equation is called the Mayer's relation for ideal gases. It states that the difference between the molar heats at constant pressure and at constant volume has a universal value for all the gases, the gas constant, within the limits they can be considered as ideal.

We notice that in the case we are considering, the difference between the two heats is due to the fact that when the gas is heated at constant pressure, it performs an external work. Contrastingly, no external work is done in the heating at constant volume. In the latter case, all the heat goes towards increasing the internal energy, while in the former, part of it goes towards external work. As we know, however, this is not the reason why $C_p > C_V$.

A further experimental result is the dependence of the molar heats of the gases on the type of their molecules. All the monoatomic gases have the same molar heat at constant volume. It is (almost) independent of temperature down to rather low temperatures. Its value is

$$C_V = \frac{3}{2}R = 12.5\,\mathrm{J\,mol^{-1}\,K^{-1}} \quad \text{monoatomic gas.} \qquad (2.50)$$

The situation of the diatomic gases is more complicated. However, the larger fraction of these gases (especially those with less massive molecules), at ambient temperature and in a rather wide interval of temperatures, have molar heats around the value

$$C_V = \frac{5}{2}R = 20.8\,\mathrm{J\,mol^{-1}\,K^{-1}} \quad \text{diatomic gas.} \tag{2.51}$$

The simplicity of the relations just determined cannot be casual. As a matter of fact, they are deeply rooted in statistical mechanics, as we shall see in Sect. 5.2.

Let us use them together with the Mayer's relation. We obtain for the molar heats at constant pressure

$$C_p = \frac{5}{2}R \text{ monoatomic gas;} \quad C_p = \frac{7}{2}R \text{ diatomic gas.} \tag{2.52}$$

In addition, defining the dimensionless parameter

$$\gamma = \frac{C_p}{C_V} \tag{2.53}$$

we have

$$\gamma = \frac{5}{3} = 1.67 \text{ monoatomic gas;} \quad \gamma = \frac{7}{5} = 1.40 \text{ diatomic gas.} \tag{2.54}$$

The above conclusions would be rigorously valid for the ideal gases and approximately so for many real gases, within large intervals of the thermodynamic variables. We shall study the real gases in Chap. 4. Here, we report the molar heats of some gases at room temperature in Table 2.2.

Table 2.2 Molar heats of some gases at 25 °C in J $\mathrm{mol^{-1}\,K^{-1}}$

		C_p	C_V	$C_p - C_V$	C_p/C_V
Monoatomic	Ar	20.8	12.5	8.3	1.67
	He	20.9	12.5	8.3	1.67
Diatomic	Air	29.1	20.8	8.3	1.40
	N_2	29.1	20.8	8.3	1.40
	O_2	29.4	21.1	8.3	1.40
	H_2	28.8	20.4	8.3	1.41
Polyatomic	CO_2	37.0	28.5	8.5	1.30
	NH_3	36.8	27.8	9.0	1.32
	CH_4	35.6	27.2	8.4	1.31

2.13 Adiabatic Processes in Gases

We shall now discuss the quasi-static adiabatic processes of an ideal gas.

For example, we can expand or compress a gas adiabatically and quasistatically by enclosing it in a cylinder with a piston. All the surrounding surfaces should be adiabatic. We shall move the piston very slowly in one or the other direction. In expanding, the gas delivers external work. In the absence of a heat exchange, all the work done corresponds to a decrease in internal energy. Internal energy being a function of temperature, the gas cools down. Contrastingly, temperature increases in adiabatic compression. You can feel that when you inflate the tire of a bike. We now seek the quantitative relation between temperature and volume.

Consider, for simplicity, a mole of gas. We apply the first law to an infinitesimal section of the process, in which, clearly, $\delta Q = 0$. We have

$$dU + pdV = 0.$$

Using Eq. (2.44) and eliminating p with the state equation, we have

$$C_V dT + \frac{RT}{V} dV = 0$$

namely

$$\frac{dT}{T} + \frac{R}{C_V}\frac{dV}{V} = 0.$$

Integrating, within the limits in which we can consider C_V to be constant, we get

$$\ln T + \frac{R}{C_V}\ln V = 0.$$

We take the exponential of this expression and obtain

$$TV^{R/C_V} = \text{constant}.$$

We want this expression in terms of the ratio γ of the specific heats, Eq. (2.53), which we write as

$$\gamma = \frac{C_p}{C_V} = \frac{C_V + R}{C_V} = 1 + \frac{R}{C_V}.$$

Hence, the expression we have found is just

$$TV^{\gamma-1} = \text{constant}. \tag{2.55}$$

This expression tells us how much a gas cools down in an adiabatic expansion. Consider, for example, a quantity of air, whose volume doubles adiabatically. The air is mainly N_2 and O_2, two diatomic gases. Hence, $\gamma = 1.4$. The temperature decreases by the factor $(1/2)^{0.4} = 0.76$.

Equation (2.55) can be written in two other equivalent forms, which are sometimes useful, using the sate equation $pV = RT$. As immediately found, they are

$$pV^{\gamma} = \text{constant} \tag{2.56}$$

and

$$Tp^{1/\gamma-1} = \text{constant}. \tag{2.57}$$

Equation (2.56) is the equation of the adiabatic transformations in the pV plane. Let us compare it with the equation of the isothermal processes, namely pV = constant. The adiabatic curves are similar to the isothermal curves, but are steeper because, in any case, $\gamma > 1$.

Figure 2.20 shows two isothermal (dotted) and two adiabatic (continuous) curves.

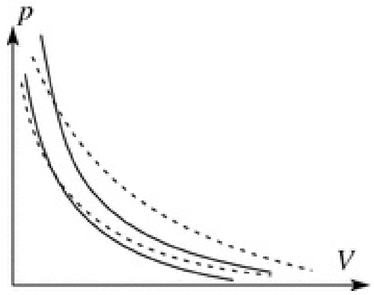


Fig. 2.20 Figure shows two isothermal (*dotted*) and two adiabatic (*continuous*) curves

---

2.14 Compressibility and Thermal Expansion

Gases are easy to compress and to expand. A small pressure increase is sufficient to change their volume considerably. The same is not true for condensed bodies, namely solids and liquids. However, even those are compressible, if subject to large enough pressures. As the volume changes if the temperature also varies, compressibility must be defined at constant temperature.

The isothermal compressibility is defined as

$$\kappa = -\frac{1}{V}\left(\frac{dV}{dp}\right)_T. \tag{2.58}$$

Note that the derivative is negative, because the volume diminishes when the pressure increases. The minus sign in the definition is used to have κ be positive. The dimensions of κ are the reciprocal of a pressure, as immediately seen in Eq. (2.58).

Let us consider some orders of magnitude. The isothermal compressibilities of liquids are in the range of 10^{-11}–10^{-10} Pa^{-1}. Suppose we want to have a relative variation of volume of one per cent, namely $dV/V = 10^{-2}$. If $\kappa = 10^{-10}$ Pa^{-1}, the change of pressure must be 1 MPa, which is about ten times the atmospheric pressure. For example, for water, $\kappa = 5 \times 10^{-10}$ Pa^{-1}, and for mercury, $\kappa = 0.4 \times 10^{-10}$ Pa^{-1}.

The isothermal compressibility of the large majority of solids is even smaller. For example, it is $\kappa = 0.6 \times 10^{-11}$ Pa^{-1} for iron, and $\kappa = 1.4 \times 10^{-11}$ Pa^{-1} for aluminum.

Let us now consider the gases. The state equation can be written as $V = nRT/p$. By derivation, we obtain

$$\kappa = 1/p. \tag{2.59}$$

At atmospheric pressure, the compressibility of a gas is $\kappa = 10^{-5}$ Pa^{-1}, about one million times more compressible than a condensed body at the same pressure.

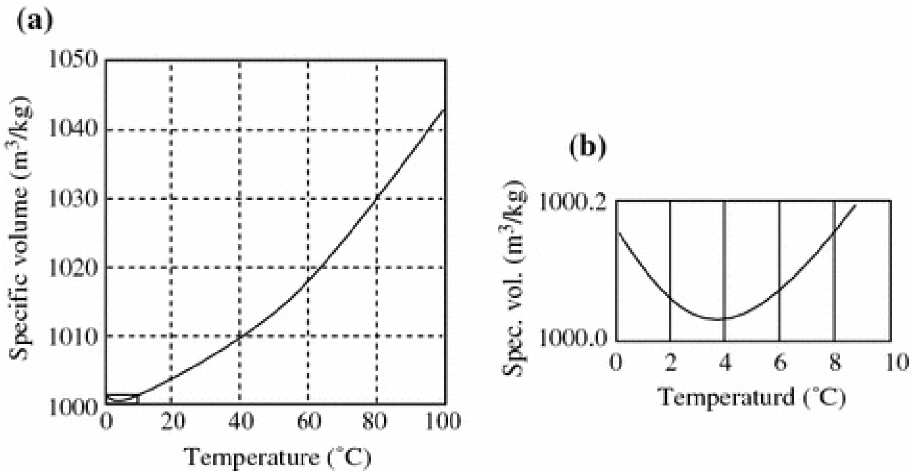


Fig. 2.21 Specific volume of water

Another quantity employed to characterize the thermal properties of the bodies is the *volumetric thermal expansion coefficient at constant pressure*, defined by the relation

$$\alpha = -\frac{1}{V}\left(\frac{dV}{dT}\right)_p. \tag{2.60}$$

The physical dimensions of the coefficient are the reciprocal of temperature, and are measured in K^{-1}. The large majority of substances expand when heated, hence α is positive. However, there are exceptions, that is, substances that expand when cooled. Their α is negative. The most common of these is water between 0 and 4 °C, while others include graphene, some complex compounds, some iron alloys, etc. Cubic zirconium tingstenate (ZrW_2O_8) has a negative expansion coefficient in the largest temperature range, namely all of them up to its fusion.

Figure 2.21 shows the specific volume (which is the inverse of the density) of water as a function of temperature in °C. The part in the small box near 0 °C in part (a) of the figure is enlarged in part (b).

We again fix the orders of magnitude. The thermal expansion coefficients of liquids are on the order of 10^{-4}–10^{-3} K^{-1}; for example, at ambient temperature, it is $\alpha = 2.1 \times 10^{-4}\ \mathrm{K}^{-1}$ for water and $\alpha = 1.1 \times 10^{-3}\ \mathrm{K}^{-1}$ for alcohol. The coefficients of solids are smaller, typically by one order of magnitude. For example, iron has $\alpha = 3.5 \times 10^{-5}\ \mathrm{K}^{-1}$. In some cases, one needs to build structures whose dimensions vary as little as possible in regard to temperature. Special alloys have been developed for this very purpose. For example, invar , an iron-nickel alloy, has $\alpha = 3.6 \times 10^{-6}\ \mathrm{K}^{-1}$.

Let us now compare these with gases. We write the state equation as $V = nRT/p$, and take the derivative with respect to temperature, obtaining

$$\alpha = 1/T. \tag{2.61}$$

Hence, under normal conditions, $T = 293$ K, and a gas has $\alpha = 3.4 \times 10^{-3}\ \mathrm{K}^{-1}$, which is not much larger than for some liquids.

Problems

2.1 Let us introduce into a calorimeter, which is at a temperature of 80 °C, 300 g of water at a temperature of 20 °C (and nothing else). We observe that equilibrium is reached when the water temperature is 60 °C. Can we find the heat capacity of the calorimeter from these data?

2.2 The air in two rooms of equal volume has the same pressure. If the temperatures are different, which room contains more air?

2.3 If we know the pressure, temperature and volume of a gas, which of the following quantities can be known: the type of gas; its number of molecules; its number of atoms?

2.4 Find the number n_p of molecules per unit volume of air at standard temperature and pressure.

2.5. Find the mass of a cubic meter of air at STP.

2.6 The air composition on the earth's surface is the following: N_2 78.08 %, O_2 20.95 %, Ar 0.93 % and other gases 0.04 %. Find the molar mass of air.

2.7 Consider the following processes of an ideal gas: (a) volume increases and pressure increases proportionally; (b) volume increases and pressure is constant; (c) the gas expands isothermally; (d) the gas expands adiabatically; (e) volume increases and pressure decreases more rapidly than in an adiabatic process. Qualitatively draw the representative curves in the Vp plane and state the behavior of internal energy in each case.

2.8 A mixture of hydrogen and oxygen gases is enclosed in a container with rigid and adiabatic walls. We fire a spark (its heat release is negligible) and the gases violently react, resulting in an increase in pressure and temperature. How much does the internal energy vary?

2.9 Suppose we want to perform the Joule experiment on the equivalence between heat and work with a calorimeter containing one liter of water ($m = 1$ kg). Each of the two weights has a mass $M = 10$ kg and the drop is $h = 2.5$ m. What is the temperature increase to be measured?

2.10 The heat capacity of a body, in the considered temperature interval, depends on temperature according to the expression $C = 10 + 0.002T + 3 \times 10^{-5}T^2\,\mathrm{J\,K^{-1}}$. How much heat is released when the temperature varies from $T_1 = 400$ K to $T_2 = 300$ K?

2.11 A system transforms from state A to state C once by process 1, once by process 2, as shown Fig. 2.22. Find the difference $Q_1 - Q_2$ between the exchanged heats (necessary data are in the figure).

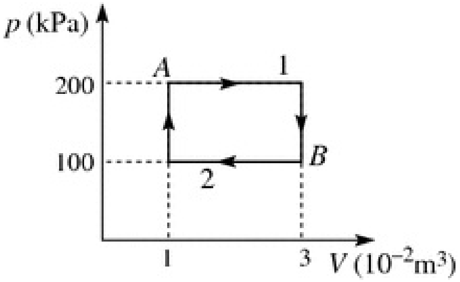


Fig. 2.22 The two processes of problem 2.11

2.12 A certain quantity of a gas passes from state 1 with internal energy $U_1 = 500$ kJ to state 2 with $U_2 = 100$ kJ, doing the work $W = 200$ kJ. How much is the absorbed heat Q if (a) the process is reversible, or (b) the process is irreversible?

2.13 A certain quantity of an ideal monoatomic gas is compressed adiabatically, reversibly changing its pressure from p_1 to $p_2 = 10\,p_1$. After that, the gas is compressed back to its initial volume, in an isothermal, reversible process, to the pressure p_3. Find p_3/p_1.

2.14 A certain quantity of an ideal monoatomic gas expands in a reversible process at constant pressure p from the volume V_1 to V_2. Find the expressions for: (a) the internal energy variation ΔU, (b) the work done by the gas W and (c) the absorbed heat Q.

2.15 One mole of an ideal gas heats from $T_1 = 273$ K to $T_2 = 373$ K in a reversible constant pressure process, absorbing the heat $Q = 6.65$ kJ.

Determine: (a) the specific heat ratio γ, (b) the internal energy increase ΔU and (c) the work done W. (See problem 2.14).

2.16 A certain quantity of an ideal gas expands at constant temperature from the volume $V_1 = 1\ \text{m}^3$ to $V_2 = 2\ \text{m}^3$. The final pressure is $p_2 = 100$ kPa. Find: (a) the internal energy increase ΔU, (b) the work done W and (c) the absorbed heat Q.

Footnotes

[1] Notice that the name is “kelvin”, not “degree kelvin” and that the symbol is K, not °K.

Alessandro Bettini, A Course in Classical Physics 2—Fluids and Thermodynamics, Undergraduate Lecture Notes in Physics, DOI 10.1007/978-3-319-30686-5_3

3. The Second Law of Thermodynamics

Alessandro Bettini[1] ✉

(1) Dipartimento di Fisica e Astronomia, Università di Padova, Padova, Italy

✉ **Alessandro Bettini**
Email: alessandro.bettini@pd.infn.it

The first law of thermodynamics is the law of energy conservation. The second law has a very different nature; it establishes the rules for the irreversibility of the natural processes. If we bring a pendulum out of its equilibrium position and let it go, its oscillations are ample at the beginning, but gradually decrease in their amplitude and finally stop after a shorter or longer period. The energy, which was initially mechanical energy, has not been lost; rather, it became internal energy of the pendulum and of the surrounding air. The first law does not forbid the inverse process, namely having a pendulum at rest starting oscillations of increasing amplitude, while its temperature and that of the air decrease. Similarly, if we leave a pot of hot coffee on a table, the liquid cools down in time, while the air heats up (not by too much, obviously). We never observe coffee at room temperature heating up while the air cools down. As we shall see, the second law forbids both types of phenomenon.

As we already mentioned, the second law was discovered before the first, in the historic period during which engineers were developing thermal engines, namely devices able to perform mechanical work using the heat produced by combustion. In this case, the work of the engineers, theoretical interpretations included, anticipated that of the physicists.

In Sect. 3.1, we state the second law. Both for historical and didactic reasons, we shall give two statements, one attributed to Clausius, one to Lord Kelvin, and then prove their equivalence.

All engines operate on cyclic, rather than open, processes. Indeed, an open process can be performed only once, while a cyclic process can continue indefinitely. The simplest cycle compatible with the second law exchanges heat with two sources. This is called the Carnot cycle, and we shall study it in the subsequent three sections, together with the fundamental Carnot theorem.

In Sect. 3.5, we shall introduce the concept of thermodynamic temperature, which, as we anticipated, allows for extending the scale down to absolute zero.

In Sect. 3.6, we shall demonstrate the fundamental Clausius theorem, which leads to the definition of a state function, entropy. This is the function directly connected to the second law, just as energy is connected to the first one. The second law is expressed, in its most precise form, as the law of increasing entropy. The statement is that the entropy of a thermally-isolated system increases if the system performs spontaneous processes; it stays constant if the system is in an equilibrium state, which is a state of maximum entropy.

In Chap. 5, we shall see how entropy is a measurement of the disorder of the internal structure of the system.

3.1 The Second Law of Thermodynamics

The first law of thermodynamics states that the total energy of an isolated system is constant. A non-isolated system can exchange energy as work or heat, the two forms of exchange being completely equivalent. For the first law, work can be transformed into heat and heat into work in a completely symmetric fashion. However, this symmetry is not present in nature. The second law quantitatively establishes how this happens.

As a matter of fact, the complete transformation of work into heat is always possible. For example, we can hammer a piece of metal, increasing its temperature and its internal energy, and then take it back to its initial temperature, extracting heat. We have completely transformed work into heat. Contrastingly, it is never possible to build a cyclic engine able to extract heat from a body and transform it completely into work. If such an engine were possible, we would have at our disposal practically infinite energy sources, given the enormous internal energy of the oceans and of the ground.

We observe a second type of asymmetry in the passage of heat between two bodies at different temperatures. A hot water pot gradually cools down until its temperature is equal to that of the room in which it sits. But it never happens that water at room temperature heats up, taking energy from the cooler environment, even if the process would not violate the first law.

The second law of thermodynamics precisely states both types of irreversibility. We shall come to that after having given a few definitions.

A *thermal reservoir* , or *thermal bath* or *heat source*, at temperature T is a thermodynamic system having all its points at that temperature, able to exchange heat but not work with another system brought into thermal contact. In addition, the heat capacity of the reservoir is large enough that, in all the thermal exchanges we shall consider, its temperature remains effectively constant. It is an effectively infinite pool of thermal energy at a given, constant temperature.

Which are the thermodynamic coordinates defining the state of a reservoir? At first sight, one might think that its thermodynamic state should not vary when we extract or inject a certain quantity of heat Q, because its temperature effectively remains constant. This argument is, however, wrong. To understand the point, let us start by considering a source of very large, but not infinite, heat capacity. The heat Q will produce a certain change of temperature ΔT. If m is the mass of the source and c its specific heat (hence the heat capacity is mc), it is $Q = mc\Delta T$. The reservoir is now in a different state, even if one near to its initial state. In particular, its internal energy has increased exactly by Q because there was no external work. Let us now increase the heat capacity to the infinite (the mass of the reservoir), keeping Q fixed. The temperature variation ΔT goes to zero, but the variation of internal energy remains constant, equal to Q. The two states remain different. In conclusion, the thermodynamic state of a reservoir, or pool, is defined by a sole coordinate, the internal energy .

A *heat engine*, or simply an *engine*, is a device made of mechanical parts (like pistons, pulleys, belts, gears, etc.) and a vessel containing a fluid (such as the water vapor in a steam engine or the gasoline air mixture in a car engine). We are interested in the fluid , its thermodynamic states and processes, its heat exchanges with the reservoirs and its external work. Consequently, we can ignore the details of the mechanical structures. The engine must operate continuously. Consequently, it must come back to its initial state periodically. The processes of the fluid will always be cycles. We shall call the engines that produce work using heat (thermal) *motors* and the engines that transfer heat from a cooler to a hotter reservoir, using work, *refrigerators*.

In practice, the cycle of every engine is always an irreversible process. However, from the theoretical point of view, it is often useful to consider reversible processes. Notice that, independently of reversibility or not, the internal energy in a cycle does not vary. Consequently, the total work done is always equal to the total absorbed heat (meaning that absorbed less that released). Going forward, when we talk of absorbed heat and external work done, we shall mean exchanged and done per cycle. Finally, we shall adopt the usual and already-mentioned convention in regard to the signs; the heat absorbed by the engine is positive, the heat released is negative; the work done is positive, the work received is negative.

We shall now give the two statements of the second law and then prove their equivalence. The two statements have been given by Rudolf Clausius (Germany, 1822–1888) and William Thomson, Lord Kelvin (UK, 1824–1907) In Sect. 3.10, we shall give a third statement.

Clausius statement . *No process is possible whose* ***sole*** *result is the transfer of heat from a cooler to a hotter reservoir*.

We stress the importance of the adjective "sole". Refrigerators, for example, transfer heat from their inside, which is cooler, to the outside environment, which is hotter. But they also absorb work. If one takes the plug out, they do not work any more (Fig. 3.1a).

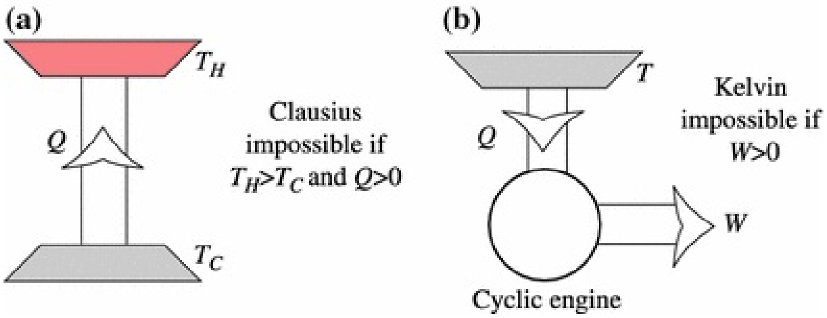


Fig. 3.1 Cartoons for the Clausius and Kelvin statements of the second law

Kelvin statement . *No process is possible whose* ***sole*** *result is the absorption of heat from a reservoir and the conversion of all this heat into work.*

Once more, the adjective "sole" is important. Consider, for example, the isothermal expansion of an ideal gas. The internal energy does not vary, because it is a function of the temperature alone. Consequently, all the absorbed heat is converted into work. But this is not the sole result; indeed, the final state is different from the initial one. The Kelvin statement refers to cyclic processes (Fig. 3.1b).

We shall now prove the equivalence of the two statements. We shall show that if the Kelvin statement was false, then the Clausius statement would be false as well, and reciprocally.

Assume the Kelvin statement to be false. We can then transform all the heat, Q, taken from a single pool at temperature, say T_C, into the work W, as shown in Fig. 3.2a. For the first law, $Q = W$. We can now, without violating the first or second laws, build a refrigerator, namely the engine in Fig. 3.2b that employs the work W, absorbs the heat Q from the pool at T_C, and delivers the heat Q_H to a reservoir at temperature $T_H > T_C$. For the first law, $Q_H = Q + Q_C$. This is positive. Hence, we have a process for which the sole result is the transfer of heat from a cooler to a hotter body. The engine composed of the two engines is a refrigerator working without external work; the Clausius statement would be false.

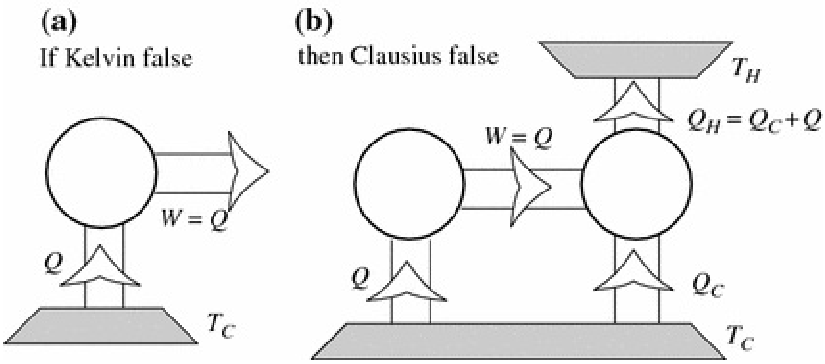


Fig. 3.2 Cartoon representation of the first part of the equivalence between the two statements

We shall give the second part of the demonstration of the equivalence between the two statements after having discussed the simplest process available for transforming absorbed heat into work, the focus of the next section

3.2 The Simplest Heat Engine

For the Kelvin statement, there is no cyclic process capable of transforming the heat absorbed from a single reservoir into work. The simplest possibility is consequently a cyclic process that exchanges heat with only two reservoirs, at different temperatures. Let us call T_C the temperature of the cooler pool, T_H the temperature of the hotter one ($T_H > T_C$). The following arguments are valid for whatever fluid and for both reversible and irreversible cycles. However, to be concrete, we shall consider an engine operating with a gas, whose state can be represented in the *Vp* plane. However, our argument will be general; in particular, we shall not use the gas equation.

Consider a gas in a cylinder, closed off by a piston. The lateral walls and the piston are adiabatic; the bottom is diathermic. We start by putting the cylinder on the (hot) reservoir at the temperature T_H. The initial state is the point *A* in Fig. 3.3 and is represented in Fig. 3.4, labeled with the same letter. We let the gas expand, isothermally reaching state *B*. This and the following processes are not necessarily reversible. When the gas is in state *B*, we move the cylinder on an adiabatic base. We perform a second expansion, now adiabatically. The temperature of the gas decreases. We stop the expansion when the temperature is at T_C (state *C*). We now move the cylinder onto the (cold) reservoir at temperature T_C and compress the gas until we reach state *D*, which is on the same adiabatic curve as *A*. Finally, we move the cylinder on the adiabatic base and compress the gas adiabatically until we reach temperature T_H. We are back to the initial state; the cycle is complete.

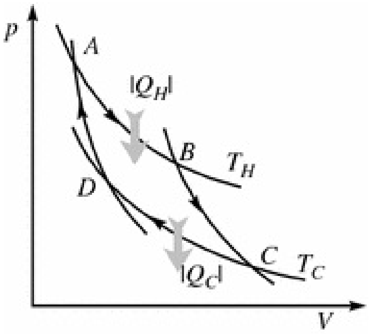


Fig. 3.3 The simplest cycle

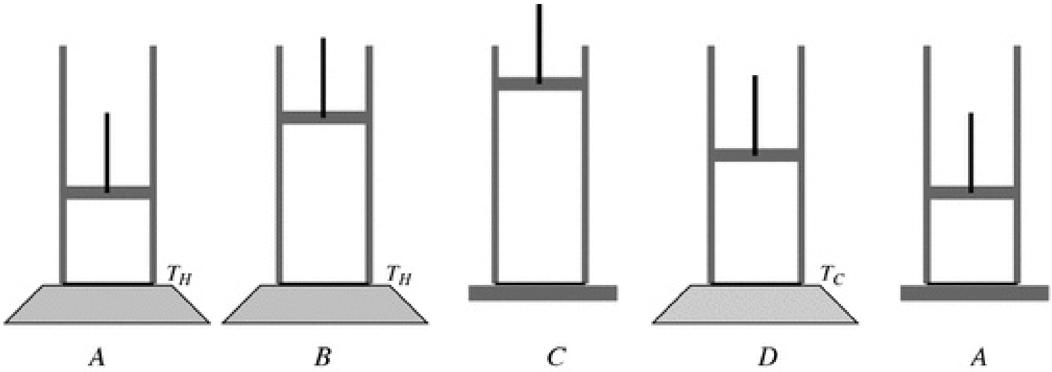


Fig. 3.4 The states of the Carnot cycle

In the isothermal expansion, the gas has absorbed the heat Q_H from the hot reservoir. With our sign convention, it is positive. In the isothermal compression, the gas releases the heat $|Q_C|$. With the sign convention, Q_C is negative.

The total absorbed heat is $|Q_H| - |Q_C| = Q_H + Q_C$. The work *W* done by the gas is positive, because the cycle is clockwise. For this reason, the reservoir into which the heat is released must be the cold one. For the first law, the total absorbed heat is equal to the work done:

$$W = |Q_H| - |Q_C|.$$

The process being a cycle, there is obviously no variation of the internal energy. We see that only a portion of the heat absorbed from the hot reservoir is transformed into heat, while a portion is delivered to the cold reservoir.

The *efficiency* of an engine operating between two reservoirs is defined as the ratio between the work done and the heat absorbed from the hot reservoir:

$$\eta = \frac{W}{|Q_H|} = \frac{|Q_H| - |Q_C|}{|Q_H|} = 1 - \frac{|Q_C|}{|Q_H|}. \tag{3.1}$$

The efficiency is a pure number, positive and less than one. It would be one, namely 100 %, only if it were possible not to deliver any heat, i.e., if $Q_C = 0$.

We can now complete the demonstration of the equivalence of the two statements. We now assume the Clausius statement to be false. We can then transfer a certain quantity of heat Q from a cold reservoir at temperature T_C to a hotter one at temperature T_H. We use a cycle, respecting the first and second laws, that absorbs that heat Q from the hot reservoir and releases a portion of it into a cold reservoir. The cycle does positive external work. The sole result of the complete process is the transformation into work of heat taken from a unique source (which is in this case the cold reservoir). The hot source, indeed, absorbs and delivers the same heat. Its internal energy, namely its state, does not vary. In conclusion, the global process violates the Kelvin statement (Fig. 3.5).

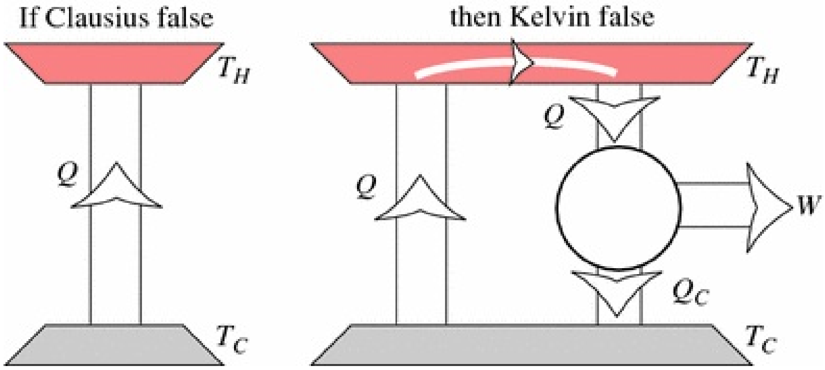


Fig. 3.5 Cartoon representation of the second part of the equivalence between the two statements

3.3 The Carnot Cycle

The *Carnot cycle* is a cycle that exchanges heat with *only two* reservoirs when all its processes are *reversible*. It is named after Nicolas Léonard Sadi Carnot (France, 1796–1832), whom we have already met as one of the founders of thermodynamics. The fluid in the cycle can be of any type; it is not necessarily a gas. In this section, we shall express the efficiency of the Carnot cycle for an ideal gas and, in the next section, demonstrate the validity of the found expression for every Carnot cycle.

Figure 3.3 is the diagram of a Carnot cycle for an ideal gas in the *pV* plane and Fig. 3.4 shows the sequence of operations necessary to perform the cycle. The four processes must now be reversible. In practice, we reduce the friction between piston and cylinder as much as possible and move the piston very slowly; all the intermediate states should be almost equilibrium states. Under these conditions, the external and internal pressures are always equal and, in the isothermal processes, the external and internal temperatures are equal too. Notice that no part of the latter condition is true for irreversible processes. The pressure and temperature of the gas cannot even be defined in this case. We also recall that heat and work exchanged in a reversible process are equal and opposite to those exchanged in the inverse process.

The efficiency of the Carnot cycle is given by Eq. (3.1), which we can write, considering that $Q_H > 0$ and $Q_C < 0$, as

$$\eta = 1 - \frac{|Q_C|}{|Q_H|} = 1 + \frac{Q_C}{Q_H}. \tag{3.2}$$

We now compute the heats exchanged in the isothermal processes. We shall consider, for simplicity, one mole of gas. The heat Q_H absorbed on the isothermal process from the state *A* to *B* is, for the first law,

$$U_B - U_A = Q_H - W_{AB} \tag{3.3}$$

where W_{AB} is the work done in the process and U_A and U_B are the internal energies in the two states. Considering that the internal energy of an ideal gas depends only on the temperature, $U_B = U_A$. Hence, the absorbed heat is equal to the work done, which is given by Eq. (2.14). We obtain

$$Q_H = W_{AB} = RT_H \ln \frac{V_B}{V_A}. \tag{3.4}$$

Similarly, for Q_C, we have

$$Q_C = W_{CD} = RT_C \ln \frac{V_D}{V_C}.$$

Let us consider the signs. W_{AB} is positive, and consequently Q_H is absorbed heat; W_{CD} is negative, and consequently Q_C is heat delivered by the cycle. In absolute value, we have

$$|Q_C| = -Q_C = RT_C \ln \frac{V_C}{V_D}. \tag{3.5}$$

We now take into account the fact that *A* and *D* are on the same adiabatic curve, and so are *B* and *C*. We use the adiabatic equation in the form linking temperatures and volumes. We have $T_H V_A^{\gamma-1} = T_C V_D^{\gamma-1}$ and $T_H V_B^{\gamma-1} = T_C V_C^{\gamma-1}$. Dividing these two equations, we obtain $V_A^{\gamma-1}/V_B^{\gamma-1} = V_D^{\gamma-1}/V_C^{\gamma-1}$. Finally, raising both sides to the $-\gamma + 1$ power, we obtain $V_A/V_B = V_D/V_C$. From Eqs. (3.4) and (3.5), we get

$$\frac{|Q_H|}{|Q_C|} = \frac{T_H}{T_C}. \tag{3.6}$$

This is a very important result. In other words, the ratio of the absolute values of the heat exchanged in an ideal gas Carnot cycle is equal to the ratio of the absolute temperatures at which the exchanges take place. The efficiency is

$$\eta = 1 - \frac{T_C}{T_H}. \tag{3.7}$$

In conclusion, the efficiency of the ideal gas Carnot cycle depends only on the temperatures of the two reservoirs between which the cycle takes place. It does not depend, in particular, on the size of the cycle, i.e., on the work done.

3.4 The Carnot Theorem

We shall now extend the results of the previous section to the Carnot engines working with an arbitrary fluid. This is the *Carnot theorem* , which is composed of the following two statements:

1. The efficiencies of all Carnot engines exchanging heat between the same two reservoirs are equal, namely they do not depend on the fluid

2. The efficiency of any engine exchanging heat with only two reservoirs cannot be larger than the efficiency of the Carnot cycle between those reservoirs.

Notice that “any engine” refers to both reversible and irreversible engines, while the Carnot engine is only reversible.

We demonstrate 2 first, through a reduction ad absurdum, and then 1. Let us call *E* the generic engine and *C* the Carnot engine. To simplify the arguments, assume that the heats absorbed at the higher temperature by the two engines are equal. With the obvious meaning of the symbols, the efficiencies η of C and η' of E are

$$\eta = \frac{W}{Q_H} = \frac{|Q_H| - |Q_C|}{Q_H}; \quad \eta' = \frac{W'}{Q_H} = \frac{|Q_H| - |Q'_C|}{Q_H}. \tag{3.8}$$

Let us assume $\eta' > \eta$ (to reduce it ad absurdum). It follows that

$$W' > W; \quad |Q'_C| < |Q_C|. \tag{3.9}$$

The Carnot engine being reversible, we are free to have it working backwards, namely as a refrigerator. The heats and the work simply change their signs. The Carnot engine now needs to absorb the work W. For that, we use part of the work $W' > W$ produced by E. The complex of engine E and backwards C is still a cyclic engine. It produces the external work $W' - W > 0$, taking heat from only one reservoir, the one at temperature T_C. Indeed, at every cycle, the hot reservoir receives the same energy from the two engines that it releases. Consequently, everything goes on as if it did not exist. Indeed, we might also think to pass the heat from one machine to the other directly (Fig. 3.6). This conclusion does not agree with the Kelvin statement. This demonstrates point 2. Let us now demonstrate the same for point 1.

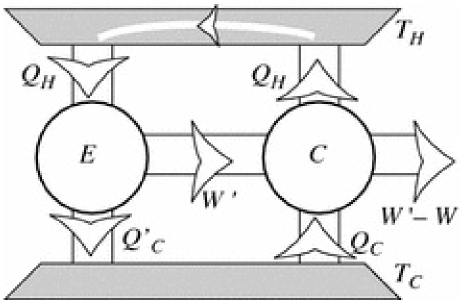


Fig. 3.6 Two engines working between the same two reservoirs, a generic one E and a Carnot one C (working as a refrigerator)

If the cycle E is reversible as well, the above argument can be inverted, showing that the assumption $\eta' < \eta$ is also false. It must then be $\eta' = \eta$.

In conclusion, we have shown that the efficiencies of all the Carnot cycles between the same two temperatures are equal. The efficiency Eq. (3.7) that we found for the ideal gas is valid for all the Carnot cycles, with whatever substance. We have also shown that the efficiencies of all the irreversible cycles between two reservoirs are less than or equal to the efficiency of the Carnot cycle between the same temperatures.

Therefore, the Carnot theorem establishes the upper limit for the efficiency of the thermal engines operating between two given temperatures. As one sees in Eq. (3.7), the closer the limit efficiency is to one, or 100 %, the smaller the ratio between the lower and higher temperatures. Very often, the lower temperature is the air temperature or, in the case of ships, the seawater. Consequently, T_C is not usually under our control. We might play with T_H. To have an idea of the orders of magnitude, take, in round figures, $T_C = 300$ K (27 °C). The maximum theoretical efficiency is, for example, 50 % for $T_H = 660$ K (327 °C) and 70 % for $T_H = 960$ K (627 °C). For this reason, thermal power stations operate close to the maximum practical steam temperature, which is about 600 °C. Their efficiency is 35–40 %, namely 50–60 % of the theoretical maximum.

In practice, the efficiency of any real engine is always less than, never equal to, the efficiency of the Carnot cycle between the same temperatures. The reasons for that are threefold. To fix the ideas, think of a piston moving up and down in a cylinder containing a fluid (this happens, for example, in our car engines).

1. Part of the work is lost against frictions that are always present.
2. In the heat exchange at the higher temperature, the mean temperature of the fluid is less than the temperature of the reservoir, $T_{H,fl} < T_H$. We have used the term “mean temperature” because the states are not of equilibrium. The opposite is true for the exchange at the lower temperature, where $T_{C,fl} > T_C$.
3. The external pressure is less than the internal one during the expansions, and consequently the (positive) work done is smaller than that in the reversible case. During the compressions, the external pressure is smaller than the internal. The work done is larger in absolute value than that in the reversible case, and is negative. Both effects reduce the total work.

In general, a thermal engine can exchange heat with several sources at different temperatures. The efficiency of a cycle is still defined as the ratio between the work done W and the absorbed heat. The total work W is the algebraic sum of the works relative to all the processes composing the cycle taken as positive if done, negative if received. Let Q_{in} be the sum of all the absorbed heats and Q_{out} the sum of all the delivered heats. As usual, Q_{in} is positive and Q_{out} is negative. The efficiency is defined as

$$\eta = \frac{W}{Q_{\text{in}}} = \frac{Q_{\text{in}} + Q_{\text{out}}}{Q_{\text{in}}}. \tag{3.10}$$

Let T_H be the highest temperature and T_C the lowest. We shall show in Sect. 3.8 that the efficiency of an engine exchanging heat with more than two reservoirs cannot be larger than the efficiency of a Carnot cycle working between the extreme temperatures. We give an example below.

Example E 3.1

Figure 3.7 shows the Stirling cycle , which is a reversible cycle made of two isothermal transformations, at T_H and T_C, respectively, and two isochoric transformations, made by an ideal gas (*n* moles). It is named after Robert Stirling (UK, 1790–1878). The cycle absorbs heat in the isochoric process *DA* (at temperatures gradually growing from T_C to T_H) and in the isothermal expansion *AB* (at T_H). It releases heat in the isochoric transformation *BC* and in the isothermal compression *CD*.

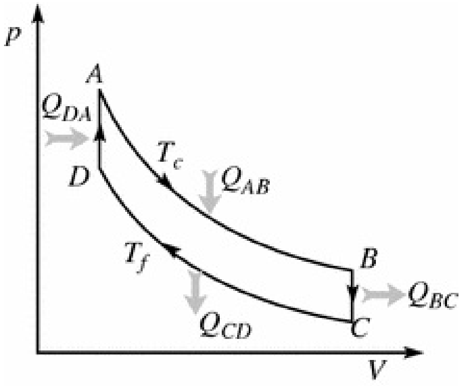


Fig. 3.7 The Stirling cycle

The heats exchanged in the isothermal processes are $Q_{AB} = RT_H \ln(V_B/V_A)$, which is positive, and $Q_{CD} = RT_C \ln(V_D/V_C)$, which is negative. The heats exchanged in the isochoric processes, in which the works are zero, are equal to the variations of internal energy. Clearly, one is the opposite of the other. Assuming the molar heat to be constant, they are $Q_{DA} = nC_V(T_H - T_C) = -Q_{BC}$. The efficiency of the Stirling engine is then

$$\eta = \frac{Q_{AB} + Q_{DA} + Q_{CD} + Q_{BC}}{Q_{AB} + Q_{DA}} = \frac{Q_{AB} + Q_{CD}}{Q_{AB} + Q_{DA}} = \frac{W}{Q_{AB} + Q_{DA}}.$$

This is smaller than the efficiency of the Carnot cycle between the same temperatures, which is

$$\eta = \frac{W}{Q_{AB}}$$

because $Q_{DA} > 0$. ▮

A thermal cycle can be used as a refrigerator , to move heat from lower to higher temperatures, rather than as a motor. From the theoretical point of view, this is just a motor working backwards. The cycle in this case absorbs work. The efficiency of a refrigerator is defined as the ratio between the heat absorbed from the low temperature reservoir and the work spent to do that,

$$\xi = \frac{|Q_C|}{|W|} = \frac{|Q_C|}{|Q_H| - |Q_C|}. \tag{3.11}$$

which, we note, is always larger than one. If the cycle is a Carnot cycle, we can also write

$$\xi = \frac{T_C}{T_H - T_C}. \tag{3.12}$$

3.5 Thermodynamic Temperature

Let us start from the conclusion of the preceding section that the efficiency of any thermal engine working between two temperatures cannot be larger than the efficiency of the Carnot engine between those temperatures. We can write

$$\frac{Q_H + Q_C}{Q_H} \le \frac{T_H - T_C}{T_H}$$

or

$$1 + \frac{Q_C}{Q_H} \le 1 - \frac{T_C}{T_H}$$

and also

$$\frac{Q_C}{T_C} + \frac{Q_H}{T_H} \le 0. \tag{3.13}$$

In other words, the sum of the ratios of the exchanged heats over the temperatures of the exchange is less than or equal to zero. It is equal to zero if the cycle is reversible, namely for the Carnot cycle .

Equation (3.13) is an important relation that we shall use in the following. Here, we use it for a new definition of the absolute temperature. Considering a reversible cycle, Eq (3.13) gives

$$\frac{T_C}{T_H} = -\frac{Q_C}{Q_H} \tag{3.14}$$

where the meaning of the minus sign is that the two heats have opposite signs (one is absorbed, the other is released).

We want to measure the temperature T of a body. We have a reservoir at the reference temperature of the water triple point $T_{tr} = 273.16$ K. We build a Carnot engine, as in Fig. 3.8, working between the body at the temperature to be measured, used as the cold source, and the hot source at T_{tr}. We measure the heats absorbed by the cycle, say $Q = Q_C$ and $Q_{tr} = Q_H$.

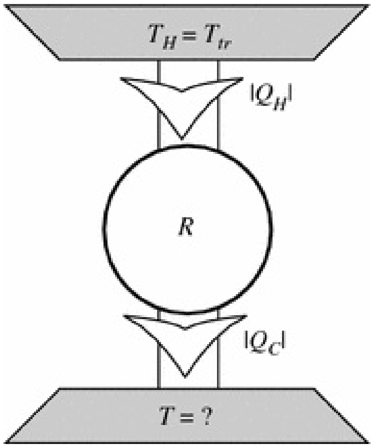


Fig. 3.8 Scheme to define the thermodynamic temperature

We now define the temperature of the body as

$$T = -T_{tr}\frac{Q}{Q_{tr}} \tag{3.15}$$

where, as you may notice, the two heats have opposite signs. The so-defined temperature is called the *absolute thermodynamic temperature* . It coincides with the ideal gas temperature that we defined in Sect. 2.2 in the interval where both definitions are meaningful, but it is also valid at the lowest temperatures, where the gas temperature does not work (below about 1 K).

The thermodynamic definition does not require measuring temperatures. We can measure only heats. For example, we can use a melting ice calorimeter to measure Q. In such a calorimeter, the heat is measured by measuring the mass of ice that melts. Similarly, we can measure how much ice melts at T_{tr}. In both cases, we weigh using balances, not thermometers. However, building a reversible engine is impossible. We must try to reduce the frictions as much as possible and apply to the measurements the corrections needed to eliminate the effects of the remaining ones.

3.6 The Clausius Theorem

We shall now demonstrate the Clausius theorem , which has a central importance in thermodynamics. We start from Eq. (3.13), valid for a cycle working between two sources, and we generalize the result to any number of reservoirs.

Consider a thermal engine, which we call M. To be concrete, we consider a motor, meaning that its W is positive. The cycle is completely general: it may be reversible or not, and it may exchange heat with any number of sources. If this number is infinite, we can approximate the system with a finite number of sources N, provided N is large enough.

Let T_i be the temperature of the generic reservoir and Q_i the heat absorbed from it by the cycle. With the usual sign convention, Q_i is positive if absorbed by the engine, negative if released. Notice that T_i is the temperature of the **source**. This is equal to the temperature of the fluid only if the process is reversible. Otherwise, the temperature of the fluid might not even be defined, because the fluid is not in an equilibrium state.

We are going to show that the sum of the ratios between absorbed heats and temperatures of the exchanges is equal to or smaller than zero. We imagine the "super-engine", shown in Fig. 3.9. Beyond the engine M, we have an additional reservoir at the temperature T_0 and Carnot cycles N. The ith Carnot cycle C_i works between the source at T_0 and the source at T_i, to which it delivers the same heat, Q_i (meaning that it absorbs—Q_i) that is absorbed by M. Let Q_{i0} be the heat absorbed by C_i at T_0. Consider now the super-engine made of M and all the C_is. At every cycle, all the engines are back to their initial state, and all the sources at T_i have received and delivered the same heat, so that they are in their initial state as well. The only component of the system that has changed is the source at T_0.

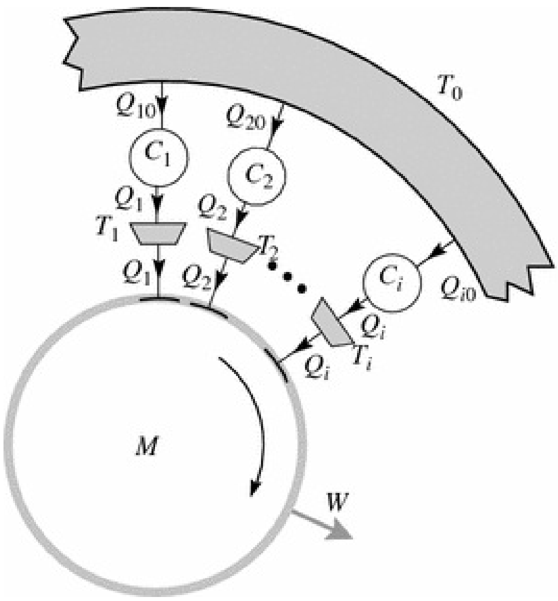


Fig. 3.9 The “super-engine” for the Clausius theorem

The work done by the super-engine is the sum of the works of M (positive) and of all the C_i (some positive, some negative). Considering that the super-engine performs cycles too, the work it does is equal to the sum of the absorbed heats. As we have just seen, all the exchanges take place with a unique source, the reservoir at T_0. It cannot be positive, for the Kelvin statement.

$$Q_0 = \sum_{i=1}^{N} Q_{i0} \leq 0. \tag{3.16}$$

On the other hand, each C_i is, by construction, a Carnot cycle; hence, it is reversible. As we showed, it absorbs heat $-Q_i$ from the source at T_i and Q_{i0} from the source at T_0. Equation (3.13) holds with the equal sign, namely

$$\frac{Q_{i0}}{T_0} - \frac{Q_i}{T_i} = 0$$

or

$$Q_{i0} = T_0 \frac{Q_i}{T_i}.$$

Substituting in Eq. (3.15), we get

$$T_0 \sum_{i=1}^{N} \frac{Q_i}{T_i} \leq 0.$$

Finally, taking into account that $T_0 > 0$, we have the desired result

$$\sum_{i=1}^{N} \frac{Q_i}{T_i} \leq 0. \tag{3.17}$$

We can tell more if the engine *M* is reversible. In this case, the super-engine is reversible too, and we can have it working backwards. The work done is exactly the opposite, $-W$ (the engine absorbs work), and all the absorbed heats are opposite too, namely $-Q_i$. All the C_is are now working backward and absorb the heats $-Q_{i0}$ from the source at T_0. We conclude that

$$\sum_{i=1}^{N} \frac{-Q_i}{T_i} \leq 0.$$

Both this relation and Eq. (3.17) hold for a reversible engine. We conclude that, for such an engine,

$$\sum_{i=1}^{N} \frac{Q_i}{T_i} = 0. \tag{3.18}$$

If the engine absorbs heat from an infinite number of sources, each exchange is infinitesimal, and Eqs. (3.17) and (3.18) become, for an arbitrary cycle,

$$\oint \frac{\delta Q}{T} \leq 0 \tag{3.19}$$

and, for a reversible cycle,

$$\oint_R \frac{\delta Q}{T} = 0. \tag{3.20}$$

In other words, the Clausius theorem states that the sum of the ratios of the heats absorbed by a cyclic engine and the absolute temperatures of the sources from which the heats are absorbed is less than or equal to zero. The equality holds if the cycle is reversible; the *R* footer in Eq. (3.20) stands to recall that it applies to a reversible cycle only.

3.7 Entropy

As we have remarked several times, the infinitesimal heat exchange δQ is not an exact differential, independent of the process being reversible or not. Equation (3.20) implies, however, that, for a reversible process, δQ divided by the temperature T, say $(\delta Q/T)_R$, is an exact differential . Notice that we have not currently specified that T is the temperature of the source, because, the process being reversible, the temperatures of the engine and the source are equal. We can express the conclusion in other words; given two equilibrium states *A* and *B* of any thermodynamic system, the integrals of $(\delta Q/T)_R$ in all the reversible processes between them are equal. We can also say that the integral of $(\delta Q/T)_R$ in a reversible process depends on both the origin and the end of the process, but not on the process itself that occurs between them. The demonstration proceeds exactly as in other similar cases. We shall give it here anyway, to be complete.

Consider, as shown in Fig. 3.10, two reversible processes, 1 and 2, between the equilibrium states A and B. The sum of process 1 and the inverse of process 2 are a reversible cycle. Hence,

$$\int_{1A}^{B} \frac{\delta Q}{T} + \int_{2B}^{A} \frac{\delta Q}{T} = 0$$

or

$$\int_{1A}^{B} \frac{\delta Q}{T} = -\int_{2B}^{A} \frac{\delta Q}{T}.$$

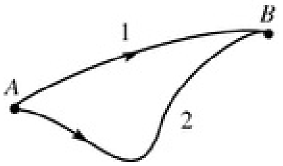


Fig. 3.10 Two reversible processes from *A* to *B*

Process 2 being reversible, we can invert the limit and we get

$$\int_{2B}^{A} \frac{\delta Q}{T} = -\int_{2A}^{B} \frac{\delta Q}{T}.$$

Finally, we have

$$\int_{1A}^{B} \frac{\delta Q}{T} = \int_{2A}^{B} \frac{\delta Q}{T}. \tag{3.21}$$

The integral from A to B in a reversible process can then be written as the difference of the values in B and A of a state function , which we call *entropy* and indicate with S, namely

$$S(B) - S(A) = \int_{A_R}^{B} \frac{\delta Q}{T}. \tag{3.22}$$

This is the definition of entropy, or, better still, of entropy differences. As the internal energy, the entropy is defined modulo and additive constant. The units of entropy are heat divided by temperature, namely J/K.

For the vast majority of thermodynamic systems, entropy is an additive quantity. Consider, for example, a thermodynamic system s composed of two subsystems s_1 and s_2. Suppose the internal energy of s to be the sum of the internal energies of subsystems s_1 and s_2. This is usually, but not always, the case. It is not so in the presence of energy associated with surfaces. Consider, for example, two drops of water. The internal energy of a drop is the sum of two terms, one proportional to the volume and one proportional to the surface. The latter is due to the surface tension, as we shall see in Sect. 4.9. We anticipate that the surface tension corresponds to the work that is needed to move a molecule from the bulk of the drop to its surface. If we put the two drops together, a larger drop forms; its volume is the sum of the two volumes, but its surface is smaller than the sum of the two surfaces. The internal energy of the composite system is not the sum of the internal energies of its components. In the majority of systems, however, the dimensions are large and the surface energy can be neglected in comparison with that of the volume. For them, the internal energy of the system is the sum of the internal energies of its parts.

We now suppose, as is often the case, that the work done on the system is equal to the sum of the works done on its parts. Then, for the first law, the heat absorbed by the system is also equal to the sum of the heats absorbed by its parts. In addition, in the reversible processes we are considering, the temperatures of all the parts are equal. The entropy is additive for the systems having these characteristics.

Equation (3.22) should be used whenever we want to calculate the entropy difference between two given equilibrium states A and B of a thermodynamic system. In general, one deals with a system performing a given process from A to B and is needed to find the corresponding entropy variation. The reader should be very careful not to be deceived by the words "*corresponding variation*" or similar terms. Indeed, the entropy variation corresponds to the fact that the system was initially in equilibrium state A and, after the process, is in equilibrium state B. It does NOT "correspond", in general, to the process actually performed. If, for example, the system performed a spontaneous process, which is irreversible, the integral of $\delta Q/T$, called the Clausius integral , in that process has nothing to do with the “corresponding” entropy variation. We shall see that immediately through a few examples.

The correct procedure for calculating entropy differences is as follows:

1. Fix your attention on the initial and final states only, forgetting the process performed by the system.

2. Consider all the *reversible* processes between the two states and choose the one that makes calculation easiest.

3. Do the calculation.

We shall now give a few examples.

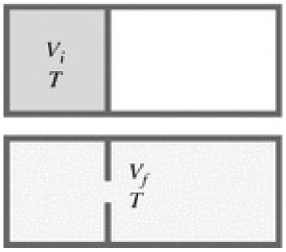


Fig. 3.11 Free expansion of a gas. Initial and final states

Example E 3.2

Free expansion of a gas. A container of volume V_t with rigid walls is divided into two parts by a septum. Initially, one of the parts, of volume V_i, is filled with n moles of a gas, which can be considered to be ideal, at the temperature T; the other part is empty (Fig. 3.11). We open an orifice in the septum and the gas expands to occupy the entire container. Calculate the corresponding variation of entropy.

If one falls into the trap and calculates the Clausius integral of $\delta Q/T$ in the actual process, considering that there is no heat exchange, he/she finds

$$\int_{A;\ \text{free expansion}}^{B} \frac{\delta Q}{T} = 0 \tag{3.23}$$

and claims that entropy does not vary. This is wrong. The Clausius integral is <u>not</u> the entropy variation because it is in an irreversible process. To calculate the entropy variation, we take the above-specified steps.

1. The initial state is $A = (p_i, V_i, T)$. We know the volume and temperature, and we can calculate the pressure with the gas equation, if we need it. The final state is $B = (p_f, V_t, T)$; we know this volume and temperature too, and we can again calculate the pressure if needed.

2. Considering that the two states have the same temperature, it is clearly convenient to take, for the calculation, the reversible isotherm process between them.

3. As the internal energy does not vary in an isothermal process of an ideal gas, the absorbed heat is equal to the work done $\delta Q = \delta W = p dV = nRT dV/V$. Hence, $\delta Q/T = nR dV/V$. Finally, integrating, we have

$$S(B) - S(A) = \int_{A;\,\text{isothermal}}^{B} \frac{\delta Q}{T} = nR \int_{A;\,\text{isothermal}}^{B} \frac{dV}{V} = nR \ln \frac{V_f}{V_i}. \tag{3.24}$$

We see, in particular, that entropy increases during the process. This is the case for all spontaneous processes of isolated systems, as we shall see (Fig. 3.12). ■

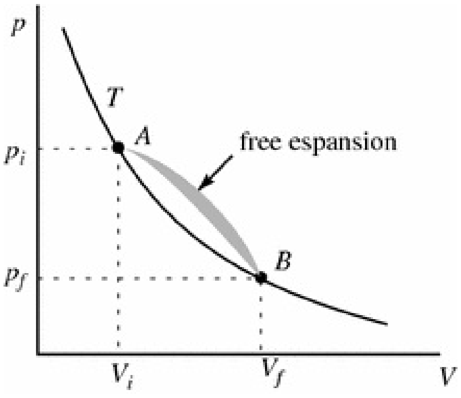


Fig. 3.12 A free expansion and isothermal processes between two equilibrium states of a gas

Example E 3.3

Spontaneous heat transfers from a hotter to a colder body. Consider two solid bodies, of masses m_1 and m_2, specific heats c_1 and c_2, and temperatures T_1 and T_2, with $T_1 < T_2$. We lodge the two bodies in a thermally insulated container and put them in thermal contact. Heat passes from the hotter to the colder body until the two temperatures are equal. The equilibrium temperature T_f can be easily calculated.

We calculate the entropy variation following the rules. We first notice that the system consists of two parts. Its entropy is the sum of the entropies of its parts. The state of each part is identified by the only thermodynamic variable of a solid, namely temperature. Indeed, we can consider the volume of a solid as invariable. Body 1 passes from T_1 and T_f. We need a reversible transformation between the two states. We can think to put the body into contact with a series of sources having temperatures between T_1 and T_f. When the body (that we are heating) has reached the generic temperature T, we put it into contact with the source of temperature $T + dT$. The body then reversibly absorbs the heat $\delta Q = m_1 c_1 \, dT$ from the source. The entropy variation of the body is then

$$\Delta S\ (1) = \int_{T_1}^{T_f} \frac{dT}{T} = m_1 c_1 \ln \frac{T_f}{T_1}$$

and, analogously, the entropy variation of the second is

$$\Delta S\ (2) = m_2 c_2 \ln \frac{T_f}{T_2}.$$

The total entropy variation is, in conclusion,

$$\Delta S = m_1 c_1 \ln \frac{T_f}{T_1} + m_2 c_2 \ln \frac{T_f}{T_2}. \tag{3.25}$$

We can make a couple of observations. First, notice that if we had calculated the entropy variation as the integral of $\delta Q/T$ in the actual process, we would have followed a faulty procedure, because that process is not reversible. However, the result would have been the correct one. This is a more of a unique than a rare case. The second observation is that the process is spontaneous and the variation of entropy is positive, as in E. 3.2. ▮

Example E 3.4

Consider a pendulum enclosed in an adiabatic container with air at atmospheric pressure. In the initial state, the pendulum is moved out of its mechanical equilibrium position and is let go (Fig. 3.13a). The temperatures of the pendulum and the air are both T_i. The pendulum will oscillate for a while, with oscillations of decreasing amplitude, due to the resistance of the air. Consider as the final state Fig. 3.13b, the state in which the pendulum is at rest and its temperature and that of the air are equal. The final temperature, T_f, is obviously larger than T_i.

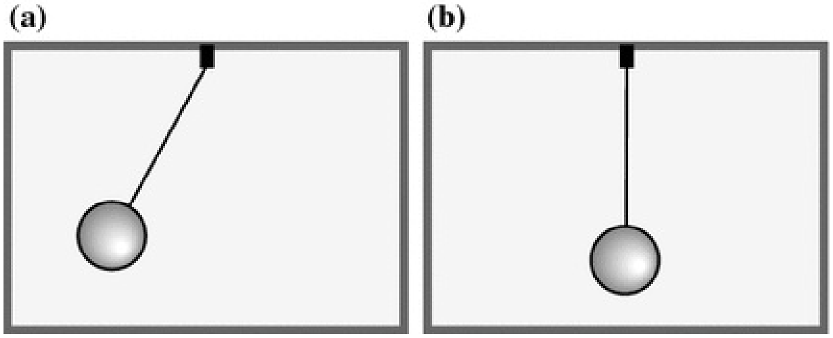


Fig. 3.13 A pendulum in an adiabatic container. **a** Initial state; **b** final state

The thermodynamic system is composed of two subsystems: a gas, which we shall consider as ideal, and a solid. The entropy variation of the system is the sum of the entropy variations of its parts. Let m be the mass and c the specific heat of the pendulum. Its initial and final states are identified by their temperatures. Indeed, the initial motion is thermodynamically irrelevant. Let n be the number of moles of the air and C_V its molar heat. The initial and final states of the air are identified by their pressures, volumes and temperatures. As the initial and final states have the same volume, we choose a reversible isochoric process, in which $\delta Q = dU = nC_V dT$. Hence,

$$\Delta S\ (\text{air}) = nC_V \int_{T_1}^{T_f} \frac{dT}{T} = nC_V \ln \frac{T_f}{T_1}$$

where we have considered C_V to be constant.

The entropy variation of the pendulum is what we just calculated for a solid body:

$$\Delta S\ (\text{pendulum}) = mc \ln \frac{T_f}{T_i}.$$

In conclusion, the entropy variation of the system is

$$\Delta S = (nC_V + mc) \ln \frac{T_f}{T_i}. \tag{3.26}$$

Once more, we see that, in a spontaneous process, the entropy variation is positive, being $T_f > T_i$. ∎

3.8 Engines Exchanging Heat with More Than Two Sources

We already mentioned in Sect. 3.4 that the efficiency η of an engine, reversible or not, which exchanges heat with more than two sources having temperatures between a minimum T_C and a maximum T_H, is smaller than the efficiency η_C of a Carnot engine working between T_C and T_H. We now demonstrate that statement, using the Carnot theorem .

Let us first consider a reversible cycle. In that case, Eq. (3.20) holds. Now, let us explicitly consider the contribution of the absorbed heats, namely the positive ones, which we shall call δQ_{in}, and the released ones (negative), δQ_{out}, and write

$$\oint \frac{\delta Q_{in}}{T} - \oint \frac{|\delta Q_{out}|}{T} = 0. \tag{3.27}$$

We wrote the left-hand side as a difference between two positive terms. Considering that T_H is the highest temperature, we can state that

$$\oint \frac{\delta Q_{in}}{T} > \oint \frac{\delta Q_{in}}{T_H} = \frac{Q_{in}}{T_H}$$

and, similarly,

$$\oint \frac{|\delta Q_{out}|}{T} < \oint \frac{|\delta Q_{out}|}{T_C} = \frac{|Q_{out}|}{T_C}.$$

Then, for Eq. (3.27) it must be

$$\frac{Q_{in}}{T_H} - \frac{|Q_{out}|}{T_C} < 0$$

or

$$\frac{T_C}{T_H} < \frac{|Q_{out}|}{Q_{in}}$$

and finally

$$\eta = 1 - \frac{|Q_{out}|}{Q_{in}} = 1 - \frac{T_C}{T_H} < \eta_C. \tag{3.28}$$

that proves the statement.

If the cycle is irreversible, Eq. (3.19) holds in place of Eq. (3.20) and the same arguments are valid *a fortiori*.

3.9 Entropy of Remarkable Systems

In this section, we shall express the entropy variations of some remarkable systems. The infinitesimal heat absorbed in a generic reversible process is

$$\delta Q = dU + \delta W = dU + pdV$$

and the infinitesimal entropy variation is

$$dS = \frac{\delta Q}{T} = \frac{dU + pdV}{T}. \tag{3.29}$$

The entropy difference between two equilibrium states *A* and *B* is

$$S(B) - S(A) = \int_A^B \frac{dU + pdV}{T}. \tag{3.30}$$

Solid body . We have already found this expression in the previous section. We repeat it here for completeness. If *m* is the mass, *c* is the specific heat and T_A and T_B the initial and final temperatures, the entropy variation is

$$S(B) - S(A) = mc \ln\frac{T_B}{T_A}. \tag{3.31}$$

Ideal gas . Let T_A and T_B be the initial and final temperatures, and V_A and V_B the initial and final volumes. We have $dU = nC_V dT$ and $pdV = nRT\frac{dV}{V}$. Hence, for Eq. (3.30),

$$S\,(B) - S\,(A) = nC_V \ln\frac{T_B}{T_A} + nR\ln\frac{V_B}{V_A}. \tag{3.32}$$

Reservoir . A reservoir is an ideal heat source with infinite heat capacity. Whatever heat it absorbs or delivers, its temperature does not vary, however, its internal energy does vary, and its state as well. As we know, the only thermodynamic coordinate of a reservoir is its internal energy. The initial and final states, A and B, are identified by the internal energies $U(A)$ and $U(B)$. We consider a reversible transformation in which the reservoir absorbs infinitesimal heat quantities δQ, such that their integral is equal to the total absorbed heat $Q = U(B) - U(A)$. All these heats are exchanged at the same temperature. The entropy change is then simply

$$S\,(B) - S\,(A) = \frac{Q}{T}. \tag{3.33}$$

Isentropic processes . Any *reversible* adiabatic process is a process at constant entropy, and is said to be an isentropic (meaning at equal entropy) process. Indeed, in such a reversible process, all the heat exchanges are zero and so are, consequently, all the $\delta Q/T$ and their integral between the initial and the final state. We here repeat that entropy is not constant, but rather increases, in an irreversible adiabatic process.

Consider now a hydrostatic system of a given mass. As we know, two thermodynamic coordinates are necessary and sufficient to define its equilibrium states. The choice of the couple of coordinates in a given problem should be done according to convenience. As a matter of fact, we can use not only pressure, temperature and volume, but also any state function (as we did for the reservoir). An often-useful choice is the entropy, temperature pair. The processes are represented in the TS plane, as shown in Fig. 3.14.

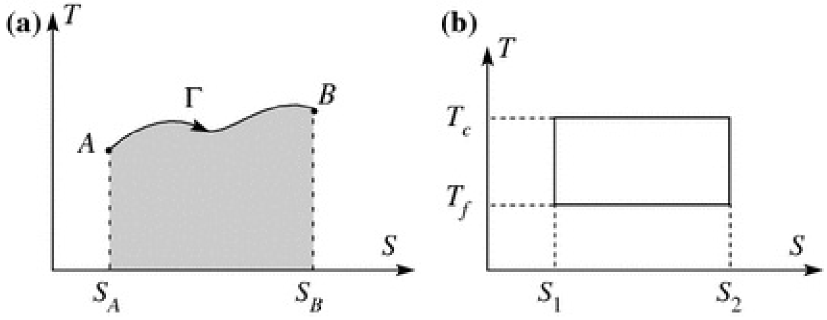


Fig. 3.14 Diagrams in the *TS* plane. **a** A reversible process, **b** a Carnot cycle

Figure 3.14a shows a reversible process from A to B. The area under the curve, grey in the figure, has the physical meaning of absorbed heat Q. Indeed, the area is the integral

$$\int_A^B T\,dS = \int_A^B T\frac{\delta Q}{T} = Q. \tag{3.34}$$

This is positive (absorbed) in the example. It would have been negative (released) if the process was in the opposite direction.

The *TS* diagram of the Carnot cycle is particularly simple. Being composed of two isothermal and two isentropic processes, it is just a rectangle, as in Fig. 3.14b.

In general, for whatever cycle, the area enclosed in the *TS* diagram represents the absorbed heat and consequently also the work done.

In Sect. 3.8, we showed that the efficiency η of any engine that exchanges heat with more than two sources between a minimum T_C and a maximum T_H is smaller than the efficiency η_C of a Carnot engine working between T_C and T_H. The demonstration is extremely simple on the *TS* plane. Consider a reversible cycle, which is represented by the closed curve in Fig. 3.15. The rectangle is the Carnot cycle between the extreme temperatures. We have profited from the fact that the efficiency of the Carnot cycle does not depend on the "length" of the isothermal transformations, but only on their temperatures, to draw the Carnot cycle tangent to the generic one. The heat input (which is positive) in the generic cycle $Q_{G,in}$ takes place on the segment *ABC;* the heat input in the Carnot cycle $Q_{C,in}$ is on the isothermal transformation at T_H. A look at the figure is enough to conclude that $Q_{G,in} < Q_{C,in}$. Similarly, one sees for the output heat, in absolute values, that $|Q_{G,out}| > |Q_{C,out}|$. These observations are enough to prove the statement.

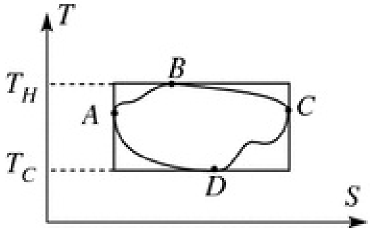


Fig. 3.15 A generic reversible cycle and a Carnot cycle

3.10 Principle of Maximum Entropy

In Sect. 3.7, we saw three spontaneous processes in which the entropy of the system increases. These are just examples of a fundamental general property of thermodynamics that we shall now prove.

Consider a thermodynamic system spontaneously transforming from the equilibrium state *A* to the equilibrium state *B*. In Fig. 3.16, the process is labeled *IR* for "irreversible". Consider also a reversible process (*R* in the figure) taking the system back to *A*. The cycle composed of these two transformations is irreversible, because it contains an irreversible process. The Clausius theorem states that

$$\int_{A;IR}^{B} \frac{\delta Q}{T} \le \int_{A;R}^{B} \frac{\delta Q}{T}.$$

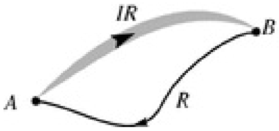


Fig. 3.16 Cycle composed of a spontaneous process from *A* to *B* and a reversible one from *B* to *A*

The right-hand side is the entropy variation, being the integral in a reversible process. We have

$$\int_{A;IR}^{B} \frac{\delta Q}{T} \le S(B) - S(A). \tag{3.35}$$

We see once more that the Clausius integral of $\delta Q/T$ is equal to the entropy variation if the process is reversible; otherwise it is smaller than it.

Consider now a generic adiabatic process from the equilibrium state *A* to the equilibrium state *B*. The Clausius integral is zero. These processes may be reversible or not and Eq. (3.35) gives

$$S(B) - S(A) \ge 0. \tag{3.34}$$

This very important relation establishes that the entropy of a thermally-isolated system can never decrease. As a consequence, any thermally-isolated system spontaneously evolves towards its state of maximum entropy. Once having reached this state, the system remains there indefinitely (in absence of external perturbations).

The conclusion is very general. Indeed, a non-isolated system exchanges heat with a certain number of sources. It is always possible to locate these sources, which are in the neighboring space of the system. We shall call the *environment* of the system the set of sources with which it can exchange heat. Then, the system composed of the system we are considering and its environment is thermally isolated, and Eq. (3.34) holds for it. The conclusion is called the *principle of maximum entropy* , which can be formulated as "*the entropy of the sum of any thermodynamic system and its environment can never decrease*."

This is clearly still another way to state the second law of thermodynamics.

The principle of maximum entropy expresses the irreversibility quantitatively. The spontaneous processes of an isolated system, or of any system plus its environment, always lead to an entropy increase. As a matter of fact, it distinguishes the arrow of time , the arrow that moves from the past to the future. If we know two states of an isolated system, or of a system plus its environment, and we do not know which came sooner and which later, we just have to look at their entropies. The state of larger entropy is the one that came later.

Before concluding, we note that one can find the principle of maximum entropy stated as being that the entropy of the Universe cannot decrease. This statement does not really have any meaning in physics, considering that nobody can measure the entropy of the Universe.

Problems

3.1. A Carnot engine that operates with a cold source at 7 °C has the efficiency $\eta = 40\ \%$. To what temperature should the hot source be raised to have the efficiency $\eta' = 45\ \%$?

3.2. Two vessels of 1 and 2 m^3 volumes contain equal masses of an ideal gas at the same temperature. What is the difference between the two internal energies? If any, which is larger? What is the difference between the two entropies? If any, which is larger?

3.3. A gas performs a reversible cycle made of two isochoric and two isothermal processes, as in Fig. 3.17a. Which are the signs of the variations (a) of the internal energy, and (b) of the entropy in each process of the cycle. (c) Under which conditions is the heat positive? (d) Under which conditions is the work positive?

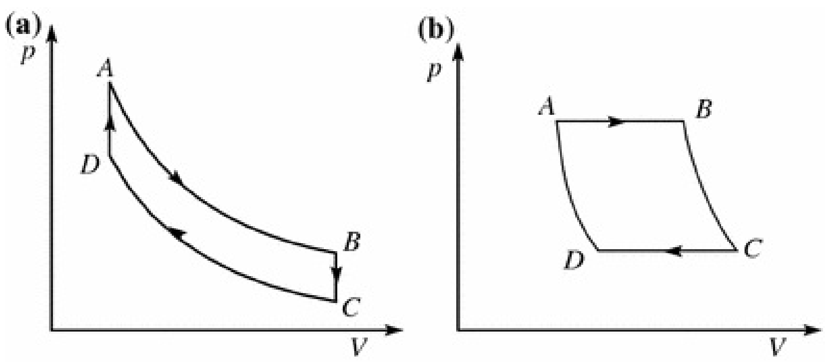


Fig. 3.17 The cycles of problems **a** 3.3 and **b** 3.4

3.4. An ideal gas performs the cycle in Fig. 3.17b made of two isobaric and two adiabatic processes. Which are the signs of the variations (a) of the internal energy, (b) of the entropy in each process of the cycle, and (c) in which the heat is positive? (d) Under which conditions is the work positive?

3.5. Consider the Carnot cycle in Fig. 3.3. The fluid is oxygen, to be considered an ideal gas. Its mass is $m = 0.4$ kg. In the warmer isothermal process at $T_H = 500$ K, the gas expands from the volume $V_A = 10^{-2}$ m^3 to $V_B = 2 \times 10^{-2}$ m^3. The work done in a cycle is 14.4 kJ. The molar heat $C_V = (5/2)\,R$ is constant. Calculate the work W_{DA}.

3.6. How does the entropy of an ideal gas behave in the following reversible expansions: (a) adiabatic, (b) isobaric, (c) isothermal processes?

3.7. Can the entropy of a system increase in a process in which the system delivers heat to the environment?

3.8. A certain quantity of gas goes from state A to B, both of equilibrium: (a) once through a reversible adiabatic process, and (b) once through an irreversible process. Which are the corresponding entropy variations? (c) Can the second process also be adiabatic?

3.9. A system passes from the equilibrium state *A* to the equilibrium state *B* through a certain process. Subsequently, it passes to the equilibrium state *C*. Knowing that the entropy variations are the opposite of each other, which is the relation between states *A* and *C*?

3.10. A certain quantity of ideal gas passes from the equilibrium state *A* to the equilibrium state *B* through a reversible isothermal process at $T = 300$ K, making the work $W_{AB} = 3$ kJ. Subsequently, it passes to the equilibrium state *C* through a reversible transformation. *A* and *C* are on the same reversible adiabatic transformation. What is the entropy variation when the system goes from *B* to *C*?

3.11. A mole of an ideal monoatomic gas is reversibly heated from 0 to 273 °C. Find the entropy variation if the process takes place: (a) at constant volume, and (b) at constant pressure.

3.12. An ideal gas expands in a reversible isothermal process at $T = 300$ K, producing the work $W = 6$ kJ. Find the entropy variation.

3.13. An ideal gas expands in a reversible isothermal process at the temperature *T* from a state of entropy S_1 to a state of entropy S_2. How much is the work?

3.14. The heat capacity of the elements near absolute zero varies with temperature as $C = \alpha T^3$, where α is a constant. Find the expression of the entropy.

Alessandro Bettini, A Course in Classical Physics 2—Fluids and Thermodynamics, Undergraduate Lecture Notes in Physics, DOI 10.1007/978-3-319-30686-5_4

4. Thermodynamic Properties of Real Fluids

Alessandro Bettini[1] ✉

(1) Dipartimento di Fisica e Astronomia, Università di Padova, Padova, Italy

✉ **Alessandro Bettini**
Email: alessandro.bettini@pd.infn.it

The ideal gas that we have studied up until now is a very important idealization with which to study the basic laws of thermodynamic. An ideal gas does not exist, but real gases behave approximately as such for low enough pressures and high enough temperatures. In the real world, gases are made of molecules that are small but have non-zero dimensions and exert forces upon one another called van der Waals forces. The ideal gas would be made of point-like non-interacting molecules. As we shall see in the next chapter, the molecules move continuously with an average kinetic energy that is proportional to the absolute temperature. Consequently, a real gas better approximates an ideal one at higher temperatures, when the potential energy of the intermolecular forces is a smaller fraction of the kinetic energy. Also, if the pressure of a given gas mass is lower, the average intermolecular distance is larger and the intermolecular forces are weaker. However, when the temperature of a gas decreases or its pressure increases, its behavior differs more and more from that of the ideal gas. At low enough temperatures, the fluid condensates, and the gas becomes a liquid. At even lower temperatures, the liquid becomes solid. A substance can be in different aggregation states (or phases). We shall discuss that in Sect. 4.1.

In Sect. 4.2, we shall see how the isothermal curves, for different temperatures, on the p,Vplane for a fluid (gas or liquid) are measured and then discuss the results. In the subsequent section, we shall see how a state equation, the van der Waals equation, is able to give an approximate description of the data.

While in an ideal gas, the internal energy is a function of the temperature alone, for the real ones, it depends on the volume as well (namely on the average distance between molecules). We shall see how this was experimentally established by Joule and Thomson in a famous experiment in Sect. 4.4. We shall then deal with the two important state functions, the internal energy and the enthalpy.

In the next three sections (Sect. 4.6–4.8), we shall deal with the aggregation phase transitions, in particular, between liquid and vapor.

In the final three sections of the chapter (Sects. 4.9–4.11), we shall study the capillary phenomena, which appear at the interfaces between different aggregation phases, a liquid and a gas, a liquid and a solid, and a gas and a solid. In particular, a liquid and its own vapor are always present in boiling and vaporization phenomena.

4.1 States of Matter

In this section, we shall give some basic hints on the structure of matter. Our main interest is in understanding the orders of magnitude of the relevant physical quantities. All macroscopic bodies are made of very small particles: molecules and atoms . Atoms characterize the elements, molecules the chemical substances. The number of atoms in a molecule ranges from a single one (monoatomic molecule, which is just an atom) up to hundreds of millions. The idea that observed differences between substances are due to the behavior of elementary objects was developed by the Greek philosophers. The origin of the atomic theory is credited by Aristotle to Leucippus , who flourished in the Vth century BC in Miletus, a Greek island near the coast of Asia Minor. Unfortunately, only a few fragments of Leucippus’s writing remain. The theory was fully developed, in philosophical terms, by his pupil Democritus of Abdera (460–370 BC). The question posed by Democritus was as follows. One can divide a piece of matter, for example, a piece of iron, into two parts, obtaining two pieces of iron. One can then break one of the pieces in half, and in half again, and in half again. Can the process continue forever? His answer was no, the process must end at some point when the smallest bit of matter is reached. This bit is indivisible, “atom” in Greek. Atoms move continually in a vacuum and are able to aggregate in different configurations corresponding to all the substances we observe. This was indeed a very brilliant idea, but still a philosophical one, which lacked experimental control. Only 23 centuries later, the existence of atoms was experimentally established. It was not a single discovery, but rather the result of a gradual process. The first important finding was by John Dalton (UK, 1766–1844) who, performing experiments with various chemical species, showed, in 1803–1808, that matter is made of “elementary” objects, the atoms, taking place in the chemical reactions. His conclusions are summarized in the *Dalton law of multiple proportions* . The law states that, *if two elements form more than one compound between them, then the ratios of the masses of the second element that combine with a fixed mass of the first element will be ratios of small whole numbers*. Soon after, in 1811, Amedeo Avogadro (Italy, 1776–1878) formulated the law we mentioned in Sect. 2.3, establishing that gas behaves as if composed of molecules. In 1897, Joseph John Thomson (UK, 1856–1940) discovered the electron.

We know, however, that the atoms, even if they are the elementary objects in the chemical reactions, have an internal structure. Atoms are composed of a central nucleus , which has a positive electric charge, and electrons , which are negative, and which form a “cloud” around the nucleus. Atoms are electrically neutral; the binding force is electromagnetic . The atomic nucleus has an internal structure as well; it is made of protons and neutrons . The force keeping the nucleus together is called the nuclear force . Protons and neutrons are composite objects too; they are made of quarks , bound by the so-called color force , for which “color” is a funny name given by physicists, one which has nothing to do with color as we know it. The nuclear force, to be precise, is a consequence of the color force. The physical laws at the atomic and sub-atomic scales are quantistic and cannot be discussed at the level of this course. However, we are interested here in giving the information that is necessary to interpret the macroscopic behavior of matter, the thermodynamic phenomena. The smallest scale objects that intervene in these phenomena (such as in the chemical reactions that we have mentioned) are molecules and atoms.

The geometrical dimensions of atoms, different from one atomic species to another, are the dimensions of this negative “cloud”. The order of magnitude is the tenth of a nanometer, or 10^{-10} m. The diameters of the nuclei are four orders of magnitude smaller, between 1 and 10 fm (10^{-15}–10^{-14} m). If we were to magnify a nucleus to the size of the dot above an “i” on this page, the size of the atom would be on the order of meters.

The number of electrons (symbol *e*), called *Z*, characterizes the element, and varies from 1 for hydrogen to 92 for uranium (Mendeleev table). Electrons inside atoms behave according to quantum, not classical, laws. In particular, electrons do not have well-defined trajectories; we cannot properly speak of electron orbits around the nucleus (even this is found in many books). Atomic electrons move very quickly compared with macroscopic objects; their speeds are on the order of 10^4 m/s, which, however, are much smaller than the speed of light. The characteristic times of electron motion are much smaller than the resolving times of our instrument and, consequently, we observe an average configuration of the atom. We see the electron charge as continuously distributed in a region around the nucleus, in a greater density where the probability of finding one electron is larger and a more sparse assemblage where the probability is smaller. We can then think of a cloud of charge, even if there is only one electron.

The atomic nucleus is made of protons (*p*) that are positive and neutrons (*n*) that are neutral. Different electric charge apart, protons and neutrons are very similar and are collectively called nucleons . For every element, the number of protons is equal to the number of electrons. Protons and electrons have equal and opposite charges; atoms, as we said, are globally neutral.

The proton and electron electric charge is the smallest existing *free* charge in nature, and is called an *elementary charge* . As a matter of fact, quarks have smaller charges. Nucleons contain two types of quark, called *up* (*u*) and *down* (*d*). Their charges are 2/3 and −1/3 of the elementary charge, respectively. The proton contains 2 *u* and 1 *d*, the neutron 1 *u* and 2 *d*. However, quarks are never free; they live inside the nucleons and other particles of the same category. The charges of all the other objects are integer multiples of the elementary charge. One might think to adopt the elementary charge as the unit, but this is not convenient because enormous numbers would represent all the usual charges. The unit of electric charge in the SI is the *coulomb* (C). We shall give the precise definition of the coulomb in the 3rd volume of the course; it is not relevant to our discussion here. The value of the elementary charge, in round numbers, is

$$q_e \cong 1.60 \times 10^{-19}\ \mathrm{C}. \tag{4.1}$$

This value is commonly used to define an energy unit at the scale of the atomic and molecular energies, the electronvolt . The electornvolt is the kinetic energy gained by an electron falling under the potential difference of one volt. As such, its numerical value in joule is equal to the elementary charge, namely

$$1\ \mathrm{eV} \cong 1.60 \times 10^{-19}\ \mathrm{J}. \tag{4.2}$$

The presence of neutrons in the nucleus is necessary to guarantee its stability. Inside the nucleus, the repulsive electric force between protons tends to destroy it. The nuclear force is, however, attractive and, under the same conditions, has the same intensity between protons, between neutrons, and between a proton and a neutron. The balance between electric and nuclear forces is realized when the number of neutrons is somewhat larger than the number of protons. The nucleon excess increases with increasing nuclear size. The number of neutrons is indicated with N, and the total number of nucleons (protons plus neutrons) with A ($A = N + Z$), which is called the *atomic number* . For a given atomic species (namely a given Z), more than one nuclear species may exist, with different values of N and, consequently, of A. All of them have the same chemical properties and are lodged in the same box of the Mendeleev table . For this reason, they are named *isotopes* (meaning “same place” in Greek). The percentages of the different stable isotopes of the same element are fixed in nature.

For example, hydrogen has two stable isotopes: ^{1}H (the superscript is A), the nucleus of which is simply the proton, and ^{2}H, the deuteron, the nucleus of which is made of a proton and a neutron. A third isotope, the tritium ^{3}H, exists but is unstable, having a half-life of 12.32 year. It is continuously produced by cosmic ray collisions in the atmosphere. The second element is helium , which has two stable isotopes, ^{3}He ($2p$,$1n$) and ^{4}He ($2p$,$2\,n$), and so on.

The masses of electron and nucleons are known through many significant figures. Here, we give values with a few digits only. The electron mass is

$$m_e \cong 9.11 \times 10^{-31}\ \text{kg}. \tag{4.3}$$

The proton mass is about 1836 times larger:

$$m_p \cong 1.673 \times 10^{-27}\ \text{kg}. \tag{4.4}$$

The neutron mass is very close to the proton mass, but a bit larger.

$$m_n \cong 1.675 \times 10^{-27}\ \text{kg}. \tag{4.5}$$

From the above values, we see that the largest fraction of the atom mass, and with them, of the mass of matter, is concentrated in the nucleus. The electrons’ contribution is only a few parts in ten thousand. One might think that the atomic masses of the elements are integer multiples of the proton mass. This is so only in a rough approximation, for three reasons. First, every element is a mixture of different isotopes with different values of A, in some proportions; second, the proton and neutron masses are almost, but not exactly, equal; third, the mass of the nucleus is not equal to the sum of the masses of its nucleons; it is smaller than that due to the binding energy, as we have seen in Chap. 6 of the 1st volume.

The mass of the u and the d quarks are about 2/1000 and 4/1000 of the nucleon mass, respectively. This is really surprising. From where does the largest fraction of the mass of the nucleons, and consequently of the nuclei, of the atoms, or of matter in general, come? The answer is in the very peculiar behavior of the quantistic, color force. On one side, it increases with the distance so much that quarks cannot be taken apart; on the other, its binding energy is *positive* and very large. Consequently, instead of a mass defect, such as in atoms and nuclei, in the nucleons, there is a mass *excess*. Namely, the mass of the nucleon is much larger than the sum of the masses of its components. This excess is the largest fraction of the mass of matter.

As far as we know, electrons and quarks do not have an internal structure and are point-like. Namely, their sizes, if any, are smaller than the experimental resolution, which is of the order of the attometer (10^{-18} m). Nucleons have a radius smaller than, but comparable to, nuclei, of a few femtometers, and, as we have already mentioned, are composed of quarks. The internal structure of the nuclei, and even more, the structure of the nucleons, does not have any influence on the thermodynamic processes we are discussing. Consequently, we shall not enter into any further detail.

In conclusion, matter is made of an enormous number of very small, charged elementary constituents, with electric charges of both signs, which are so intimately and precisely mixed that their effects, which are enormous inside the atoms, almost disappear outside of them. Some effect, however, remains outside the atom.

Two or more atoms, depending on their species, may form a molecule . The nature of the several different existing molecular bonds is, in all cases, quantistic. We can only give a few hints here. The simplest molecules are made of two atoms. The simplest types of bond are the ionic bond and the covalent bond . There are atoms, such as, for example, Na, that are happy to lose an electron and to become positive ions (Na^+), and there are those that can capture an electron, like Cl, becoming a negative ion (Cl^-). Two opposite charged ions bind together, forming a molecule (NaCl, in the example) in an ionic bond. The quantum binding interaction corresponds to the macroscopic electrostatic attractive force between opposite charges. The two atoms in molecules like H_2, O_2, HCl and many others are bound by a different interaction that has no classical analogue, called a covalent, because both partners have the same chemical valence. The two nuclei share some of the electrons and the resulting common part of the electron cloud binds the two partners. The common electron cloud is thicker in the region between the two nuclei, opposing, in this way, the repulsive force between the two nuclear positive charges.

Figure 4.1a gives an example of the interaction of potential energy in a covalent bond, the HCl molecule. This is just an example for the purpose of discussing the general features in a semi-quantitative form. Keep in mind that there is a quantitatively important difference between different molecules. The potential energy U_p is given in an electronvolt as a function of r, which is one half of the distance between the nuclei. The function has a rather deep minimum at a certain value, r_0. At smaller distances, the curve grows rapidly. The force is the derivative of the potential, which is the slope of the curve in the figure. For $r < r_0$, the force is strong and negative, meaning that the two atoms repel each other. This is due to the electric action of the two electron clouds that become compressed. The two atoms behave almost the same as two rigid spheres. The distance r_0 is the stable equilibrium distance and is, in general, on the order of tenths of a nanometer (0.12 nm for HCl in the figure). The depth of the potential minimum U_{p0} characterizes the strength of the bond between the two atoms. Indeed, U_{p0} is the minimum energy we must give to the system in order to break the molecule. We can do so, for example, by hitting it with another molecule. Typical values of U_{p0} are on the order of the electronvolt (−4.5 eV for HCl).

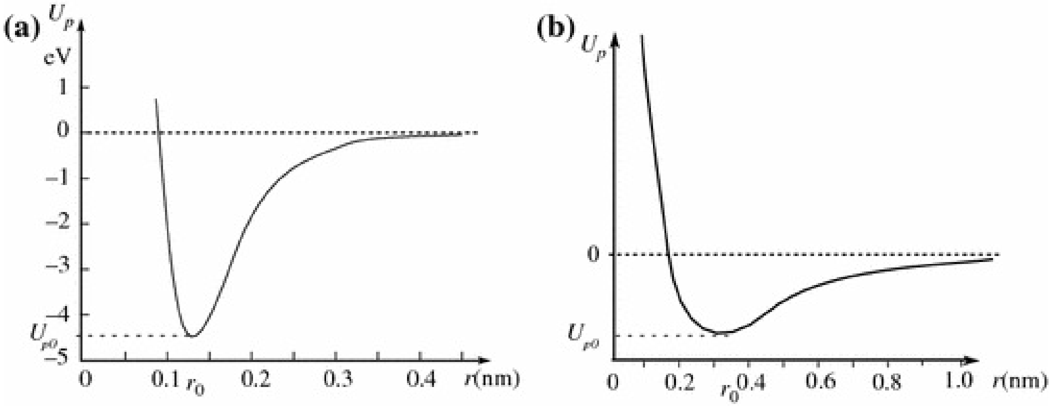


Fig. 4.1 **a** Interaction potential in the HCl molecule. **b** van der Waals potential

An important characteristic of the forces binding the molecules, which we shall call chemical forces , is the *saturation* . This means that, once the molecule has been built, the chemical force will not act on another atom that may pass nearby.

Molecules interact with each other through a force called the van der Waals force , after Johannes Diderik van der Waals (The Netherlands, 1837–1910), who gave it its expression in 1873. These forces are different from chemical forces. Rigorously speaking, they are rather complicated; for example, they depend on the relative orientation of the interacting molecules. We are interested, however, in the average value of all possible orientations. Figure 4.1b shows a typical mean van der Waals potential (the force is obviously its derivative) as a function of half of the distance between the molecular centers r. At first sight, its behavior is qualitatively similar to the potential of the chemical force, in particular, with a minimum at, using the same symbol, r_0. However, there are also very important differences. The similarity is that the van der Waals force is also repulsive and rapidly growing at short distances ($r < r_0$) and attractive and slowly decreasing at large distances ($r > r_0$). The molecules can be roughly approximated with small rigid bodies of radius r_0 as well. Notice that the "radius" r_0 of the molecules is larger than the "radius" r_0 of the atoms, even for the simplest molecules. For the diatomic ones, such as HCl, $r_0 = 0.3$–0.4 nm.

The (feeble) attractive force between molecules gradually vanishes at large enough distances. We can consider that a molecule acts only within a certain distance, which we call the *molecular action radius* and indicate with r_a. This cannot be defined very rigorously, but we can think of it as being a few times greater than r_0, typically less than 1 nm.

We did not show any energy scale in Fig. 4.1b, because the energy values vary by more than an order of magnitude for different molecules. In any case, however, the depth of the potential minimum of the intermolecular force is much smaller than that of the chemical ones. It typically ranges between 10 meV and a few 100 meV.

Another very important difference is that van der Waals forces, as opposed to chemical ones, do not saturate. Two nearby molecules interact with other molecules as well. As a result, all molecules that are closer than the action radius attract each other. This feature is at the origin of the liquid and solid states, which we shall now briefly describe.

In a gas, the typical distance between neighboring molecules is much greater than the molecular size. Their motion is completely disordered. It is called *thermal motion* . The average kinetic energy of the molecules in a gas is considerably larger than the potential energy of the van der Waals force, and, as we shall see, increases with temperature. We shall study the microscopic interpretation of thermodynamics in more detail in Chaps. 5 and 6.

Let us evaluate the order of magnitude of the distance between molecules. As the reader will remember, a mole of any gas, considering it to be ideal, occupies the same volume under the same conditions of temperature and pressure. The molar volume at STP is 22.4 L. The average distance between molecules $\langle r \rangle$ is the cubic root of the average volume available to one molecule, which is the molar volume, divided by the Avogadro number. Hence,

$$\langle r\rangle = \sqrt[3]{22.4\times 10^{-3}\ \mathrm{m}^3/N_A} = 3.3\times 10^{-9}\ \mathrm{m} = 3.3\ \mathrm{nm}.$$

The mean distance at STP is substantially larger than the molecular action radius, and the van der Waals force is weak and can be neglected in a first approximation; the gas behaves approximately as an ideal one. If the pressure increases, however, the distances between molecules become smaller and smaller, and the van der Waals force effects come into play. The state equation becomes, approximately, the van der Waals equation, which we shall study in Sect. 4.3.

If the average intermolecular distance further decreases, due to increasing pressure, or the kinetic energy decreases, due to decreasing temperature, the attractive van der Waals force effects become more and more important. Finally, the gas liquefies. This is the second aggregation phase, different from the gas, in which the molecules are, so to speak, touching one another. The average distance between molecules is substantially $2r_0$. The repulsive force becomes very strong if we try to further reduce the distance. Consequently, the volume of a liquid is almost constant. On the other hand, the molecules in a liquid are free to slip one over the other. Consequently, a liquid has no definite shape, taking the shape of the container. The inter-molecular distances are one order of magnitude smaller than for a gas at STP. Consequently, the density of the liquid is three orders of magnitude larger than that of a gas at STP. As a rule of thumb, the density of a liquid is 10 % smaller than that of the corresponding solid.

If we further decrease the temperature, the average kinetic energy of the molecules decreases too and the substance, under certain conditions, becomes solid . This is the third phase, or aggregation state, of matter. In the solid state, the constituent particles, which may be ions, atoms or molecules, are closely packed together. The van der Waals forces are so strong that the particles cannot move freely. They can, and do, vibrate around fixed positions. As a result, the solid not only has a (almost) fixed volume, like the liquid, but also a (almost) fixed shape. In the properly-named solid state, the particles arrange themselves in symmetric structures, forming a crystal . In this structure, a *unit cell* repeats itself periodically. The shape of the cell depends on the constituent ions, atoms or molecules and may be a cube, a prism, a tetrahedron, etc. The vertices of the periodic structure are the equilibrium positions around which the particles oscillate. The distances between particles are on the same order as those in the liquids. As already mentioned, the densities are similar, some 10 % larger in solids. Figure 4.2 shows two examples of crystalline structures: (a) represents 12 (not rectangular) parallelepiped cells; (b) represents three cells that are hexagonal section prisms, with an atom at the center of the face.

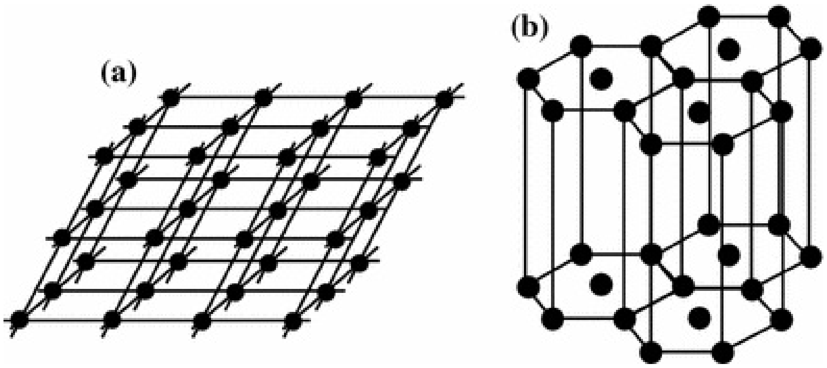


Fig. 4.2 Crystal structures with **a** parallelepiped cell; **b** face centered hexagonal prismatic cell

Crystallography is the branch of physics that studies crystals and their symmetries experimentally and theoretically. The existing shapes are many, but we shall not deal with them. We will only mention that the symmetry rarely appears at the macroscopic level. In these fascinating cases, one talks of crystals in the common language as well. Much more frequently, the crystal structure, which exists, does not appear. This is the case with metals, for example. In these cases, the macroscopic body is an aggregate of microcrystals . Their sizes are on the order of the micrometer, which is too small to be seen with the naked eye, but very large compared to the atomic sizes. The microcrystals can be easily seen with a microscope. Their arrangement appears to be completely disordered. Consequently, the body does not have a symmetric structure at the macroscopic level.

Several bodies exist, like plastic materials and glasses, which are solid in the common language, as they have a fixed shape. They are not properly solid, according to the definition given in physics. Indeed, they do not have, even at the microscopic level, any symmetrical structure. As a matter of fact, they are liquid, with an extremely high viscosity.

4.2 Isothermal Transformations of Real Fluids

As we know, isothermal transformations of the ideal gas on the plane $V\,p$ are hyperbolae with the coordinate axes as asymptotes, corresponding to the state equation $pV = \text{const}$. In the real world, this law is approximately valid for gases at high enough temperatures and small enough pressures. The knowledge of the isothermal curves of a gas is equivalent, for any practical point of view, to the knowledge of its state equation, even if the latter cannot be expressed in an analytical form. An isothermal curve of a gas can be determined experimentally by measuring the pressure as a function of the volume, keeping the temperature constant. We enclose the gas in a cylinder with a piston, so as to be able to change its volume. Clearly, we must eliminate the air from the cylinder, producing a vacuum with a pump, before introducing the gas. We put the cylinder into thermal contact with a bath at the desired temperature and proceed slowly, varying the volume and measuring volume and pressure. Then, we draw the isothermal curve through the points we have measured. We repeat the process for many temperatures.

To be concrete, let us consider carbon dioxide. The systematic measurements were done by Thomas Andrews (Ireland, 1813–1885) in the 1860s and are shown in Fig. 4.3. Consider working with one mole. At high temperatures, say above 80 °C, the curves are similar, but not identical, to those of an ideal gas. The similarity is closer the higher the temperature. At lower temperatures, the curves develop an inflection point, at which the slope is negative. These curves —such as the one of 40°—resemble hyperbolae at large volumes, i.e., small pressures, but behave very differently at small volumes. The curve has a vertical asymptote, which is not, as in the ideal gas, the ordinate axis.

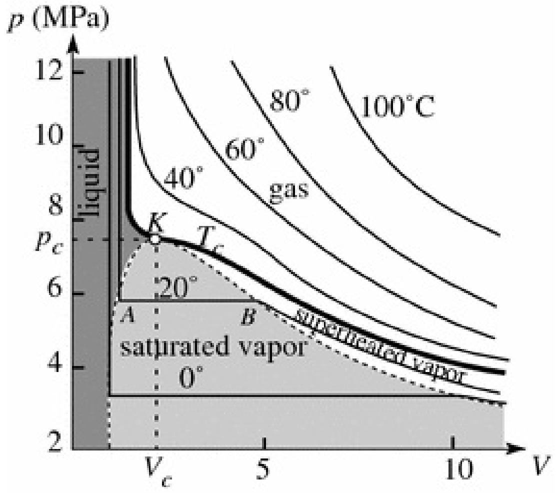


Fig. 4.3 The isothermal curves of CO_2

Going further down in temperature, we meet an important isothermal, characterized by having zero slope at its inflection point. It is drawn thicker in the figure and is called a critical isothermal . The inflection point is the *critical point* , *K* in the figure. Its coordinates are the critical temperature (T_c = 31.04 °C for CO_2), critical pressure and critical volume (molar because we are dealing with one mole). As we shall now see, gases can condensate only at temperatures lower than that which is critical for each one. Under these conditions, the gas is said to be a vapor .

Consider a temperature lower than the critical one, 20 °C, for example. Let us start from a large volume, hence from a low pressure. Decreasing the volume, when the pressure reaches a well-defined value (point *B*), which is about 6 MPa (60 atm) for CO_2, the pressure ceases to vary, even if the volume decreases. The representative point moves on a horizontal line, i.e., at constant pressure. In this segment, the isothermal process is isobaric as well. At the end of the horizontal segment (point *A*), the pressure suddenly increases and it is practically impossible to further decrease the volume. The substance that was easy to compress at larger volumes has become incompressible.

Let us look inside the cylinder. We see that beyond the point *B*, part of the gas is liquid. The gas to liquid phase transition is called condensation . In the segment *BA*, both liquid and gas are contemporarily present. The two phases coexist in equilibrium. For a given substance at a given temperature, this can happen only at a certain pressure, which is called the *saturated vapor pressure* . Let us make clear that the equilibrium between the two phases can be reached only if the following conditions are satisfied. First, the fluid must be in a closed container; otherwise, the vapor dissipates. Second, there must be no other gas (air, for example) present in the container; otherwise, the pressure of the liquid is equal to the pressure of the vapor plus the foreign gas. We will also notice that the saturated vapor pressure depends on the temperature, but not on the volume. If we repeat the experiment with a different quantity of CO_2, we find the same pressure, at a different volume.

The lengths of the parts *PB* and *PA*, in which the representative point *P* divides the line *AB*, are proportional to the fractions of liquid and vapor, respectively. In *A*, all the substance is liquid. Its volume cannot be further reduced. The slope of the isothermal curve becomes almost infinite.

At still lower temperatures (for example, at 0 °C), we observe similar behavior, but the horizontal stretch gets longer, with much more on the larger volume side than on the smaller volume side. This is because the volume of the liquid has only a small dependence on temperature. During condensation, the thermal bath absorbs heat from the substance. The heat absorbed for the complete condensation is a well-defined quantity proportional to the mass, characteristic of the substance. An equal heat is delivered to the bath, rather than absorbed, in the inverse process, which is the vaporization . The heat for the phase transition of the unit mass is called the *specific latent heat of vaporization*

. For a mole of the substance, it is called the *molar* . The heat of vaporization depends upon temperature, being smaller for higher temperatures and becoming zero at the critical temperature.

The critical temperature T_c is the maximum temperature at which the substance can be liquid. At higher temperatures, only the gas state is possible, for whatever pressure. The critical point , *K* in the figure, is particularly interesting. In its neighborhood, infinitesimal variations of pressure or temperature make the substance change from liquid to vapor or vice versa. The vapor heat is zero at the critical temperature. As a matter of fact, under these conditions, the vapor pressure is high and the density is very close to the liquid density. The difference between the two phases is vanishingly small.

In summary, inside the bell-shaped curve, grey in the figure, called the *coexistence curve* , the saturated vapor and liquid phases are in equilibrium. On the left-hand side of the coexistence curve and of the critical isothermal, the substance is in the liquid phase (darker grey in the figure). On the right-hand side of the coexistence curve, below the critical isothermal, the substance is non-critical vapor, or overheated vapor ; this is a vapor at a pressure smaller than the saturated vapor pressure (at the given temperature).

The behavior we have described for carbon dioxide is similar for the other fluids, once the values of the critical constants, which are different for different substances, are considered. As a matter of fact, the approximate validity of the *law of the corresponding states* has been experimentally established. We must use the so-called reduced variables , which are pressure, volume and temperature divided by their critical values, namely

$$\pi = \frac{p}{p_c}, \ \phi = \frac{V}{V_c}, \ \tau = \frac{T}{T_c}. \tag{4.6}$$

If we draw the reduced temperature isothermal curves (τ = cost) on the plane π, ϕ, we find that the same curves approximately represent the behavior of all the fluids.

Table 4.1 reports the critical parameters of several substances.

Table 4.1 Critical parameters of several substances

Substance	**T_c (K)**	**p_c (MPa)**	**V_c (cm^{3}mol^{-1})**	**ρ_c (kg m^{-3})**
Helium	5.3	0.23	58	69
Hydrogen	33.3	1.3	64.5	31
Nitrogen	126.2	3.39	90	311
Argon	150.9	4.91	85.0	531
Oxygen	154.8	5.08	74.4	430

Methane	191	4.64	98.8	162
Ethylene	283.1	5.19	124	225
Ethane	305.4	4.88	143	211
CO_2	304.2	8.1	95.6	460
Propane	369.9	4.2	200	220
Ammonia	405.6	11.4	72.4	235
Water	647.3	22.12	45	400

4.3 Van der Waals Equation

As we have seen in Sect. 4.2, when the density of a gas increases, its behavior differs more and more from that of an ideal gas, finally becoming liquid. These phenomena depend upon the forces between molecules. It is not possible to take into account analytically the effects of these interactions and theoretically establish a precise state equation. We can find it experimentally, as we have seen in Sect. 4.2.

It is, however, possible to take into account the main characteristics of the intermolecular forces and write a state equation describing *approximately* the behavior of a real fluid. This is the van der Waals equation . We shall find it starting from the ideal gas equation, which we write as

$$p = \frac{nRT}{V} \tag{4.7}$$

and introduce the necessary correction terms.

We have already discussed the intermolecular forces in Sect. 4.1. We recall their principal characters: the force is repulsive and rapidly increases when the distance between the centers decreases below a certain value $2r_0$, while it is weakly attractive and decreases with increasing distance above $2r_0$.

The rapid increase at small distances makes it possible to consider, in a first approximation, the molecules as rigid spheres of radius r_0. The volume of the fluid cannot be reduced below the volume taken by the molecule. This volume is called the *covolume* , which is, substantially speaking, the volume of the liquid, which we will indicate with b. We correct the ideal gas law Eq. (4.7) to take into account that the available volume is only $V–b$ and write

$$p = \frac{nRT}{V - b}.$$

We see, in particular, that the pressure diverges when V tends towards b.

We now have to take into account the attractive part of the intermolecular force. We start by observing that the pressure of the gas on the surrounding walls is due to the collisions, enormous in number, of the molecules. We shall come back to that in Sect. 5.1. The pressure is higher for larger momenta of the colliding molecules (hence, of the impulse they produce) and for larger numbers of molecules. The attractive interaction between molecules reduces the momenta and consequently the pressure.

A molecule far from the walls feels the attraction of the other molecules in a sphere of radius r_a, the molecular action radius. The distribution of the molecules is uniform, and consequently, the resultant of the forces is zero (Fig. 4.4a). When the molecule is near a wall, it feels the attraction of the molecules in half of the sphere only (Fig. 4.4b). The resultant is a force directed away from the surface. This force, which slows down the molecule, is proportional to the number of molecules in the half sphere, and hence to the density of the gas. On the other hand, the pressure is proportional to the number of molecules hitting the wall per unit time, hence also to the gas's density. In conclusion, the pressure very near the wall is lower than the pressure in the bulk by a quantity that is proportional to the square of the gas's density, or, in an equivalent manner, inversely to the square of its volume. Notice that whenever we measure the pressure, we introduce an instrument into the fluid, an external body. The pressure we measure is the one close to the surface of the instrument (even if we put it in the middle of the fluid). In conclusion, the measured pressure p is the pressure near to an external surface, not the pressure in the bulk. We take this effect into account, adding to p the term a/V^2, where a is another constant

$$p+\frac{a}{V^2}=\frac{nRT}{V-b},$$

which can also be written as

$$\left(p+\frac{a}{V^2}\right)(V-b)=nRT. \tag{4.8}$$

This is the *van der Waals equation* . The constants a and b are different for different fluids. They also depend on the mass of fluid, as we shall see at the end of the section. The constants cannot be obtained from theory, but are rather free parameters to be determined in the manner best suited to reproduce the experimental data.

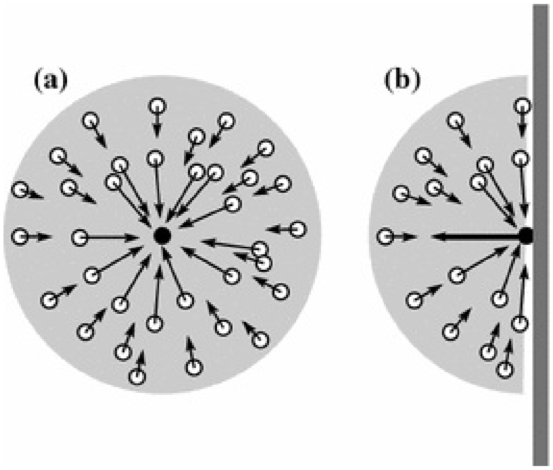


Fig. 4.4 **a** A molecule in the bulk of the fluids and its action sphere; **b** a molecule near a surface

Notice that, if the gas is very rarefied, i.e., if its volume is large, the correction terms we introduced become negligible and Eq. (4.8) goes back to the ideal gas law. But Eq. (4.8) can also describe phenomena at the opposite limit, at high density and high pressure levels. Let us study the shape of the isothermal curves in the plane *Vp*. To this aim, we write Eq. (4.8) as

$$V^3 - \left(b + \frac{nRT}{b}\right)V^2 + \frac{a}{p}V - \frac{ab}{p} = 0. \tag{4.9}$$

On the isothermal curves, obviously, *T* is a constant, and we interpret Eq. (4.9) as the relation giving *V* as a function of *p*. This is an algebraic equation of the third degree in *V*. Such equations have three roots, which can be real and different, real and coincident, or have one that is real and two that are complex. Only real roots have a physical meaning. (To be precise, to have physical meaning, they must also be positive; however, it can be shown in the present case that if a solution is real, it is also positive, provided that $p > 0$). Consequently, the isothermal curves are cut by the horizontal lines $p = \text{const}$ at three points or just one. For the higher temperatures (only one real root), the shapes of the curves are not very different from the hyperbolae of the ideal gas. At lower temperatures (three real roots), the isothermal curves become completely different. The curve corresponding to the temperature at which the three roots coincide separates the two cases. It has an inflection point with a horizontal tangent. These are the critical isothermal and the critical point, respectively.

Figure 4.5 shows the van der Waals isothermal curves. Comparison with the empiric examples of Fig. 4.3 shows that, in the region of the permanent gas, which lays above the critical isothermal, in the region of the liquid, and in the region of the super-heated gas, the van der Waals equation follows the experimental data reasonably well (provided the constants *a* and *b* are properly chosen), but not perfectly, as we shall see. In the region of the saturated vapor, under the coexistence curve, the empirical isothermal curves have a horizontal segment, while the van der Waals are *S* shaped.

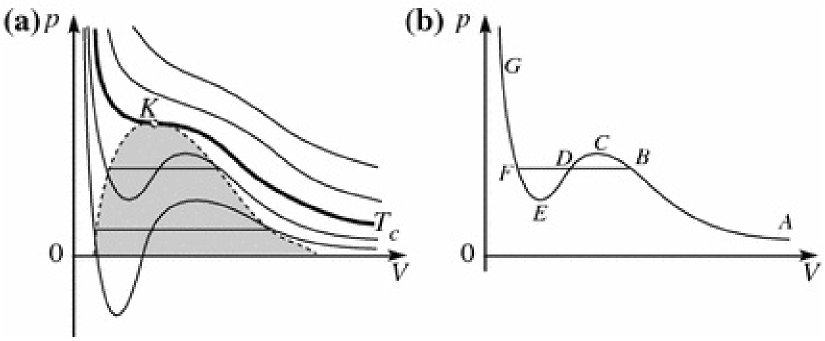


Fig. 4.5 **a** Isothermal curves of the van der Waals equation. **b** An isothermal below critical temperature

Let us follow an isothermal process in detail, using the one represented in Fig. 4.5b. We start from a low density (large volume). In segment *AB*, the substance is over-heated vapor. As we have mentioned, the behavior of the van der Waals curve approximately reproduces the real one. Below *B*, the van der Waals curve continues smoothly, while the experimental curve shows the horizontal segment *BF*, corresponding to the condensation.

As a matter of fact, the first part of the segment *BC* is also experimentally recheable. Indeed, provided that the vapor and its container are very clean, we can take the gas pressure above the saturated vapor pressure (at the given temperature) without condensation. As we shall see in Sect. 4.11, the condensation process needs to be initiated by so-called *condensation nuclei* , which are, in fact, impurities in the volume or roughness of the walls. However, the states represented by the segment *BC* are not stable (they are said to be *metastable*); any small perturbation is sufficient to take the system on the horizontal segment suddenly and irreversibly: the two phases have formed.

On the other side of the coexistence curve, the segment *GF* represents the liquid. We now start from a high pressure state on *GF* and gradually expand. If we proceed smoothly enough, we are able to experimentally reach part of the segment *FE*, provided the liquid and the walls are very clean. In fact, one can take the pressure below the saturated vapor pressure for some time without the liquid beginning to boil. As we shall see in Sect. 4.11, boiling is also triggered by the presence of nuclei (powder, irregular points on the walls, ions, etc.). The segment *FE* corresponds to the superheated liquid . The corresponding states are metastable. The smallest perturbation causes the representative point to “jump” onto the horizontal segment: suddenly, the liquid boils.

There are also isothermal curves in the part of the segment *FE* that can reach negative pressures. In these states, the liquid is *expanded* . Some of these states can be reached, for example, as follow. We put the liquid in a capillary tube, closed at one extreme and open at the other. We fix the capillary horizontally at a vertical axis passing through the closed extreme and put the system in rotation with a high angular velocity. In the reference frame of the rotating capillary, the centrifugal force (pseudo force in the non-inertial frame) pulls the free extreme of the liquid column outside. Cohesion forces forbid the column from breaking (as long as the centrifugal force is not too strong) and the pressure goes negative.

The segment *EC*, on the other hand, does not correspond to any physically achievable state, not even in principle. This segment is indeed anomalous, in contradiction to Le Chãtelier’s principle , because pressure would increase when the volume increases. Assume an equilibrium state to exist on this segment. Then, the response of the system in reaction to a change in the volume would be to enhance, rather than to decrease, the variation. Indeed, suppose that the volume of a small portion diminishes momentarily under the effect of fluctuations, which are always present (the molecules in the small volume have approached each other for a moment at a little more than the average rate). The decrease in the volume implies a reduction of the pressure, which, in turn, causes a further reduction of the volume, and so on. The zone becomes smaller and smaller indefinitely. There is no equilibrium, and the state is so completely unstable, they cannot exist.

When temperature increases, the three roots of the van der Waals equation approach each other and become coincident at the critical temperature. All the states represented by the van der Waals equation above the critical temperature are reachable, although some do not exist below it. In this way, the van der Waals equation interprets the necessity of the phase separation below the critical temperature.

In conclusion, the van der Waals equation gives an approximate description of the behavior of the real fluids. It does that, however, only in a first approximation. Let us compare the van der Waals equation predictions with reality.

The two parameters a and b in the van der Waals equation must be determined through the experimental data. The simplest way to do this is to start imposing on the van der Waals critical isothermal at T_c so as to go through the experimentally determined critical point (V_c, p_c). In this way, we determine two unknowns of a problem with three degrees of freedom. We (arbitrarily) take the gas constant R as the third unknown without giving it a priori the value for the ideal gases.

The van der Waals critical isothermal curve is the solution with three real coincident roots. We take the relations amongst them from a book of algebra and have

$$V_c = 3b,\ p_c = \frac{1}{27}\frac{a}{b^2},\ T_c = \frac{8}{27}\frac{a}{bR}. \tag{4.10}$$

We now invert these expressions to have the constants a, b, and R as functions of the critical variables and get

$$a = 3p_c V_c^2,\ b = \frac{V_c}{3},\ R = \frac{8}{3}\frac{p_c V_c}{T_c}. \tag{4.11}$$

The first two equations give the constants a and b on the basis of the experimental values of the molar critical pressure and volume. The third equation is, contrastingly, a relation amongst the critical constants of any fluid that should satisfy if the van der Waals equation is correct. Here, we find the first failure of the theory: the values of R calculated with the third Eq. (4.11) are different for different fluids and different from the gas constant (R = 8.3 J mol $^{-1}$ K^{-1}). For example, for Nitrogen and Oxygen, one gets values around 6.5 J mole^{-1} K^{-1}, for Helium and Hydrogen, values around 6.7 J mol^{-1} K^{-1}, for Ammonia, 5.4 J mol^{-1} K^{-1}, for water, 4.1 J mol^{-1} K^{-1}, and for Argon, 7.3 J mol $^{-1}$ K^{-1}.

Even having treated R as a free parameter, the result is only approximately satisfactory. Figure 4.6 shows, as an example, three isothermal curves for Argon; the true ones as continuous lines, the van der Waals ones (with the three parameters determined as above) as dotted lines. As we can see, even outside the coexistence curve, the theory fails by several percentage points.

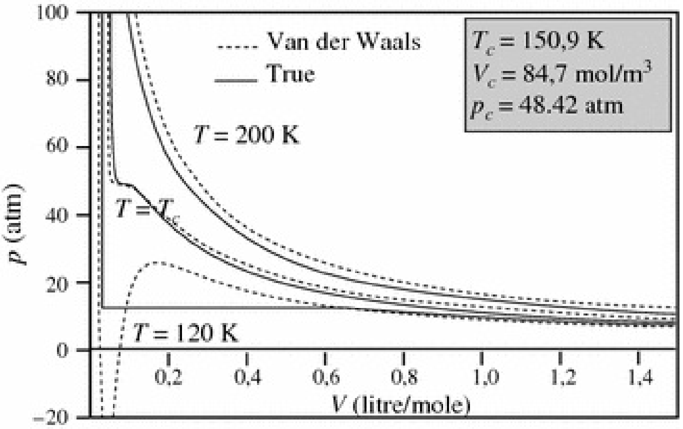


Fig. 4.6 True (*continuous*) and van der Waals (*dotted*) isothermal curves for Argon

Before concluding, we observe that the van der Waals constants a and b depend not only on the fluid but also on its mass, say on the number of moles. The values for a mole, which we indicate now with a_{mol} and b_{mol} [as they are indicated in Eq. (4.11)], depend only on the fluid. These are the values you usually find in the tables of the manuals. The constant b is the total volume of the molecules, and consequently is proportional to the number of moles. From the van der Waals equation, one also sees that the constant a is proportional to the square of the number of moles. Hence, in formulas, for n moles,

$$a = n^2 a_{\text{mol}} \qquad b = n b_{\text{mol}}. \tag{4.12}$$

The units of the van der Waals constants for an arbitrary quantity of fluid and for one mole are

$$[a] = \text{Pa m}^6,\ [a_{\text{mol}}] = \text{Pa m}^6\ \text{mol}^{-2},\ [b] = \text{m}^3,\ [b_{\text{mol}}] = \text{m}^3\ \text{mol}^{-1}. \tag{4.13}$$

4.4 Joule-Thomson Effect

In Sect. 2.12, we saw that the Joule free expansion of a gas takes place at a constant temperature. As a consequence, the internal energy of an ideal gas does not depend on the volume. As we have already noticed, however, the Joule experiment has a low sensitivity, because the heat capacity of the gas is small compared to the heat capacity of its container. The conclusion can consequently be considered as valid only in a first approximation. As already mentioned, Joule himself, along with W. Thomson, designed and performed the far more sensitive experiment that we shall now describe.

We start with a logically simplified scheme of the Joule-Thomson apparatus, as shown in Fig. 4.7. It consists of two cylindrical sections, each closed by a piston and separated by a porous wall. The function of the separator is to let the gas through with a very small velocity, when there is a pressure difference between its sides. Let us call these pressures p_1 and p_2. The walls and the pistons are thermal insulators.

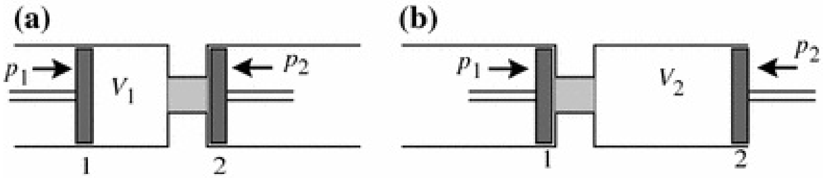


Fig. 4.7 The Joule-Thomson experiment. **a** initial state, **b** final state

Initially, all the gas is in the left container, as in Fig. 4.7a. Its volume is V_1. We move piston 1 to the right to push the gas into the second container, constantly keeping pressures p_1 and p_2 on the pistons. Finally, all the gas is in the second compartment with the volume, say, V_2 and the pressure p_2, as in Fig. 4.7b.

In both compartments, the process was at a constant pressure. Consequently, the work done by piston 1 to push the gas out has been $p_1 V_1$; while the work done by piston 2 is—$p_2 V_2$. There has been no heat exchange, because the system is isolated. Hence, the total work is equal to the variation of the internal energy $U_2 – U_1$ of the gas:

$$U_2 - U_1 = p_1V_1 - p_2V_2, \tag{4.14}$$

which we can write as

$$U_1 + p_1V_1 = U_2 + p_2V_2 \tag{4.15}$$

or

$$H_1 = H_2 \tag{4.16}$$

where $H = U + pV$ is the enthalpy . In conclusion, the enthalpy is conserved in the Joule-Thomson experiment.

In practice, for better sensitivity, the operation is a bit different. The pistons are not present. Rather, we force a continuous gas flow through the porous separator, maintaining a constant pressure difference $p_1 – p_2$ with a pump. We measure the temperatures T_1 and T_2 of the gas on the two sides of the separator. If the gas were ideal, the temperature would be equal. Indeed, the enthalpy of the ideal gas, like its internal energy, depends only on the temperature. The experiment shows that the temperatures are different. The majority of the gases, around STP, cools down in the expansion, namely $T_1 > T_2$. Air at room temperature, for example, cools 0.3 °C, expanding from $p_1 = 2$ atm to $p_2 = 1$ atm, and 45 °C, expanding from 200 to 1 atm. Hydrogen and oxygen are exceptional; at room temperature, they heat up in an adiabatic expansion.

To be precise, all the gases heat up in Joule-Thomson if the temperature is high enough and cool down if it is low enough. A, pressure dependent, *inversion temperature* exists, at which the temperature variation in the Joule-Thomson expansion changes sign. For hydrogen, the inversion temperature at atmospheric pressure is about 200 K, while for helium, it is about 40 K.

The Joule-Thomson effect shows that the enthalpy of the real gases, and consequently their internal energy, does not depend on their temperature alone, but also on their volume. This is a consequence of the attraction of the van der

Waals forces between molecules. When the volume increases, the mean distance between molecules increases as well. The potential energy decreases in absolute value, and the internal energy increases, being the potential negative energy.

The Joule-Thomson effect is used, in practice, to liquefy gases. The gas is cooled, pushing it through a narrow orifice. The process is repeated until the point when the condensation temperature is reached. Hydrogen and helium must be preliminary cooled with a service gas below the inversion temperature.

4.5 Internal Energy and Entropy of Gases

The Joule-Thomson effect has shown us that the internal energy of a real gas depends not only on temperature but also on volume. We shall now see how the equation of state of a fluid contains the information on the dependence or not of its internal energy on the volume. Our arguments will be general, valid for any homogeneous thermodynamic system, the state of which is identified by pressure, volume and temperature. We shall find a general expression for the internal energy in terms of measurable quantities. We shall then apply that expression to the ideal gas and to the van der Waals fluid.

Consider such a thermodynamic system of *n* moles. The three state variables are, in all cases, linked by a state equation, which, for the moment, we do not specify. The independent variables are two. We choose the temperature *T* and the volume *V*. The internal energy is a function, *U*(*T*,*V*), of the two variables, which we generally do not know. Its total differential is

$$dU = \left(\frac{\partial U}{\partial T}\right)_V dT + \left(\frac{\partial U}{\partial V}\right)_T dV \qquad (4.17)$$

where, as usual, the subscript near the derivative symbol specifies the variable that is kept constant. We know that the derivative with respect to the temperature is

$$\left(\frac{\partial U}{\partial T}\right)_V = nC_V. \qquad (4.18)$$

We shall now find that the other partial derivative, the derivative with respect to the volume, is

$$\left(\frac{\partial U}{\partial V}\right)_T = -p + T\left(\frac{\partial p}{\partial T}\right)_V. \qquad (4.19)$$

To show that, we start from entropy, which is also a function of state, of the two variables *V* and *T*, *S*(*V*, *T*). Being a function of two variables, *S* has two mixed second derivatives. One is obtained deriving with respect to *V* first and then with respect to T, the other proceeding in the opposite order. Calculus shows that the two derivatives are equal:

$$\frac{\partial^2 S}{\partial V \partial T} = \frac{\partial^2 S}{\partial T \partial V}. \tag{4.20}$$

We shall use this relation shortly. The total differential of S is

$$dS = \left(\frac{\partial S}{\partial T}\right)_V dT + \left(\frac{\partial S}{\partial V}\right)_T dV. \tag{4.21}$$

On the other hand $dS = \delta Q/T$, and, for the first law, $\delta Q = dU + pdV$. Using Eq. (4.17) for dU, we obtain

$$dS = \frac{1}{T}\frac{\partial U}{\partial T}dT + \frac{1}{T}\left(\frac{\partial U}{\partial V} + p\right)dV \tag{4.22}$$

where we did not write the subscripts, because, in this section, we shall always use the variables T and V. Comparing with Eq. (4.21), we have

$$\frac{\partial S}{\partial T} = \frac{1}{T}\frac{\partial U}{\partial T}, \frac{\partial S}{\partial V} = \frac{1}{T}\left(\frac{\partial U}{\partial V} + p\right). \tag{4.23}$$

We now take the partial derivative with respect to V of the first equation and that with respect to T of the second and make them equal. Considering that the mixed partial derivatives of U are also equal and simplifying, we immediately get Eq. (4.19).

We shall now use this for the two cases in which we have the equation of state in analytical form, the ideal gas and the van der Waals fluid.

Ideal gas . The equation of state is

$$p = \frac{nRT}{V}. \tag{4.24}$$

Equation (4.19) gives us

$$\left(\frac{\partial U}{\partial T}\right)_V = -p + T\frac{nR}{V} = -T\frac{nR}{V} + T\frac{nR}{V} = 0. \tag{4.25}$$

We have found what we already knew, that the internal energy of the ideal gas does not depend on the volume. We have learnt that this fact is included in the equation of state.

van der Waals fluid . The equation of state is

$$p = \frac{nRT}{V - b} - \frac{a}{V^2}. \tag{4.26}$$

We proceed as in the case of the ideal gas. We take the derivative of the pressure with respect to the temperature (at constant V) and we substitute the result along with p as given by Eq. (4.26). We find

$$\left(\frac{\partial U}{\partial V}\right)_T = \frac{a}{V^2}. \tag{4.27}$$

As we expected, the internal energy depends on the volume as well.

We want now to find the expression of the internal energy. We start by integrating Eq. (4.27) on the volume

$$U = \int \left(\frac{\partial U}{\partial V}\right)_T dV = \int \frac{a}{V^2} dV = -\frac{a}{V} + C(T). \tag{4.28}$$

The quantity $C(T)$ on the right hand side is the integration "constant". Notice that it is constant relative to V, but not necessarily relative to T. On the other hand,

$$\left(\frac{\partial U}{\partial T}\right)_V = nC_V = \frac{dC(T)}{dT}.$$

We cannot proceed further if we do not know how C_V depends on the temperature. Let us assume, as usual, that we can consider it to be constant (which is true in limited temperature intervals). Under this hypothesis, we get

$$C(T) = nC_VT + \text{const}$$

where const is the integration constant. In conclusion, the internal energy of the van der Waals fluid is

$$U(T, V) = nC_VT - a/V + \text{const}. \tag{4.29}$$

Let us discuss what we have found. The energy of a van der Waals fluid is the sum of two terms. The first term is equal to the unique term of the ideal gas. As we shall see in Chap. 5, in both cases, it is the kinetic energy of the molecules, which is proportional to the absolute temperature. The second term is negative and inversely proportional to the volume. It includes the potential energy of the interaction between molecules, and is not present in the ideal gas, in which the interaction between molecules is neglected.

Let us now find the expressions for the entropy. We already know it for the ideal gas (Eq. 3.32). The entropy difference between state f and state i is

$$S_f - S_i = nC_V \ln\frac{T_f}{T_i} + nR\ln\frac{V_f}{V_i}. \tag{4.30}$$

For the van der Waals fluid, we proceed as we did for the ideal gas in Sect. 3.7. For a reversible process from i to f, we have

$$S_f - S_i = \int_i^f \frac{dU + \delta W}{T}.$$

The total differential of the internal energy is

$$dU = \left(\frac{\partial U}{\partial T}\right)_V dT + \left(\frac{\partial U}{\partial V}\right)_T dV = nC_V dT + \frac{a}{V^2} dV.$$

The elementary work for the equation of state Eq. (4.26) is

$$\delta W = p\,dV = \left(\frac{nRT}{V-b} - \frac{a}{V^2}\right) dV$$

We substitute and simplify, obtaining

$$dS = \frac{dU + \delta W}{T} = nC_V \frac{dT}{T} + nR \frac{dV}{V-b}.$$

Integrating this expression, we get

$$S_f - S_i = nC_V \ln \frac{T_f}{T_i} + nR \ln \frac{V_f - b}{V_i - b}. \qquad (4.31)$$

The expression is very similar to that of the ideal gas. The only difference is that the volumes diminished by the covolume appear in place of the volumes. This takes into account the fact that the volume occupied by the molecules is not available. Notice that the term a/V^2, which takes into account the interaction between molecules in the van der Waals equation, does not have any effect. The entropy does not change when one introduces the intermolecular force. The microscopic interpretation that we shall discuss in Sect. 5.11 will explain this effect too.

4.6 Clapeyron Equation

In Sect. 4.2, we saw that the vapor-liquid phase transition of a substance at a given pressure takes place at a certain temperature. In a closed container, maintained at constant temperature, with no other gas present, the vapor above the free surface of the liquid naturally reaches the pressure at which the two phases are in equilibrium: the saturated vapor pressure at that temperature. The same is true for the other transitions between aggregation phases; they are in equilibrium at a given temperature only at a certain pressure. In an equivalent manner, we can say that the aggregation phase transitions take place at a constant, well-defined temperature, once the pressure has been fixed. As we have seen in the case of vaporization, throughout the aggregation phase transition, a certain quantity of heat is absorbed by, or delivered to, the substance. These quantities are called *latent heats* of vaporization , fusion and sublimation , for the liquid-vapor, the solid-liquid and solid-vapor phase transitions. The latent heats are *specific* if referred to a kilogram, *molar* if referred to a mole. Notice that by solid phase, we mean the crystalline phase. Amorphous solids like wax, black tar, glass, etc., are not really solids, but liquids with extremely high viscosity. These materials become softer and softer with increasing temperature. There is no latent heat in these cases.

For completeness, we mention that additional phases exist along with the aggregation phases. Some substances can crystallize in different forms, which are called *allotropic phases* . Allotropy is a rather common phenomenon. For example, carbon has four allotropes: the diamond (in which the lattice is tetrahedral), the graphite (in which the carbon atoms are arranged in sheets of hexagonal cells), the grapheme (a single graphite sheet) and fullerene (in which the atoms are in a closed surface at the vertices of 20 hexagons and 12 pentagons, as in a soccer ball). Another example is ice, of which nine allotropic phases are known. Latent heats also exist for the transitions between two allotropic phases.

The phase transitions usually take place at constant pressure. Consequently, the heat corresponding to the transition between phases 1 and 2, Q_{12}, is equal to the enthalpy difference between the two states. For this reason, the latent heats are also called latent enthalpies . Indeed, we have $Q_{12} = U(2) - U(1) + p\left[V(2) - V(1)\right] = U(2) + pV(2) - \left[U(1) + pV(1)\right]$, which is also

$$Q_{12} = H(2) - H(1). \tag{4.32}$$

The enthalpy being a function of state, the heat exchange in the opposite phase transition is equal and opposite, namely

$$Q_{21} = H(1) - H(2) = -Q_{12}. \tag{4.33}$$

We shall now find a relation between the derivative of the equilibrium pressure with respect to temperature, the latent heat and the change of volume in the corresponding transition. The equation holds for any transition between any pair of aggregation phases. It was found in 1834 by Benoît Paul Émile Clapeyron (France, 1799–1864) and is called the Clapeyron equation . It was later rediscovered by R. Clausius in 1857, and is sometimes called the Clapeyron-Clausius equation.

To be concrete, we consider the liquid to vapor transition. Figure 4.8 shows two isothermal curves, one at T and one at the infinitesimally higher temperature $T + dT$.

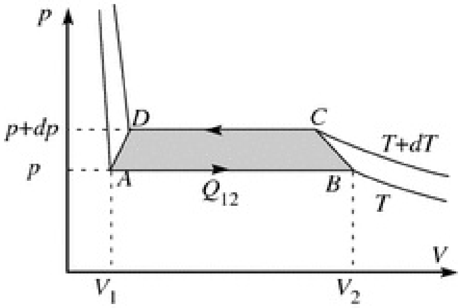


Fig. 4.8 Two isothermal curves in the liquid-vapor transition at infinitesimally small temperature difference

We consider a unit mass, 1 kg, of the substance. Its volume is the *specific volume* . In the initial state, it is liquid (point *A* in the figure) at the saturated vapor pressure, in thermal contact with a bath at temperature *T*. We deliver heat, and the fluid starts to evaporate, at constant pressure. The system has two phases, liquid and vapor, in equilibrium. The representative point in the diagram moves along the line *AB*. When it is in *B*, the absorbed heat is Q_{12}.

Once the system is in *B*, we move it into contact with a bath at the temperature $T + dT$. The state is point *C* in the figure. We now subtract heat, moving the point to *D*. This is, again, an isobaric transition at the saturated vapor pressure $p + dp$ corresponding to $T + dT$. We close the cycle by taking the system back to *T*. We performed all the processes reversibly.

The process can be considered a Carnot cycle . Indeed, it is composed of two isothermals (*AB* and *CD*) and the two other processes (*BC* and *DA*). Rigorously speaking, the latter are not necessarily adiabatic. However, the corresponding exchanged heats are infinitesimal, the temperature difference being infinitesimal, and can be neglected in comparison with the finite heats exchanged in the isothermal transformations. The efficiency is consequently dT/T.

The work is given by the area of the cycle, which is $(V_2 – V_1)dp$, higher order infinitesimals apart, where V_1 is the volume of the liquid and V_2 the volume of the vapor. This work is equal to the absorbed heat Q_{12} times the efficiency. We have

$$(V_2 - V_1)\,dp = Q_{12}\frac{dT}{T},$$

from which we obtain

$$\frac{dp}{dT} = \frac{Q_{12}}{T\,(V_2 - V_1)}. \tag{4.34}$$

This is the Clapeyron equation .

We observe that the volumes of the phases and the latent heat are functions of the temperature. Consequently, Eq. (4.34) is generally not sufficient to know the function $p(T)$. However, it contains a lot of information. Let us see.

First, we observe that the slope of the equilibrium curve $p(T)$ is inversely proportional to the difference of the volumes of the phases. Consequently, it is small for the vaporization (because the volume of the vapor is much larger than that of the liquid) and very large for the fusion (because the volumes of liquid and solid do not differ much).

Consider the important example of water . At $T = 273$ K (0 °C), the specific volumes of the solid (ice) and liquid are $V_{\text{sol}} = 1.09 \times 10^{-3}\ \text{m}^3/\text{kg}$ and $V_{\text{liq}} = 1.0 \times 10^{-3}\ \text{m}^3/\text{kg}$, very close indeed. The specific fusion heat of ice at 273 K is $Q_f = 335$ kJ/kg. The slope of the equilibrium curve given by the Clapeyron equation is, then, $dp/dT = -1.4 \times 10^7\ \text{Pa K}^{-1} \approx -140\ \text{atm K}^{-1}$. This value agrees very well with the experimental one. The curve is extremely steep. For example, to change the ice fusion temperature by one degree only, we must change the pressure of 140 atmospheres. Notice the minus sign. It means that the equilibrium pressure diminishes with increasing temperature. This behavior is exceptional and characteristic of water (and a few other substances), corresponding to the fact that water expands when cooling ($V_{\text{liq}} - V_{\text{sol}} < 0$) between 0 and 4 °C (see Fig. 2.21b).

This property is the cause of the *regelation* phenomenon, in which ice melts under pressure and freezes again when the pressure is reduced. Regelation helps the glaciers to flow under the effect of their weight. Consider, for example, a narrow in the valley of the glacier. The pressure in the ice near the bottom in contact with the rocks of the narrow is particularly high due to the weight of the upstream part of the glacier. The fusion temperature is consequently lower than the ice temperature, which might be around −10 °C, and some ice melts. The resulting water flows beyond the narrow, where the pressure is lower, and refreezes. The glacier can flow in its bed, overtaking curves and narrows somewhat, as if it were plastic.

We can make a qualitative demonstration looping a thin metal wire over an ice block and attaching two heavy metallic weights to its ends. The weights have two functions. They give a tension to the wire, which exerts a constant pressure on the ice, and they act as heat reservoirs providing the melting heat, through the conductive wire. Under pressure, the ice under the wire melts; the wire descends a bit through the resulting water, which refreezes behind the wire, and so on. The result, after several minutes, is that the wire has passed through the ice, leaving the ice block apparently intact.

At T = 373 K (100 °C), the volume of the water vapor is about 1600 times the liquid volume and, for a kilo, is $V_{\text{vap}} = 1.7\ \text{m}^3/\text{kg}$. The condensation heat at 373 K is $Q_{\text{con}} = 2250$ kJ/kg. The Clapeyron equation gives the slope of the equilibrium pressure as $dp/dT = 3.5 \times 10^3\ \text{Pa K}^{-1} \approx 3.5 \times 10^{-2}\ \text{atm K}^{-1}$, in good agreement with the experimental value. To change, for example, the equilibrium temperature by one degree, it is enough to change the pressure by 3.5 hundredths of atmosphere.

Table 4.2 reports the fusion temperatures, in °C, of the molar and specific fusion heats for several substances.

Table 4.2 Fusion temperature and specific and molar latent heats

Substance	**Fusion temp. (°C)**	**Q_{fm} (kJ mol^{-1})**	**Q_{f} (kJ kg^{-1})**
Aluminum	658.5	10.68	395.6
Argon	–190.2	1.21	30.4
Calcium	851	9.34	233.2
Cesium	28.3	2.09	15.5
Cobalt	1490	15.24	260
CO_2	–57.6	7.95	180.9
Copper	1083	13.02	205.1
Hydrogen	–259.25	0.12	57.8
Iodine	112.9	15.28	59.9
Iron	1530	14.9	266.7
Lead	327.3	5.12	24.7
Mercury	–39	2.33	11.3
Nitrogen	–210	0.72	25.7
NaCl	800	30.23	517.1
NaFl	992	29.31	697.9
Oxygen	–218.8	0.45	13.8
Potassium	63.4	2.4	61.1
Silicon	1427	39.65	1410.9
Silver	961	11.3	104.7
Tin	231.7	7.2	60.3
Tungsten	3387	35.25	191.8
Water	0	6.01	335.0

4.7 Vaporization

The Clapeyron equation for vaporization of one mole of fluid is

$$\frac{dp}{dT}=\frac{Q_{\text{vap}}}{T\left(V_g-V_l\right)} \tag{4.35}$$

where Q_{vap} is the molar vaporization heat and V_g and V_l are the molar volumes of the gas and liquid phases, respectively. The three quantities are functions of the temperature. For the liquid-gas transformation, it is possible to find an approximate expression of the saturated vapor pressure function $p(T)$.

We have already noted that the volume of the liquid is usually much smaller (a few per mille) than the volume of the gas. Consequently, we neglect V_l in the above equation. At temperatures substantially lower than the critical temperature, the saturated vapor pressure is rather low and the vapor behaves, in first approximation, as an ideal gas. In this approximation, we can write $V_g=RT/p$. Equation (4.35) becomes

$$\frac{dp}{dT}=\frac{pQ_{\text{vap}}}{RT^2}. \tag{4.36}$$

The next approximation is to consider the vaporization heat independent of temperature. To have an idea of its variations, consider that, for example, the vaporization heat of water varies by about 10 % between 0 and 100 °C. In these approximations, we can separate the variables and write

$$\frac{dp}{p}=\frac{Q_{\text{vap}}}{R}\frac{dT}{T^2}, \tag{4.37}$$

which is immediately integrated into $\ln p=-\frac{Q_{\text{vap}}}{RT}+\text{const}$ and, taking the exponentials,

$$p=C\exp\left(-\frac{Q_{\text{vap}}}{RT}\right) \tag{4.38}$$

where C is now the integration constant. Notice that T is in the denominator of the negative exponent. The saturated vapor pressure increases very rapidly in temperature (and exponentially decreases with the inverse of the temperature). We shall learn the physical reason for that in Chap. 5.

Here, we observe that the molecules in a liquid are subject to the van der Waals forces, which are practically irrelevant in the vapor. This means that, in the vaporization process, work must be done against these forces to bring each molecule from the liquid to the vapor. This work is equal to the potential difference of the molecule in the vapor and in the liquid, which is the molar vaporization heat divided by the Avogadro number Q_{vap}/N_A . Then, by measuring Q_{vap}, we can evaluate the depth of the minimum of the van der Waals potential . One obtains values ranging, for the different gases, from a few tenths of an electronvolt to a few electronvolt. These are the values we mentioned in Sect. 4.1.

Table 4.3 gives the vaporization temperatures and specific heats for several substances.

Table 4.3 Vaporization temperatures T_v and heats Q_{ev} of several substances

Substance	**T_v (°C)**	**Q_{ev} (kJ kg^{-1})**
Water	100	2250
CO_2	–60	365
Ammonia	–33.4	1369
Ethyl alcohol	78.3	854
Benzene	80.2	395
Oxygen	–182.9	213
Nitrogen	–195.6	199
Hydrogen	–252.8	452
Helium	–268.6	25

4.8 Pressure-Temperature Dyagrams

The Clapeyron equation Eq. (4.34) expresses the derivative of the equilibrium pressure $p(T)$ between any two phases. We shall now discuss more of its properties in the plane p,T, as in Fig. 4.9.

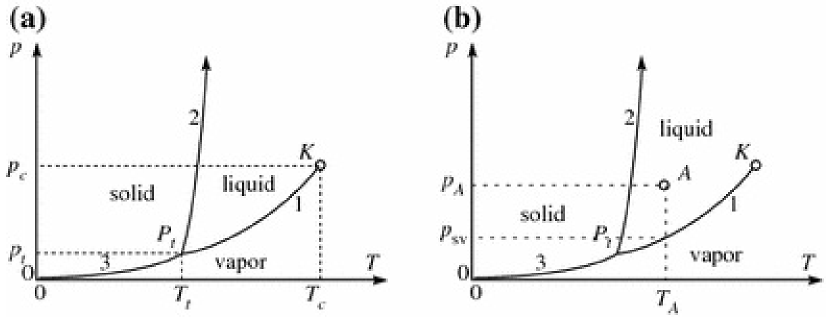


Fig. 4.9 The coexistence curves between the aggregation phase couples

Consider Fig. 4.9 a. The three branches of the diagram, 1, 2 and 3 in the figure, represent the function $p(T)$ for the equilibrium pressure between the three pairs of phases. We repeat that the substance under study must be in a closed container, with no other foreign substance, air, in particular, being present. Otherwise, the vapor pressure is the partial pressure of the mixture, while the liquid, or the solid, are at the total pressure, which is the sum of the vapor and extraneous gas pressures. Let us consider the state A in Fig. 4.9b, where, according to the diagram, the substance is completely liquid. In practice, the piston, with which we exert the pressure (p_A), must be in direct contact with the liquid. Indeed, assume that some air is present between the free surface of the liquid and the piston, which still exerts the (total) pressure p_A. Under these conditions, part of the liquid vaporizes. The equilibrium between the two phases is reached when the (partial) pressure of the vapor is the saturated vapor pressure at the existing temperature T_A (p_{sv}) in the figure. The two phases are in equilibrium at a total pressure different from the saturated vapor pressure.

Consider now branch 1 in Fig. 4.9a, which is the curve of the saturated vapor pressure as a function of temperature. We have found the approximate expression of this function of Eq. (4.38). When the temperature increases, the saturated vapor pressure, and consequently the density, grows very rapidly. The vapor density becomes closer and closer to the liquid density. When the state reaches the critical point, the vapor has the same density as the liquid and the two phases are indistinguishable. In other words, the liquid-vapor equilibrium curve ends in the critical point (K in the figure).

Consider two states at the same temperature, A in the liquid phase, B in the vapor phase. Figure 4.10a shows the situation in the p,T plane, Fig. 4.10b in the V, p plane. We can pass from A to B, compressing the fluid at constant temperature, as shown by the continuous line in Fig. 4.10a. During the process, the two phases separate, becoming different from one another. However, we can go from A to B in another way too, which is the dotted curve in Fig. 4.10a and b. We start by compressing the gas, increasing its temperature, in a way that does not liquefy it. Once we are above the critical temperature, we continue the compression, but now decreasing the temperature. In this way, as shown in the figures, we go around the critical point. In this process, there is no sharp change of phase; the substance always remains homogeneous and it is impossible to determine where the vapor became liquid.

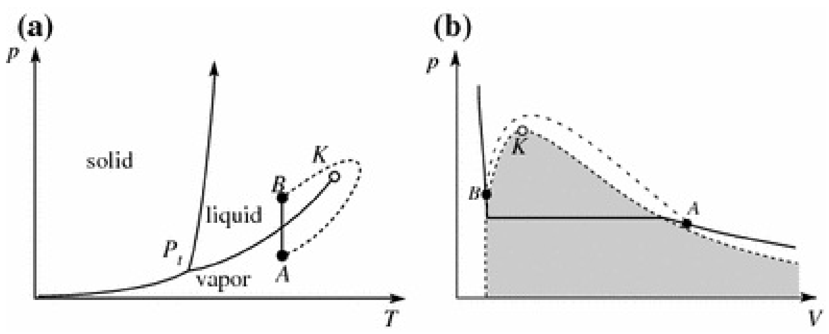


Fig. 4.10 Two processes for the condensation from state *A* to state *B*. **a** on the *p*, *T* plane, **b** on the *p*, *V* plane

An important difference between the solids (crystals) and the liquids is the anisotropy of the crystals. In a solid, as opposed to a liquid, privileged directions exist. Consequently, the transition between liquid and solid cannot in any way take place continuously as is the case between vapor and liquid. As a consequence, no critical point exists on the solid-liquid coexistence curve. The curve continues indefinitely, as indicated with an arrow on branch 2. Also notice that, as already discussed in Sect. 4.7, the liquid-solid coexistence curve (branch 2) is much steeper than the liquid-vapor curve (branch 1).

In Fig. 4.10, the liquid-solid coexistence curve is represented in the most common situation of bodies that expand when melting. As we noticed in Sect. 4.6, the opposite is true in a few cases, such as water , for which the density of the solid is smaller than the density of the liquid. The slope of the coexistence curve is opposite, as in Fig. 4.11.

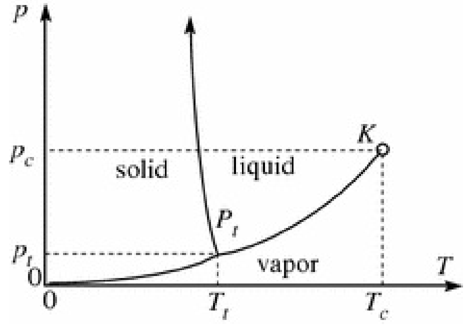


Fig. 4.11 The coexistence curves for water

The equilibrium state between the three phases must lay contemporarily on all the coexistence curves. The three curves intersect at a point, which is the *triple point* , P_t in the figures. The equilibrium of the three phases exists only at well-defined values of the three variables: temperature, pressure and, for a given mass, volume.

Consider the example of water. The equilibrium between ice, water and water vapor is only at a pressure of about a 600 Pa (about 6 thousandths of an atmosphere) and the temperature of +0.01 °C. The temperature of the triple points being completely determined, in particular, independent of pressure, they are extremely useful as standards of the temperature scale. As we have seen, the temperature unit, the kelvin , is defined by fixing the water triple point temperature at 273.16 K.

Consider now the solid-vapor coexistence curve , branch 3 of the diagrams. At temperatures lower than the triple point temperature (or a little less than that for water), the liquid phase does not exist. Heating (giving heat to) the solid, it passes directly into the vapor phase, and vice versa when taking out heat. The former process is called *sublimation* , the latter *deposition* .

For example, the triple point of carbon dioxide is at the pressure of 510 kPa (about 5.1 atm) and at the temperature of—56.6 °C. Under normal conditions, the solid CO_2 does not melt, it sublimates. For this reason, it is called dry ice . Similarly, snow and ice in the glaciers and elsewhere sublimate when the atmospheric temperature is below 0 °C. An example of deposition is the formation of snow in the clouds and frost on the ground.

For the same reasons discussed for the liquid-solid, continuous transition between vapor and solid phases is impossible. The solid-vapor coexistence curve does not have an end point.

4.9 Surface Tension

Up to now, we have studied three-dimensional thermodynamic systems. The relevant geometrical quantity was the volume. The free surfaces of the bodies are the sites of another class of phenomena, the *surface phenomena* , which we shall study in this section and the following two.

Consider, for example, water in a glass, filled close but not all the way to the rim. Its volume is limited by an upper surface in contact with the air, which is almost a horizontal plane, but not completely so, as we shall soon see, and by the lateral and bottom surfaces that are in contact with the glass. The water-glass and the water-air contact surfaces intersect along a line, which is a circle in this case. Similarly, the air volume in the glass above the water is limited by the same water-air contact surface and by a glass-air contact surface. The three contact surfaces meet in the intersection circle.

In the surface phenomena, only the molecules near enough to the surfaces intervene. Surface phenomena are important in two different, and correlated, cases that we shall now discuss: when a liquid is in a capillary tube (a tube of millimetre scale diameter) and for a drop on a solid surface .

The conditions of a molecule at a distance from the surface smaller than the molecular action radius are different from those in the bulk. The latter are surrounded by equal molecules in all directions, the former only on one side. Suppose now that we want to bring a molecule located in the bulk of the volume to the surface. We shall need to perform a positive work against the resultant of the van der Waals forces, which is directed inside. This work is equal to the difference between the energy of the molecule on the surface and in the bulk of the fluid.

We have already discussed this issue in Sect. 4.3; let us repeat the argument more precisely, also taking into account the molecules of the medium that limits the fluid (for example, the walls of the container or the air). Figure 4.12 shows the molecular action sphere of a molecule on the surface; half of the sphere contains molecules from the fluid, half molecules from the external medium, which are generally different and have different density. The force on the molecule is the resultant of the forces of the two types of molecules. To take a molecule to the surface, we must work against this resultant. Consequently, the surface energy of a molecule depends on *both* media.

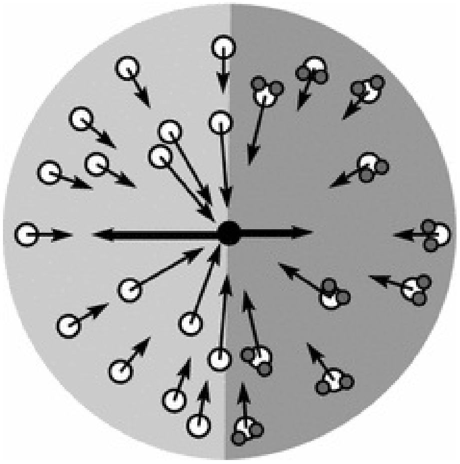


Fig. 4.12 Cartoon showing the van der Waals forces exerted by the molecules of two media in contact on a surface molecule and their resultants

The number of molecules on a surface, or better yet, in a thin layer one molecular action radius thick, of area A is proportional to A. And so, obviously, is their energy, which is called *surface energy* and which we indicate with U_{sur}, writing

$$U_{\text{sur}} = \tau A. \tag{4.39}$$

The proportionality constant τ is the surface energy per unit area and is called *surface tension* . It depends, as we mentioned, on both media in contact and is a function, in general decreasing, of temperature.

The presence of the surface energy is evident in the drops. A small quantity of liquid spontaneously takes a spherical shape, because this is the shape that minimizes the surface area for a given volume (of liquid). As a matter of fact, drops are perfectly spherical only in absence of other forces, in particular, of weight. As an illustration, this is what happens in a spacecraft. The shape of a drop of mercury, for example, lying on a plane is somewhat flattened. This is now the condition of minimal total energy, which is the sum of the surface and volume (weight) energies. We shall come back to this phenomenon later on. Here, we observe that it follows from the above considerations that the surface energy cannot be negative. If that were the case, two media in contact could not exist separately; the separation surface would tend to increase indefinitely and the two media would mix completely.

We now consider a simple experiment showing the action of the surface tension. Figure 4.13 represents a small metal frame with three fixed sides; one of the short sides is free to slide. We momentarily block the mobile side and dip the frame in soapy water. If we take it out smoothly we are able to have a film inside the frame like those of the soap bubbles. We can evaluate the order of magnitude of the thickness of the film, which is clearly very thin, with a simple argument. The beautiful colors of the soap bubbles are due to the interference of light reflected at the forward and backward surfaces of the film. The phenomenon will be explained in the 4th volume of the course. It happens when the thickness of the film is on the order of half a wavelength, which is a few hundredths of nanometers. In conclusion, the film thickness is two orders of magnitude larger than the molecular action radius.

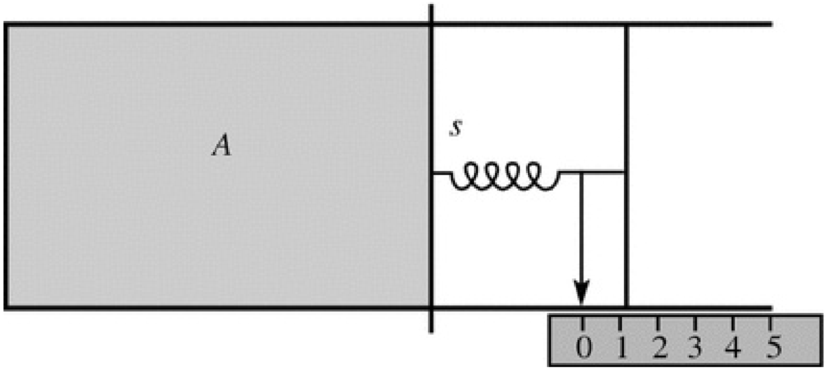


Fig. 4.13 Measuring the force on a side of a soapy water film

If we now unblock the mobile side, we notice the presence of a force exerted by the film, which tends to reduce its area as much as possible. We can measure this force by inserting a spring between the mobile and a fixed side, as in the figure, and measuring its stretch. We find that the force is proportional to the length *s* of the side:

$$F = 2\tau' s. \tag{4.40}$$

We have introduced the factor two, because the soapy water film has *two* free surfaces, one on each side. Consequently, the proportionality constant τ' is the force per unit length tangent to the film.

Let us now move the mobile side by an infinitesimal distance dx, increasing the total free surface by $dA = 2sdx$. The work to be done against the force F is equal to the increase in the surface energy due to this increase of surface. Notice that the surface increases but the volume of the liquid remains constant; an increasing number of molecules pass from the interior to the surface of the film. In formulae, we have $F dx = 2\tau' s dx = 2\tau dA = 2\tau s dx$. Hence, it is just $\tau' = \tau$.

We see that the surface tension has two physical meanings; it is the surface energy per unit area and the tension force on the surface per unit length. This force is present at all the surface points, not only on the borders. It is the two-dimensional analogue of the pressure in a fluid, which is present at all its points, not only on the surface. On the borders, the surface tension force is directed perpendicularly toward the interior (this is equivalent to saying that $\tau > 0$).

We can see that we have a very simple experiment. We spread a film of soapy water on a metallic frame. We lay on it a sewing thread the ends of which we had joined together to make a loop (see Fig. 4.14a). The thread forms an irregular shape. We now break the film with a pin inside the loop. The thread immediately becomes a circle (see Fig. 4.14b). Indeed, all its elements are pulled by the surface tension. Before the film was broken, they were pulled by equal and opposite forces on both sides.

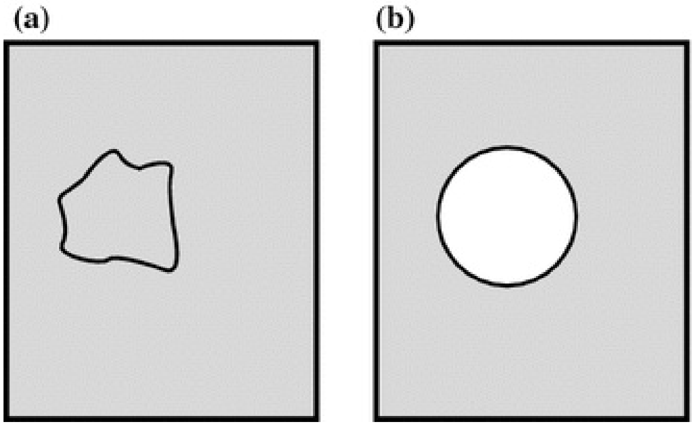


Fig. 4.14 A soapy water film **a** with a wire loop, **b** after breaking the film in the loop

The physical dimensions of the surface tension, on the basis of its two meanings, are of an energy divided by an area or of a force divided by a length, $[\tau] = \mathrm{J\,m^{-2}} = \mathrm{N\,m^{-1}}$.

As we mentioned, the surface tension depends on the two media in contact. However, we also speak of surface tension of one liquid, meaning of the liquid and its vapor. The surface tension of a liquid decreases with increasing temperature, reaching zero at the critical point. Here, we know, there is no difference between liquid and vapor and no separation surface can exist.

To get an idea of the orders of magnitude, at 20 °C, the surface tension of water is 73 mN/m, of mercury, 480 mN/m. Table 4.4 gives the values of the surface tension for several combinations of media and, for the water-air pair, at a number of temperatures.

Table 4.4 Surface tension τ for several combinations of media

Media	T (°C)	τ (mN/m)
Water-air	0	75.6
	10	74.22
	20	72.75
	30	71.18
	40	69.56
	50	67.91
	60	66.18
	70	64.4
	80	62.6
	100	58.9
Benzene-Mercury	20	357

Benzene-Water	20	35
Water-Ethyl ether	20	10,7
Water-Mercury	20	375
Argon liquid-vapor	–188	13.2
Hydrogen liquid-vapor	–255	2.31
Nitrogen liquid-vapor	–183	6.6
Oxygen liquid-vapor	–183	13.2
Ethyl alcohol liquid-vapor	20	17.01
Benzene-Air	20	28.88
Ethyl alcohol -Air	20	17
Ethyl ether-Air	20	17

4.10 Capillary Phenomena

Consider a liquid in a container; let us stick with water in a glass. We are dealing with three media, the solid of which the container is made (medium *S*), the liquid (medium *L*) and the gas or vapor above the liquid (medium *G*) (glass, water and air in the example), and with three surfaces separating the three combinations of media. We shall call them S_{SL}, S_{SG} and S_{LG}. We observe that the free surface of the liquid is not flat 1 mm or so near its border (where the water touches the glass) but slightly deformed. The water rim is higher than the free surface. The rim is called the *meniscus*, specifically the *concave meniscus* , because it rises. It is shown magnified in Fig. 4.15a. For mercury in glass, the rim is turned downward, and is called the *convex meniscus* (Fig. 4.15b. Water in a Teflon container behaves in the same way as mercury in glass; its meniscus is convex.

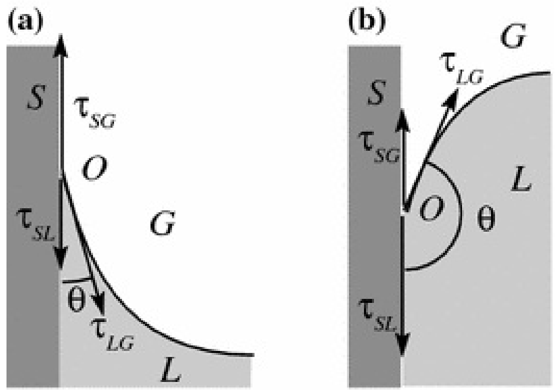


Fig. 4.15 **a** Concave and **b** convex menisci

The border of the liquid is the curve where the three separation surfaces S_{SL},

S_{SG} and S_{LG} meet. It cuts the plane of the figures at point O. The border of the liquid is also the border of each separation surface. Consequently, along each segment of the border, the forces resulting from the three surface tensions are present. They are all perpendicular to the segment, each on the plane of the corresponding surface toward the inside. The magnitudes of the three forces per unit length are the surface tensions relative to the two media in contact (τ_{SL}, τ_{SG}, τ_{LG}). This situation is shown in Fig. 4.15. The shapes of the solid-liquid and solid-gas surfaces S_{SL} and S_{SG} are a priori defined because medium 1 is solid, while the shape of the liquid-gas interface S_{LG} is not because both media are deformable. This surface adjusts itself at the angle with the solid wall θ at which the equilibrium of the forces is reached. This is called the *contact angle* .

At equilibrium, the component of the resultant of the three tension forces tangent to the solid surface must be zero. Otherwise, the border would move up or down. Notice that the normal component can be and is different from zero, being equilibrated by the adhesion forces.

In formulae, we have $\tau_{SG} = \tau_{SL} + \tau_{LG} \cos\theta$, or

$$\cos\theta = \frac{\tau_{SG} - \tau_{SL}}{\tau_{LG}}. \tag{4.41}$$

This is the Young equation , after Thomas Young (UK, 1773–1829), who established it in 1804. One sees how the contact angle depends on all **three pairs** of media. Somewhat surprising, it also depends on τ_{SG}, the solid-gas (glass-air in the example) surface tension. To understand the reason for this, let us find the Young equation in a different way, using the virtual works principle .

Consider each separation surface in turn. Let us start with the liquid-solid interface (S_{SL}). The work needed to take a molecule from inside the liquid to the contact surface with the solid depends on the nature of both media, namely on the van der Waals forces of the molecules of both of them. The surface energy per unit area of S_{SL}, which is the surface tension τ_{SL}, consequently depends on both as well. Similarly, the work needed to take a molecule from inside the gas to the contact surface with the solid, namely to build S_{SG}, depends on the gas and the solid. And so it is for τ_{SG}. The liquid-gas interface S_{LG} is a bit more complicated. In this case, to build a piece of surface, we must take molecules from inside the liquid to the surface as well as gas molecules from inside the gas to the surface, because both must increase. The surface energy per unit area τ_{LG} includes both works. Note that a similar problem did not exist in the other two interfaces because the solid molecules did not move.

We now use the virtual works principle to establish the equilibrium conditions of the meniscus. Let z be a coordinate vertically directed upward through point O in the figure. Consider a segment of the border of length Δl near O and its virtual displacement dz. The corresponding variations of the three surfaces are $dS_{SL} = \Delta l dz$, $dS_{SG} = -\Delta l dz$ and $dS_{LG} = \Delta l dz \cos\theta$. Notice that, in particular, the area of the contact surface of the gas with the solid wall varies too; hence, the energy of the S_{SG} surface varies. This explains why the equilibrium conditions must include τ_{SG}. The necessary condition for equilibrium is that the total virtual work be zero, namely $\tau_{SL}\Delta l dz - \tau_{SG}\Delta l dz + \tau_{LG}\cos\theta\Delta l dz = 0$. This immediately gives Eq. (4.41).

We now discuss the principal consequences of the Young equation for the cases of the menisci (Fig. 4.15) and of a liquid drop on a solid surface (Fig. 4.16). We call the ability of the liquid to maintain contact with the solid surface *wetting*. If the solid-gas surface energy is larger than that of the solid-liquid, $\tau_{SG} > \tau_{SL}$, then $\cos\theta > 0$ and the angle θ is acute. We talk of *high wettability*. The meniscus is concave, as in Fig. 4.15a. A drop of liquid on a solid surface has a shape as shown in Fig. 4.16a. The extreme case, $\cos\theta = 1$, is called *complete wetting* . In Fig. 4.16b, the wetting is almost complete.

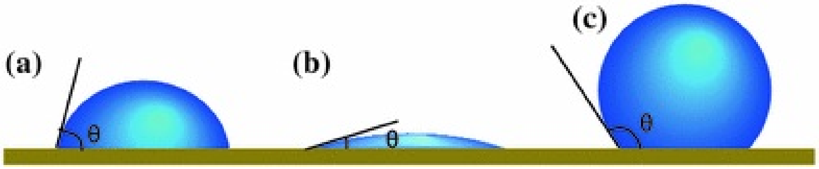


Fig. 4.16 A drop on a solid surface in air with wettability **a** high; **b** almost perfect; **c** low

If the solid-gas surface energy is smaller than the solid-liquid surface energy, $\tau_{SG} < \tau_{SL}$, then $\cos\theta < 0$, and the angle θ is obtuse. We talk of *low wettability*. The meniscus is convex, as in Fig. 4.15b; the form of the drop is as shown in Fig. 4.16c. For example, the mercury-glass contact angle in the air is $\theta \sim 150°$. The extreme case, $\cos\theta = -1$, is called *perfect non-wetting*. Particularly important is the case in which the gas is air and the liquid is water. Solid surfaces are called *hydrophobic* (from the Greek words for “fearing water”) if the contact angle with water in air is $\theta > 90°$, and *hydrophilic* (from the Greek for “loving water”) if $\theta < 90°$.

We notice that the Young equation assumes a perfectly flat and rigid solid surface. In practice, the surface tensions, and consequently the contact angle, strongly depend on the structure of the surface at microscopic and nanoscopic levels. For example, the water-glass contact angle in the air depends on the treatment of the glass surface and on the purity of the water. Untreated glass is hydrophilic for pure water, the contact angle being in the range $\theta = 25°–30°$. As another example, Teflon is hydrophobic, the contact angle with water in the air being $\theta \sim 110°$. The study of these properties is an important chapter in contemporary research. Microscopic and nanoscopic architectures are used by nature on the surfaces of vegetables and animals and, once understood, can be used by man to produce new products. One example is the use of a nano-structure to minimize the droplet's adhesion to the surface (superhydrofobicity) by the lotus flower (Nelumbo nucifera and N. lutea) for self-cleaning. Water droplets do not remain on these surfaces stably. They spontaneously roll off with only a slight tremble, picking up and removing any dust particles they may touch. Water striders are insects using superhydrofobicity to walk on water. Figure 4.17 shows a gerridae (Aquarius paludum), on a pond surface. The insect has four long legs used to walk and two short ones to catch prey. The surfaces of the long legs are covered with specialized tiny hairs, spaced by a several micrometres, resulting in a hydrophobic surface.

Fig. 4.17 Aquarius paludum. Photo of Andrej Gogala, by kind permission

As we saw in Sect. 1.4, a consequence of the Stevin law is that the free surfaces of a homogeneous liquid in communicating vessels in equilibrium have equal heights. This is not true if one of the containers is a capillary, namely if its section is small, on the order of one millimeter. Consider two communicating vessels, one capillary and one not. If we fill them with a wetting liquid, the level of the free surface in the capillary is higher than the level in the wide vessel. Contrastingly, if the liquid does not wet, the level of the free surface in the capillary is lower. This phenomenon, called *capillary ascension* , is shown in Fig. 4.18. It is found that the level difference h is larger the smaller the diameter of the capillary. Let us try to understand these phenomena.

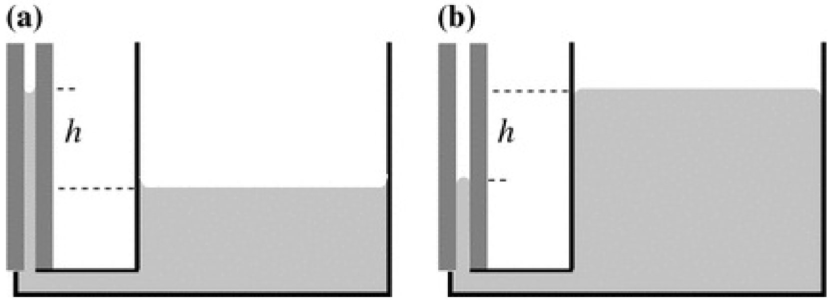


Fig. 4.18 Capillary ascensions. **a** Concave meniscus, **b** convex meniscus

Let us start by considering a spherical water drop of radius r in air. The surface tension tends to reduce the free surface as much as possible (to minimize the surface energy) and consequently the volume. The pressure inside the drop becomes higher than the pressure of the air outside. Let us call p_{dif} the overpressure, namely the *difference* between the internal and external pressures. To calculate this quantity, we observe that the work done by the surface forces to decrease the free surface by a generic dS is $\tau_{LG}\, dS$, where the water-air surface tension is. The same work can be expressed as the work to reduce the volume by dV, namely $p_{\mathrm{dif}}\, dV$. In formulae, we have

$$\tau_{LG}\, dS = p_{\mathrm{dif}}\, dV. \tag{4.41}$$

The surface of the spherical drop of radius r is $S = 4\pi r^2$, and its volume is $V = 4\pi r^3/3$. Differentiating and using Eq. (4.41), we obtain

$$p_{\mathrm{dif}} = 2\tau_{LG}/r, \tag{4.42}$$

which is called the *capillary pressure* . We can also say that Eq. (4.42) gives the overpressure under a surface of radius r relative to the pressure under a flat surface ($r = \infty$). The argument is still valid for a bubble of gas in liquid; the pressure is larger in the bubble than in the liquid. In general, there is an over-pressure under any concave surface. Contrastingly, there is an under-pressure under a convex surface. Equation (4.42) is valid, in absolute value, in both cases. Note that p_{dif} is inversely proportional to the curvature radius. As we have already noticed, when $r \to \infty$, when the surface becomes flat, the pressure difference goes to zero. The equilibrium pressures on the two sides of a flat separation surface are equal. The tendency of the surface tension to reduce the surface exists in this case as well, but the forces are parallel to the surface, with no component towards the interior.

To get an idea of the orders of magnitude, consider an air bubble in pure water at 25 °C. The surface tension is τ_{LG} = 72 mN/m (see Table 4.2). The overpressure in a bubble of $2r$ = 1 mm diameter is negligible (288 Pa), but it is already 96 kPa (i.e., almost one atmosphere) for a diameter $2r$ = 3 μm and 960 kPa (almost 10 armpsphere) $2r$ = 0.3 μm.

Let us come back to the capillary ascensions. Observations show that, when the radius *a* of the capillary is small enough, the free surface of the liquid is a spherical cap ("meniscus" comes from the Greek for crescent). The geometrical relation between the radius of the cap and *a* is immediately found looking at Fig. 4.19 (for a concave meniscus), namely

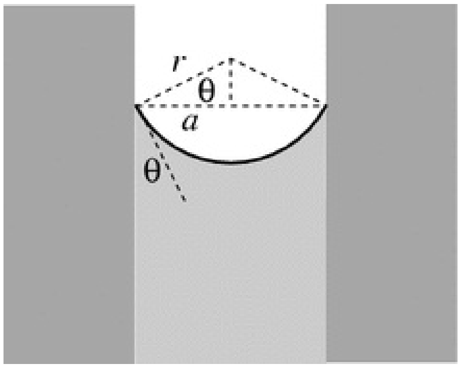


Fig. 4.19 Geometry of a concave meniscus in a capillary

$$a = r\cos\theta. \tag{4.43}$$

From Eq. (4.42), we then have

$$p_{\mathrm{dif}} = \frac{2\tau_{LG}\cos\theta}{a}. \tag{4.44}$$

The pressure in the liquid in the capillary beyond a concave separation surface from the air is *smaller* than the air pressure by p_{dif} in Eq. (4.44). The atmospheric pressure, on the other hand, is the same in the capillary and in the wider vessel. Consequently, the liquid in the capillary must rise to a height *h* such that the weight of the liquid column equilibrates the pressure difference, i.e., $\rho g h = p_{\mathrm{dif}}$, where ρ is the liquid density (Fig. 4.18a. Equation (4.44) gives

$$h = \frac{2\tau_{LG}\cos\theta}{\rho g a}. \tag{4.45}$$

If the meniscus is convex, the internal pressure in the liquid is larger than the atmospheric pressure, and the height of the meniscus in the capillary is lower than that in the wider vessel (Fig. 4.18b).

Notice that Eq. (4.44) also correctly foresees the ascension h to be inversely proportional to the capillary diameter.

Resolving Eq. (4.45) for τ_{LG}, we have

$$\tau_{LG} = \frac{h\rho g a}{2\cos\theta}, \tag{4.46}$$

which can be used for a simple determination of the surface tension, by measuring the capillary ascension h and the contact angle θ and knowing the other quantities by construction.

4.11 Boiling and Condensation

In this section, we consider the boiling and condensation phenomena. In the first case, we have a liquid containing bubbles of its vapor, in the second, a vapor containing droplets of its liquid. In both cases, we must take into account that the saturated vapor pressure under a curved liquid-vapor interface is different from that under a flat surface. The overpressure is given by Eq. (4.42), with the gas being the saturated vapor.

A supersaturated vapor in contact with its liquid on a flat surface immediately condensates. Consider, however, a closed container with only the supersaturated vapor phase inside. The condensation must initiate with the formation of droplets. Very small droplets do indeed form spontaneously by local fluctuations. They are, however, unstable, and soon re-evaporate, because the vapor that is supersaturated relative to a flat surface is not even saturated relative to the curved surface of the droplets, whose radius is very small. To be able to expand, rather than evaporate, the radius of the droplet must be larger than the so-called *critical radius* . This is the radius relative to which the vapor, under the given conditions of temperature and pressure, is saturated. When this happens, the drop starts growing and growing, and the condensation proceeds. The drop has acted as a *condensation nucleus* . If the vapor is extremely clean, the condensation nuclei form only for random thermal fluctuations. These are very rare, all the more so the larger the droplets. The boiling does not start immediately.

If the vapor is not very clean, and, for example, contains thin powder grains, these can act as condensation nuclei. The small drops that spontaneously form, and that are too small to imitate the condensation alone, now wet the grains, forming a liquid layer on their surface. The curvature radius of the so-formed liquid surfaces may be larger than the critical radius, and the condensation proceeds. For example, the fog one sees over a boiling water pot under certain conditions is formed by the supersaturated vapor condensing on tiny powder particles. A similar condition occurs, if the surfaces of the container are not very smooth, at the micrometric scale. The under-critical droplets form a liquid film on the small asperities, triggering the condensation.

Ions are another class of condensation nuclei. When present, they attract the vapor molecules, strongly facilitating the formation of large enough droplets. This phenomenon was discovered by Charles Thomas Rees Wilson (Scotland, 1869–1959), and was used to develop the *cloud chamber* . The chamber makes the tracks of charged particles, like cosmic rays and those originated by radioactive decays, visible. The cloud chamber consists of a vessel containing a vapor, with a glass window to see inside and a piston to expand the volume when needed. The pressure is initially higher than that of the saturated vapor. When a charged particle goes through, it produces a trail of ions in its path. If we now expand the chamber, namely reduce the pressure moving the piston, the vapor becomes supersaturated and the ions act as condensation nuclei. The “track” of the particle is materialized in a trail of droplets. After a fraction of a second, the droplets are big enough. We shoot a flash of light and take a picture. The cloud chambers have made vital contributions to cosmic ray and particle physics.

Consider now a liquid at a certain pressure. If we increase its temperature, when it reaches the value at which the pressure of the liquid is equal to the saturated vapor pressure, the liquid boils. However, if the liquid is very pure and the walls of the container extremely smooth, it will not begin to boil. The fluid is in the state of *superheated liquid* . Again, in this case, the beginning of the phase transition requires the presence of *boiling nuclei* . In the liquid, small vapor bubbles spontaneously form due to fluctuations. The pressure is less than that of the saturated vapor above a flat interface surface but not above the concave interface of the small bubble. Consequently, the vapor in the bubble re-condenses, and the bubble disappears. Additionally, the boiling can now start only if bubbles with a radius larger than a critical value form. These will grow, because the liquid evaporates inside them. However, the process is rare. Impurities like powders and, more often, dissolved extraneous gases that form large enough bubbles trigger the boiling.

In this case as well, ions work as boiling nuclei. The energetic charged particles tracks can be detected by devices similar to the cloud chamber, which are *bubble chambers* . These instruments have also made fundamental contributions to particle physics.

Problems

4.1.
How would the pressure of a fluid change if the intermolecular forces were suddenly to disappear?

4.2. An airtight vessel of one-liter volume is completely full of water at 27 °C. How would the pressure change if the intermolecular forces were suddenly to disappear?

4.3. A certain quantity of nitrogen is enclosed in a container with rigid walls of volume V. Its temperature is $T = 173$ K. Suppose the gas follows the van der Waals equation. Find its pressure p and compare it with the pressure p_{id} the gas would have if ideal in the following cases: (a) $V = 1$ L, $n = 1$ mol, (b) $V = 0.1$ L, $n = 1$ mol, (c) $V = 1$ L, $n = 2$ mol, (d) $V = 0.1$ L, $n = 2$ mol. The molar values of the van der Waals constant for nitrogen are: $a_{mol} = 0.135$ Pa m^6 mol^{-2} and $b_{mol} = 3.9 \times 10^{-5}$ m^3 mol^{-1}.

4.4. A van der Waals gas (n moles) expands at constant temperature from the volume V_i to the volume V_f. Find the expression of the work done, if the constants are a and b.

4.5. A mole of nitrogen adiabatically expands in a vacuum (Joule free expansion). Its volume passes from $V_i = 10^{-3}$ m^3 to $V_f = 10^{-2}$ m^3. How much does the temperature vary? Suppose that nitrogen follows the van der Waals equation with $a_{\mathrm{mol}} = 0.135$ Pa m^6 mol^{-2}.

4.6. We want to expand, at constant temperature, 2 mol of hydrogen in a vacuum from $V_i = 10^{-3}$ m^3 to $V_f = 10^{-2}$ m^3. How much heat should be supplied to the system? Suppose that hydrogen follows the van der Waals equation with $a_{\mathrm{mol}} = 0.024$ Pa m^6 mol^{-2}.

4.7. A mole of oxygen that has the initial volume $V_i = 0.5 \times 10^{-3}$ m^3 expands at the constant temperature of –100 °C to $V_f = 10^{-2}$ m^3. Suppose that oxygen follows the van der Waals equation with $a_{\mathrm{mol}} = 0.136$ Pa m^6 mol^{-2} and $b_{mol} = 3.2 \times 10^{-5}$ m^3 mol^{-1}. Calculate and compare with the case of an ideal gas: (a) the variation of internal energy ΔU, (b) the work W, (c) the exchanged heat Q, (d) the entropy variation ΔS.

4.8. What are the characteristics of the substances, the solid and vapor phases of which can be in equilibrium at the atmospheric pressure?

4.9. Find the specific volume of liquid water and of its vapor, considered as an ideal gas, at STP.

4.10. The mercury-saturated vapor pressure at the temperatures $T_1 = 373$ K and $T_2 = 393$ K are, respectively, $p_1 = 36$ Pa and $p_2 = 99$ Pa. What is the molar latent heat of vaporization in this temperature interval?

4.11. The boiling temperature of benzene (C_6H_6) at atmospheric pressure is $T_1 = 80.2$ °C. Find the pressure p_1 of its saturated vapor at $T_1 = 75.6$ °C, knowing that the latent vaporization heat in this interval is $Q = 400$ kJ/kg.

4.12. Consider, in the plane V,p, the horizontal segment of the isothermal curve at the temperature T of a certain fluid, corresponding to the equilibrium between liquid and vapor. Beyond the temperature, we know the saturated vapor pressure p_s, the mass of the substance m, the specific volumes V_l and V_v of the liquid and vapor and the vaporization heat Q_v. Call 1 and 2 the states at the extremes of the line (all liquid and all vapor, respectively). Express for the transition from 1 to 2: (a) the work W_{12}, (b) the heat Q_{12}, (c) the internal energy variation $U_2 – U_1$, (d) the entropy variation $S_2 – S_1$, (e) the enthalpy variation $H_2 – H_1$.

4.13. Consider, in the plane V,p, the horizontal segments of the isothermal curves at the temperature T of a certain fluid, corresponding to the equilibrium between liquid and vapor. What does correspond to them in the plane p,T?

4.14. The films of soapy water are similar to rubber films. How do their surface tensions differ?

4.15. On the bottom of a vessel, there are holes of radius $a = 50$ µm. Which is the maximum height h to which we can fill the vessel with water before it begins to pour out of the holes? Water wets the bottom of the vessel. The water surface tension is $\tau = 73$ mN m^{-1}.

4.16. One end of a straight glass capillary tube, of internal radius $a = 0.5$ mm, is vertically immersed in a water bath to the height $h = 2$ cm. What pressure must we apply, blowing from the upper extreme, to blow out an air bubble? The water surface tension is $\tau = 0.073$ N m^{-1}.

4.17. A capillary tube, of internal diameter $d = 0.5$ mm, is vertically immersed in a water bath. The length of the external part of the capillary is 10 cm. How much does the water rise in the capillary? The water completely wets the walls.

4.18. A small cubic box of side $a = 3$ cm and mass $m = 5$ g floats on water.

What is the distance *h* under the water surface of the bottom of the box if
(a) the water completely wets the walls of the cube (contact angle = 0°), or
(b) the water does not wet the walls (contact angle = 180°)?

© Springer International Publishing Switzerland 2016
Alessandro Bettini, A Course in Classical Physics 2—Fluids and Thermodynamics, Undergraduate Lecture Notes in Physics, DOI 10.1007/978-3-319-30686-5_5

5. Microscopic Interpretation of Thermodynamics

Alessandro Bettini[1] ✉

(1) Dipartimento di Fisica e Astronomia, Università di Padova, Padova, Italy

✉ **Alessandro Bettini**
Email: alessandro.bettini@pd.infn.it

Thermodynamics and statistical mechanics give complementary descriptions of the same physical processes, from different points of view, the former macroscopic, the latter microscopic. When observed at nanometric dimensions, matter appears to be composed of an enormous number of molecules. Molecules move according to mechanical laws, but there are so many that it would be impossible to solve the problem of motion for any single one. Thermodynamics deals with this issue using global variables, like volume, pressure, temperature, concentrations, etc. This approach is powerful, allowing us to describe considerably different systems, including those that are not made of molecules, like the electromagnetic radiation in a box.

Statistical mechanics starts from the laws of mechanics to extract the equations governing the mean values of the kinematical quantities and their statistical distributions. A rigorous treatment of statistical physics requires knowledge of mathematical tools that are beyond those available to the reader. It will, however, be possible to understand the phenomena clearly in their essential aspects, if not in their fine details. We shall be able to understand the physical meanings of the thermodynamic variables, pressure, temperature, internal energy, entropy and, in the next chapter, viscosity.

In the history of physics, the biggest steps forward happen when previous, apparently completely separate fields become “unified” in a single theory. The first historical unification is credited to Galilei and Newton, who unified terrestrial and celestial mechanics. The second one, in the second half of the XIX Century, is credited to James Clerc Maxwell (Scotland, 1831–1879) and Ludwig Boltzmann (Austria, 1844–1906), who developed statistical mechanics, leading to a unified description of mechanics on one side and thermodynamics, along with the chemistry that can be considered a part of it, on the other.

As is always the case, they had predecessors. The most important was Daniel Bernoulli (Switzerland, 1700–1782), who laid down the basis for the kinetic model of gases. In 1738, he not only developed the hydrodynamic theorem that we studied in Sect. 1.8, but he also put forward the hypothesis that gases consist of an enormous number of molecules moving in chaotic motions. He advanced the fundamental proposals that the pressure of a gas is due to the impact of molecules and that heat exchanges correspond to variations in the molecules’ kinetic energy.

The velocity distribution function of molecules was determined by James Clerc Maxwell in 1859, using, for the first time, statistical concepts in a physical law. Inspired by Maxwell, Ludwig Boltzmann started his lifelong study of statistical mechanics a few years later. This monumental contribution, composed of papers published over a number of years, eventually reached some 2000 pages. In 1902, Josiah Willard Gibbs (USA, 1839–1903) published a book in which he gave a complete formal treatment of statistical mechanics, both for microscopic and macroscopic systems. His beautiful brand of mathematics falls beyond the limits of this course.

Two years before that, in 1900, Max Planck had employed statistical mechanical methods to interpret the energy spectrum of the black body (electromagnetic radiation in a box). Classical physics proved inadequate for this particular endeavor, and Planck was consequently led to advance the hypothesis that energy is not continuous, but rather quantized in discrete minimum quantities. The development of quantum statistics followed with the work of several scientists in the first decades of the XX Century.

Figure 5.1 shows the life spans of the major contributors to classical statistical mechanics in graph form.

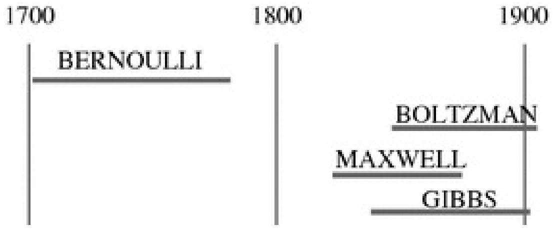


Fig. 5.1 Life spans of the main contributors to classical statistical mechanics

In the first two sections of this chapter, we shall develop the kinetic model of the (ideal) gas and see the physical meaning of pressure and internal energy. We shall control the predictions of the model on the molar heats of the monoatomic and diatomic gases and see how, for the latter (and, more generally, for the polyatomic varieties), classical mechanics reaches its limits of validity. This fact leads to quantum mechanics, which, however, is beyond the aims of this book. In Sect. 5.3, we shall deal with the molar heats of the elemental solid, meeting a similar limit.

Statistical mechanics does not deal with the means of the quantities alone, but also with the distributions of the probabilities for their different values. We shall study such distributions of kinetic energy and velocity in Sects. 5.4–5.7. Following that, we shall consider the always necessary experimental control of the laws of probability distributions.

The probability distribution of the different forms of energy follows, in a first approximation, the universal law according to Boltzmann. The corresponding Boltzmann factor is capable of describing, broadly but in their essential physics, a large number of phenomena. As such, its importance cannot be over-evaluated. We shall discuss a few examples in Sect. 5.9.

In the final two sections, we shall demonstrate physical reasons why microscopic phenomena are reversible, while macroscopic are not. We shall thus understand the physical meanings of entropy and the second law of thermodynamics.

5.1 Kinetic Model of Ideal Gas

In this section, as a first step towards the unification of thermodynamics with mechanics, we shall demonstrate the state equation of the simplest thermodynamic system, the ideal gas, starting from the Newton equation.

We start by defining a *molecular model* of the monoatomic ideal gas, assuming the following hypotheses:

1. The gas is made of molecules, which can be considered point-like. This property is well satisfied for the monoatomic gases, but not for the polyatomic ones. In addition, the density, hence the pressure, should be low enough to guarantee that the average distances between molecules is much larger than their action radius. Consider on purpose that the total volume of the molecules is substantially the volume of its liquid, the covolume.

2. Molecules do not interact except when they collide. The collision time is very short compared to the average time between one collision and the next. In the real gases, molecules interact with the van der Waals force , which, however, can be neglected in a first approximation, all the better the more rarefied the gas, namely the lower its pressure.

3. Collisions between molecules are elastic. This property is also well satisfied by the real gases (if the temperature is not extremely high).

4. The motion of the molecules is completely random; there is no privileged direction. This property is also well satisfied by the real gases.

We assume the gas to be enclosed in a container with the form of a rectangular parallelepiped, with rigid and perfectly plane walls. We choose a reference frame having its origin in a vertex and axes on the three perpendicular edges joining at the vertex, as in Fig. 5.2. We assume the collisions of a molecule with the walls to be elastic too. Consequently, the velocities of a molecule before and after a collision with one of the walls are equal in magnitude, and their directions form equal and opposite angles with that normal to the wall. In other words, the component of velocity normal to the wall inverts in the collision, while the parallel component does not vary.

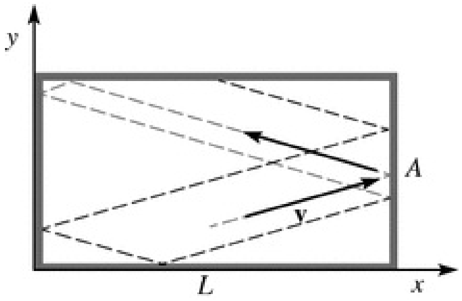


Fig. 5.2 The motion of a molecule in a container

Note that the last hypotheses we introduced, on the elasticity of the molecule-wall collisions and the perfect planarity of the walls, are not included in the molecular model of the ideal gas. In practice, the walls always have some roughness much larger than the molecular dimensions. However, the wall properties cannot have any influence on the state equation. As a matter of fact, these supplementary hypotheses are not logically necessary; we have assumed them to simplify the demonstration.

Let us start with the physical meaning of pressure. In the collision of a molecule with a wall, the normal component of its momentum changes sign. The change of momentum is equal to the impulse given to the wall. The impulse is the integral of a force over the time of the collision. It is very small, but the number of molecules is enormous. The number of collisions is very large in any area, even if very small compared to the macroscopic dimensions, and in every time interval, even if very small compared to the macroscopic times. As a consequence, the effect does not appear macroscopically as a sequence of small impulses, bong, bong, bong,… but as constant in time and uniform at all the points along the walls.

We now mathematically express what we have stated in words. Consider the two walls normal to the x-axis, one at $x = 0$, one at $x = L$ (the length of the relevant side). Let us indicate its area with A. Let m_i be the mass of a molecule and $\mathbf{v}$ its velocity just before a collision with the wall at $x = L$ (see Fig. 5.2). In the collision, the x component of the velocity inverts, while the other two components do not change. The impulse given to the wall is then $2m_i\upsilon_x$. The force exerted by that molecule on that wall is the impulse delivered in the unit time. We must consider that the molecule hits that wall many times in a second. How many? After the first collision, the molecule moves away, hits some other wall and then, after a while, comes back. We can find the time interval recalling the principle of independence of motions. We can forget the components of motion in the y and z directions and think of the motion in the x direction only with velocity υ_x. To go back and forth, it takes $2L/\upsilon_x$ seconds. The number of collisions per second is then $\upsilon_x/2L$ and the impulse delivered per unit time is $2m_i\upsilon_x^2/2L$. If the total number of molecules is N, the force in the x direction due to their entire contribution is

$$F_x = \frac{1}{L}\sum_{i=1}^{N} m_i v_{xi}^2. \tag{5.1}$$

The sum represents all the gas molecules. Note that, in general, the gas might be a mixture of gases with different molecules. This is why we did not take the mass out of the sum. We now observe that the addends are, by a factor of ½ a part, the kinetic energies of the molecules. Then, the sum in the right-hand side of Eq. (5.1) is just the average of the summed quantity times the number of molecules, namely

$$\left\langle mv_x^2 \right\rangle \equiv \frac{1}{N}\sum_{i=1}^{N} m_i v_{xi}^2. \tag{5.2}$$

On the other hand, the motion being completely random, we have $\left\langle mv_x^2 \right\rangle = \left\langle mv_y^2 \right\rangle = \left\langle mv_z^2 \right\rangle$, and, being $v^2 = v_x^2 + v_y^2 + v_z^2$, it is also

$$\left\langle mv_x^2 \right\rangle = \frac{1}{3}\left\langle mv^2 \right\rangle$$

and then, calling $\langle U_K \rangle = \frac{1}{2}\left\langle mv^2 \right\rangle$, the average kinetic energy of the molecules is

$$F_x = \frac{1}{L}\frac{N}{3}\left\langle mv^2 \right\rangle = \frac{1}{L}\frac{2}{3}N\langle U_K \rangle. \tag{5.3}$$

Here, we notice that we considered our molecule going back and forth as if it were alone. This is not so, and collisions with other molecules do change its velocity. However, this is statistically irrelevant, because when a molecule changes velocity, there is always, on average, another one that takes that velocity, and, so to speak, takes its place. We now continue the demonstration. To find the pressure, we just have to divide the force we found by the area on which it is exerted, which is *A*. Taking into account that the volume of the gas is *V* = *LA*, we write

$$p = \frac{F_x}{A} = \frac{1}{AL}\frac{2}{3}N\langle U_K \rangle = \frac{1}{V}\frac{2}{3}N\langle U_K \rangle.$$

The number *N* of molecules is the number of moles *n* times the Avogadro number, and we finally have the equation

$$pV = \frac{2}{3}nN_A\langle U_K \rangle. \tag{5.4}$$

This is the state equation of the ideal gas of our model, obtained from the Newton mechanics with statistical arguments, namely with statistical mechanics . The equation we found says that the product of pressure times the volume is proportional to the average kinetic energy of the molecules. In thermodynamics, this product is proportional to the absolute temperature, namely

$$pV = nRT. \tag{5.5}$$

We must conclude that the absolute temperature is proportional to the average kinetic energy of the molecules. We might even go further and state that it *is* the average kinetic energy. It would be enough to change the measurement unit from kelvin to joule. This, however, is not practical. Keeping the SI units, we write

$$\langle U_K \rangle = \frac{3}{2}\frac{R}{N_A}T = \frac{3}{2}k_B T \tag{5.6}$$

where, on the right-hand side dove, we have introduced the *Boltzmann constant* k_B , after Ludwig Boltzmann (Austria, 1844–1906), one of the main creators, with Maxwell, of statistical mechanics. This is a fundamental constant of physics; it is universal because the gas constant and the Avogadro number are such. Its value is

$$k_B = \frac{R}{N_A} = 1.38 \times 10^{-23}\ \mathrm{J\,K^{-1}}. \tag{5.7}$$

The result we found is valid, in particular, for a monoatomic gas. In this case, the kinetic energy of its molecules, which we considered to be point-like, is the center of mass kinetic energy. Indeed, there cannot be any kinetic energy relative to the center of mass for a point. In a polyatomic gas, the structure of the molecules must be considered and the average kinetic energy relative to the center of mass must be added to $\langle mv^2 \rangle /2$. We shall come back to this point later in the chapter.

We will now discuss the result we found. First, we notice that we have given a clear mechanical meaning to a purely thermodynamic quantity, such as the absolute temperature. Notice that only the absolute scale, the scale based on the gas thermometer, has a clear physical meaning, not the Celsius, the Fahrenheit or others. This is because a physical temperature of zero exists. One might think this would happen when the motion of molecules ceases. However, things are not so simple. Indeed, when one gets very close to absolute zero, classical mechanics, which we have used, no longer correctly describes nature. Quantum mechanical phenomena appear. In particular, at absolute zero, a well-defined non-zero kinetic energy remains, called *zero-point energy* . This is a purely quantum phenomenon, not understandable through classical physics. However, the zero-point energy is extremely small and completely negligible at the usual temperatures. This is true because the motions we have considered take place on very large geometrical scales compared to the atomic ones. Under these conditions, classical mechanics works well down to very low temperatures. But this is not the case when the motions are on molecular scales, like the vibrations of molecules. We shall come back to that in a later portion of the chapter.

We still should check whether temperature and average molecule kinetic energy have the same properties. Otherwise, we cannot identify them and the theory loses meaning. While this is certainly true, the demonstration is not simple and we shall not go into it here. Rather, we shall look at an interesting consequence. In a mixture of gases, all the components have the same temperature and, consequently, the different molecules have the same average kinetic energy.

Let us consider a mixture of monoatomic ideal gases, respectively, with N_1 molecules of the first gas, N_2 of the second, and so on. The total number of molecules is $N = N_1 + N_2 + \ldots$Recalling that we have shown Eq. (5.4) without assuming all molecules to be equal, that result is still valid and we can write

$$pV = N_1 k_B T + N_2 k_B T + N_3 k_B T + \cdots \tag{5.8}$$

This tells us that the different gases are independent from one another. This is a consequence of the assumption that their molecules do not interact outside collisions. Defining the partial pressure of the *i*th p_i gas as the pressure it would have if it were taking up the entire volume of the container by itself, we reach the conclusion that

$$p = p_1 + p_2 + p_3 + \cdots \tag{5.9}$$

which is the Dalton law : the total pressure of a (ideal) gas is the sum of the partial pressures of its components. We have seen that this is also a consequence of the laws of mechanics.

We now summarize the physics contained in the ideal gas equation. The pressure of a gas in a container is due to the collisions of its molecules with the surrounding walls. The pressure is proportional to the average square velocity. One of the powers comes from the impulse delivered in a collision, the second from the rate at which the molecule collides. The average square velocity is proportional to the average kinetic energy and the latter is proportional to the absolute temperature.

5.2 Meaning of the Internal Energy. Specific Heats of Gases

The internal energy state function of a thermodynamic system is the sum of the kinetic and potential energies of the constituent molecules, namely their mechanical energies. The kinetic energy we are talking about is the internal energy , namely the energy in a frame in which the system is at rest. For example, the internal kinetic energy of a gas in a bottle is the same whether the bottle is at rest or on a train moving at 100 km/h, or if it is rotating on a merry-go-round. Indeed, the temperature of the gas is the same in these different situations. The potential energy to be considered is also the internal one. It is the same whether the bottle is at sea level or on the top of a mountain. It is the potential energy of the van der Waals forces between molecules. The potential energy is zero for and only for the ideal gas. The first law of thermodynamics states that the internal energy is conserved if the system does not exchange heat or work. This implies that the intermolecular forces are conservative . As a matter of fact, at the microscopic level, dissipative forces do not exist.

The van der Waals force is attractive at the intermolecular distances that are characteristic of gases. Consequently, the internal potential energy is negative. If the gas expands, the average distance between molecules increases and the internal energy decreases in absolute value, remaining negative. We observe that, under the usual conditions, the variation of internal energy of gases with varying volume is modest, but it is very important with varying temperature. We must conclude that the kinetic energy is much larger than the potential energy. As a matter of fact, it is found that Eq. (5.6) also holds for the monoatomic real gases, not only for the ideal ones.

Notice that the factor 3 on the right-hand side is just the number of mechanical degrees of freedom of a point object, such as a monatomic molecule. The number of degrees of freedom of a body is the number of parameters we must know to know the mechanical state of the body. Indeed, when we know the three coordinates, we know the mechanical state of the point. We conclude that, in a monoatomic gas, the average internal kinetic energy is $k_B T/2$ per degree of freedom. This conclusion is very important because a theorem of statistical mechanics states its general validity. The *energy equipartition* theorem states that, in a system in thermodynamic equilibrium, the average kinetic energy is “equiparted”, equally divided, amongst the degrees of freedom of the molecules. If, for example, the molecule is diatomic, we must add a $k_B T/2$ contribution for each independent rotation (two) and vibration (one) degrees of freedom.

The measurable quantities that are more directly related to the internal energy are the specific or molar heats . We shall see in this section and the next the predictions by statistical mechanics for the molar heats of the polyatomic gases and of the solid, and we will compare them with the measured values.

Let us start with the measurement of the gas molar heats , or, better yet, because it is easier, of their ratio $\gamma = C_p/C_V$. Nicolas Clément (France 1779–1842) and Charles Desormes (France; 1771–1862) measured the heat ratio with the simple experiment shown in Fig. 5.3 (Clément-Desormes experiment) in

1819. The gas is initially in the S container. We transfer some gas in the spherical bottle to the pressure p_0, which is smaller than the pressure in S. Once the equilibrium is reached, we measure p_0 with the manometer M. The volume of the gas is the volume of the bottle, which we know. Let us call it V_0. The temperature is the (known) ambient temperature T_0. This is our thermodynamic system.

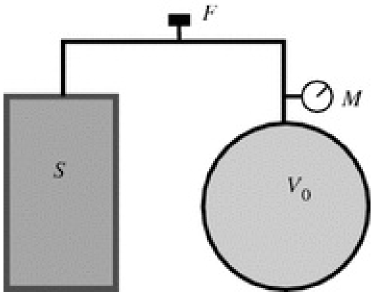


Fig. 5.3 Sketch of the Clément-Desormes experiment

We now open the valve F for a short time, admitting some more gas into the bottle. We can consider that the gas originally present, our system, now takes only a fraction, say V, of the total volume. Its temperature changes to a new value too, say T. We do not know V or T, but, as we shall see, we do not need them. We immediately read the new pressure, say p_1. The process from the initial state (p_0, V_0, T_0) to the final one (p_1, V, T) has been very fast and the heat exchange has been negligible. We can use the adiabatic equation:

$$p_0 V_0^\gamma = p_1 V^\gamma. \tag{5.10}$$

The temperature of the gas in the bottle now gradually changes to reach equilibrium at T_0. The volume of the originally present gas, our system, remains V, because neither the volume of the bottle nor the fraction it occupies vary. The pressure does vary, reaching the new value p_2, which we measure. This third state is (p_2, V, T_0). It is on the same isothermal as the initial state, and we can write

$$p_0 V_0 = p_2 V. \tag{5.11}$$

We raise both sides of this equation to the power γ and divide by those of Eq. (5.10), obtaining

$$\left(\frac{p_2}{p_0}\right)^\gamma = \frac{p_1}{p_0}. \tag{5.12}$$

Solving for γ, we have

$$\gamma = \left(\frac{p_1}{p_0}\right) / \ln\left(\frac{p_2}{p_0}\right). \tag{5.13}$$

Let us now see what statistical mechanics foresees. Let us start with the ideal monoatomic gases . The noble gases at STP behave almost as such. The molecules have three degrees of freedom. Hence, the average kinetic energy is $\langle U_K \rangle = (3/2)k_B T$. There is no potential energy between molecules (the gas is ideal) or inside the molecules (that we consider point-like). The internal energy of a mole of gas is then

$$U = \frac{3}{2}N_A k_B T = \frac{3}{2}RT. \tag{5.14}$$

The molar heats, taking into account the Mayer's relation , Eq. (2.49), are foreseen to be

$$C_V = \frac{3}{2}R, \quad C_p = \frac{5}{2}R. \tag{5.15}$$

We finally foresee that

$$\gamma = \frac{5}{3} = 1.666. \tag{5.16}$$

Experiments find very similar values, for example, 1.668 for Ar and 1.666 for He.

Consider now a diatomic gas (oxygen or hydrogen, for example). We can think of the molecule as being made of two point-like atoms of masses m_1 and m_2 (they might be different, as, for example, in the CO). The two atoms interact with a force that, in a good approximation, we can think of being elastic, with an elastic constant that we call κ. This is a harmonic oscillator with proper angular frequency (see Sect. 6.2 of the 1st volume),

$$\omega = \sqrt{\kappa/\mu} \tag{5.17}$$

where μ is the reduced mass. This is the oscillation angular frequency of the molecule.

We need six quantities to define the mechanical state of the molecule. It has six degrees of freedom. Three degrees of freedom characterize the translator motion, namely the motion of the center of mass. Two degrees of freedom correspond to the rotations about the central axes. We do not count the rotation about the axis joining the atoms, because there is no kinetic energy associated with the moment of inertia about this axis. The last degree of freedom, the distance between atoms, is internal to the molecule and corresponds to its oscillation at the angular frequency we found. For the equipartition theorem, the average kinetic energy is $\langle U_K \rangle = (6/2)k_B T$.

We should still add the average potential energy of the harmonic oscillation. We know from mechanics (see Sect. 3.2 of the 1st volume) that the average for a period of the potential and kinetic energy are equal in a harmonic oscillator. We must be careful, however. Those are *time* averages while we now need *statistical* averages , namely averages for all the elements of the system at a certain instant. For the systems in thermodynamic equilibrium we are considering, all the statistical averages are independent of time. Fortunately, for our system, the two, conceptually different averages have the same value. This is, indeed, the case for the majority of statistical ensembles, but not for all of them. We can understand that the statement is likely by thinking about taking a rapid sequence of ideal shots of a certain molecule. We shall "see" the atoms, sometimes closer, sometimes farther away, sometimes with one orientation, sometimes with another. The set of these photos cannot be different from the set of photos of many molecules at the same instant. As a matter of fact, all the molecules of a given species are identical. We can conclude that the statistical average is equal to the time average and that the average potential energy of the harmonic oscillators is equal to their average kinetic energy, namely $k_B T/2$. Finally, statistical mechanics foresees the internal energy of a mole of diatomic gas to be

$$U = \frac{7}{2}N_A k_B T = \frac{7}{2}RT.$$

For the molar heats, it foresees

$$C_V = \frac{7}{2}R, \quad C_P = \frac{9}{2}R \Rightarrow \gamma = \frac{9}{7} = 1.286. \tag{5.18}$$

The experimental values at STP for common gases like O_2, H_2, N_2, are around $\gamma = 1.40$, in clear disagreement with statistical mechanics. For gases with more massive molecules, like iodine I_2 or bromine Br_2, the experimental values are around $\gamma = 1.30$, close to the theoretical one. This is true at the ambient temperature. Experiments show that both molar heats and their ratios vary with temperature. At high enough temperatures, the γ's of all the diatomic gases tend toward the statistical mechanics value of 9/7.

The following attempt to "save" (classic) statistical mechanics has been done. We observe that the measured value at STP $\gamma = 1.40$ is just 7/5, which is the value foreseen for 5 degrees of freedom. It looks like two degrees of freedom would be missing, likely the internal to the molecule ones (the vibration). In this case, the internal energy would be $U = (5/2)RT$. We can think of the molecule, rather than being two material points linked by a force, as being a small rigid stick. Such a stick indeed has five degrees of freedom: the three coordinates of the center of mass and the two angles fixing the direction of the stick. The angle about the stick axis is irrelevant, because we can consider its section negligible. This argument is known as the *freezing out of degrees of freedom* .

This argument (which is often made) is, however, wrong. Indeed, the rigid stick is the limit of the system shown in Fig. 5.4, when the spring constant goes to infinite. But the average energy in the internal to the molecule degree of freedom, the potential plus kinetic vibration energy, is $k_B T$ *however large* the spring constant might be. Consequently, such is its value in the infinite limit, when the molecule becomes a rigid stick.

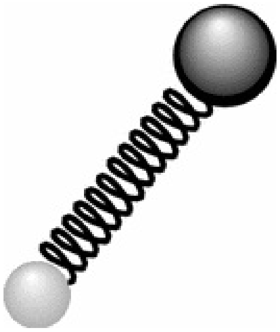


Fig. 5.4 The scheme of a diatomic molecule

In 1859, James Clerc Maxwell (UK; 1831–1879) published a fundamental article, the foundation of statistical mechanics. It was the unification of previously separated fields of knowledge, mechanics, thermodynamics and chemistry. At the end, he summarized the extraordinary successes of the theory. He had explained many known relations, such as the gas laws, the diffusion processes, and the viscosity of gases. We shall discuss the latter in the next chapter. However, he honestly concluded:

> Finally, by establishing a necessary relation between the motions of translation and rotation [*he is talking about the equipartition theorem*] of all particles not spherical, we proved that a system of such particles could not possibly satisfy the known relation between the two specific heats.

For the first time in history, classical mechanics had to face a problem that was going to determine its limits of validity. The problem of the specific heats would remain open and worry the most brilliant scientists for the entire second half of the XIX century. The next difficulty was discovered in the last decades of the century. It was again in statistical mechanics, this time in the mechanics of the photon gas. It was the problem of the black body radiation . It was solved by Max Planck (Germany, 1858–1947) in 1900, opening the way for a very deep scientific revolution, quantum mechanics.

We now briefly consider the polyatomic molecule gases. Take, for example, ethane (C_2H_6). Having 8 atoms, the molecule has 3 × 8 = 24 degrees of freedom, namely $12k_BT$ alone of kinetic energy. The internal to the molecule potential energy should be added. It is clear that γ must be very close to 1. However, the measured value is 1.22. Once more, the theory fails.

We notice that quantum effects are relevant for the motion inside the molecules, but not for the motion of the molecule center of mass in the container. This is an example of what we stated in the introduction; quantum effects, which in any case increase with decreasing temperature, are more important the more the motion is limited. The center of mass is essentially free and quantum effects on its motion appear only at extremely low temperatures. At the end of Sect. 5.9, we shall give some hints as to how quantum mechanics explains the observed values of the specific heats.

5.3 Specific Heats of Solids

The equipartition theorem holds for every system, including the condensate bodies. Even for them, the average kinetic energy is $k_BT/2$ per degree of freedom. The calculation of the contribution of the potential energy is, in general, very difficult. It is rather simple for the *elemental solids* , namely for the crystals having only one type of atom, which we will now discuss. The measured quantity is the specific heat. Generally, it is measured at constant (atmospheric) pressure. Note, however, that the differences between constant pressure and constant volume heats are very small for solids (for example, for Fe, it is $\gamma = c_p/c_V = 1.02$). This is because the dilatation coefficients of solids are small.

Table 5.1 gives, in the second column, the specific heats of several elemental solids at room temperature. They are very different from one another. If we, however, compare *molar heats* , reported in the 4th column, we see that they are rather similar. In a round figure, we can write

$$c = 25\ \mathrm{J\ K^{-1}} \cong 3R. \tag{5.19}$$

Table 5.1 Specific and molar heats and molar masses for crystals of several elements

Element	**Specific heat (kJ kg^{-1} K^{-1})**	**Molar mass (g mol^{-1})**	**Molar heat (kJ mol^{-1} K^{-1})**
Carbonium	0.50	12	6.0

Aluminum	0.89	27	24.1
Copper	0.38	63.4	24.4
Iron	0.44	55.8	25.0
Zinc	0.39	65.4	25.3
Tungsten	0.134	183.9	24.6
Lead	0.13	197	25.2

This property was discovered in 1819 by Pierre Loius Dulong (France, 1785–1838) and Alexis Thérèse Petit (France, 1791–1820) and is known as the *Dulong-Petit rule* . They also discovered the important exception of the diamond, whose molar heat is about one fourth of the others. The statistical mechanical interpretation was given by Ludwig Boltzmann in 1871. We can consider a crystal as an array of point-like atoms that oscillate about stable equilibrium positions, as we described it in Sect. 4.1. Each atom can be considered a harmonic oscillator, which can oscillate in three independent directions. For each of them, the average kinetic energy is $k_B T/2$. In addition, we must include the potential energy.

As we did for the gases, we can identify temporal averages with statistical averages. The average potential energy of each oscillation direction is equal to the average kinetic energy, namely $k_B T/2$. In conclusion, the total average mechanical energy, kinetic plus potential, of the oscillators is $6k_B T/2$. The internal energy of a mole of solid is then

$$U = N_A 3k_B T = 3RT \tag{5.20}$$

and the molar heat is

$$c = \frac{dU}{dT} = 3R \tag{5.21}$$

which is the Doulong Petit rule. We have seen that, based on the Newton laws, particularly the energy equipartition, this rule is what statistical mechanics foresees. However, at an ambient temperature, the rule is not valid for diamond. This problem worsened when experiments at low temperatures by James Dewar (UK, 1842–1923) and Walther Nernst (Germany, 1864–1941) showed that molar heats are not constant. They increase with temperature. The Dulong-Petit rule holds for all elemental crystals, including diamond, if the temperature is high enough.

When temperature decreases, the molar heats' decrease is initially slow. However, below a certain temperature, characteristic of the solid, it becomes much faster. This is the *Debye temperature* θ_D. Plotting the molar heats of different elemental solids versus the ratio of the temperature and the Debye temperature (T/θ_D), one finds an almost universal curve, as shown in Fig. 5.5. The correct, quantum explanation of the phenomenon was formulated by Peter Debye (The Netherlands, 1884–1966) in 1912. The theory foresees, in particular, in perfect accord with the data, that the specific heats tend towards zero proportionally to the third power of temperature:

$$c \propto (T/\theta_D)^3 . \tag{5.22}$$

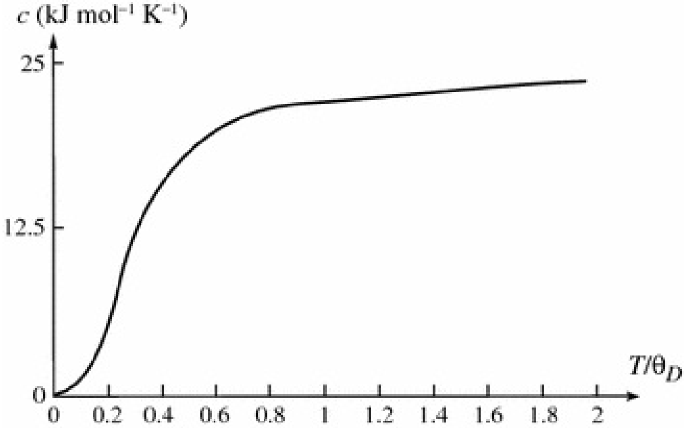


Fig. 5.5 Molar heat of elemental crystals versus absolute temperature normalized to the Debye temperature

Roughly speaking, the Debye temperature is the temperature below which classical mechanics fails. To give some examples, values of θ_D are 2230 K for diamond (atomic mass of C is $A = 12$), 428 K for Al ($A = 27$), 470 K for Fe ($A = 55.8$), 327 K for Zn ($A = 65.4$), 164 K for Au ($A = 197$) and 105 K for Pb ($A = 207$). One sees that the Debye temperature tends to decrease with an increasing atomic number. In particular, the Debye temperature of Carbonium is much higher than room temperature. For such light elements, quantum effects are already important at room temperature.

Classical mechanics failure happens for solids at temperatures much higher than those for the translatory motions in gases. This is because the oscillatory motions in a crystal are confined within much smaller dimensions. We also see that the larger the atom mass, the lower the temperature at which classical mechanics fails.

At temperatures close to absolute zero, all substances are in a condensed phase; no gas exists. Nernst postulated in 1905 the *Nernst heat theorem* . This can be formulated by stating that, at low enough temperatures, the thermodynamic state functions of condensed bodies tend to be independent of temperature. The derivatives relative to temperature of, in particular, internal energy and enthalpy, which are the molar heats C_p and C_V, go to zero. We have already noticed this behavior.

5.4 Distribution Functions

Up to now, we have considered the mean values of mechanical quantities, such as kinetic energy, over all the molecules of a macroscopic system. Obviously, the kinetic energies of molecules are different from one another. They are casually distributed. So are the positions, the velocities, etc. We shall now find the strictly connected distributions of kinetic energies and of velocities of an ideal monoatomic gas. We shall then see how the results can be compared with experiments.

We start by considering how the molecules are distributed in space, independently of their velocity. In general, density might vary with position in the gas. Let $\rho(\mathbf{r})$ be the gas density at the generic point with position vector $\mathbf{r} = (x, y, z)$. All the molecules of our ideal gas are equal, point-like and have mass m. A relevant quantity for a system composed of equal particles is the *number density*, which is the number of particles per unit volume. We shall indicate this with n_p. The relation between number and mass densities is simply

$$n_p(\mathbf{r}) = \rho(\mathbf{r})/m. \tag{5.23}$$

The number dn_p of molecules in the elementary volume between x and $x + dx$, y and $y + dy$ and z and $z + dz$ is

$$dn_p = n_p(\mathbf{r})\,dxdydz. \tag{5.24}$$

The function $n_p(\mathbf{r})$ informs us as to how the molecules are distributed in space, or, in other words, where it is more probable and where less probable for us to find molecules. Note that the question as to know *exactly* how many molecules are in a given position is meaningless, because the position is a continuous variable with infinite values, while the number of molecules is enormous, but finite. The medium is often homogeneous, for example, a liquid in a container. Then, $n_p(\mathbf{r})$ is independent of $\mathbf{r}$. But this is not always so. Consider, for example, the atmosphere. Its density becomes smaller and smaller with increasing height, due to the weight. We shall study that in the next section.

We now consider the distribution of the magnitude of the velocity. Again, the question of exactly how many molecules have that velocity is meaningless (How many cars pass at a given kilometer of a freeway in a day traveling at exactly 100,000,000… km/h?). The more meaningful question to ask is how many cars have velocities in a small interval around the value we are interested in, for example, between 199 and 101 km/h.

Let N be the number of molecules per unit volume and ΔN the number with velocity in magnitude between υ and $\upsilon + \Delta\upsilon$. We define the function

$$n_p(v) = \lim_{\Delta v \to 0} \frac{\Delta n_p}{\Delta v} = \frac{dn_p}{dv}. \tag{5.25}$$

To understand the meaning of this equation, suppose that the function $n(\upsilon)$ has, for example, the behavior shown in Fig. 5.6. The shaded area represents the number dN of molecules per unit volume with velocity between υ and $\upsilon + d\upsilon$. From Eq. (5.25), we have, obviously,

$$dn_p = n_p(v)\,dv, \tag{5.26}$$

which, geometrically, is the shaded area in the figure (consider that the ordinate, N, varies only by infinitesimals inside the interval $d\upsilon$ that is itself infinitesimal).

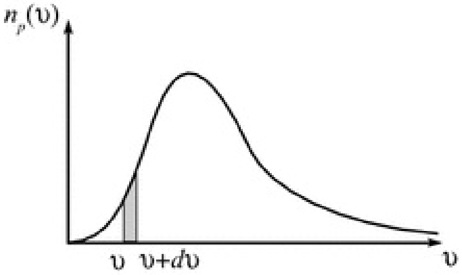


Fig. 5.6 A velocity distribution function

If we know $n_p(\upsilon)$, we know several details of the system. For example, if we want the number of molecules (per unit volume) with velocities between υ_1 and υ_2, we just have to calculate the integral $\int_{v_1}^{v_2} n_p(v)dv$, if we want the number of molecules with velocity larger than υ_1, we calculate the integral $\int_{v_1}^{\infty} n_p(v)dv$, and so on. Notice, in particular, that the total number of molecules per unit volume, namely with whatever velocity, is

$$N = \int_0^\infty n_p(v)dv. \tag{5.27}$$

$n_p(\upsilon)$ being a number per unit volume and unit velocity interval, its physical dimensions are $[n_p] = \frac{1}{L^3(L/T)} = L^{-4}T$.

The *velocity distribution function* is defined as $n_p(\upsilon)$ divided by the number of molecules per unit volume, namely

$$f(v) = n_p(v)/N. \tag{5.28}$$

Clearly, it is

$$\int_0^\infty f(v)dv = 1. \tag{5.29}$$

We say that the function is *normalized*, meaning that its integral over its entire domain is one. The quantity *f*(*υ*) *dυ* has two meanings that are strictly connected with one another. The first meaning is to be the *fraction* of molecules with velocity between *υ* and *υ* + *dυ*. We recall that, by definition, the probability of an event is the ratio between the number of cases favorable to the event and the total number of cases. Then, we ask what the probability is that a molecule chosen at random has velocity between *υ* and *υ* + *dυ*. We immediately see that the probability of this event is just *f*(*υ*)*dυ*. The function *f*(*υ*) becomes a probability when multiplied by *dυ*. For this reason, it is said to be a *probability density* . Its physical dimensions are the reciprocal of velocity.

5.5 The Ideal Gas in a Force Field

In this section, we shall consider a gas in thermal equilibrium in the field of the weight force. The mass and number densities are larger at lower altitudes than at higher. We shall find the distribution of the molecules in elevation, assuming temperature to be the same at all points of the gas. This is not really so in the earth’s atmosphere, where the temperature diminishes with increasing altitude, and in which winds exist. However, up to altitudes of 10 km, the temperature usually decreases by about 6 K/km, or, in relative terms, about 3 %/km, which we can neglect in limited elevation intervals in absence of strong winds.

Let us consider a vertical isothermal column of section *S* of ideal gas under the action of its weight. The system is in thermodynamic equilibrium (in particular, there is no wind). We consider a gas of equal molecules of mass *m*. We take a vertical coordinate axis *z* vertically upward. The gravity acceleration is **g** (directed as −z). The number of molecules per unit volume in the generic layer between *z* and *z* + *dz* is the numeric density at *z*, say *n*(*z*), times the volume of the layer, *S* dz.

The layer is in equilibrium. Consequently, the resultant force acting on it is zero. The forces are: (1) the weight, which is the number of molecules in the layer *n*(*z*)*Sdz*, times the mass of a molecule *m* times gravity acceleration, namely *n*(*z*)*Sm gdz* directed vertically downward. (2) The pressure force on the lower face (elevation z) $Sp(z)$, vertically downward, (3) the pressure force on the upper face (elevation $z + dz$) $S\,p(z + dz)$, vertically downward. At equilibrium: $-S\,p(z+dz) + S\,p(z) - n_p(z)\,S\,mg\,dz = 0$. We immediately get

$$\frac{dp}{dz} = -mg n_p(z)\,. \tag{5.30}$$

The two functions, numeric density $n_p(z)$ and pressure $p(z)$, are linked by the state equation. Indeed, let us consider a generic volume V and indicate with n the number of moles it contains. The number of molecules per unit volume is $n_p(z) = n(z)N_A/V$. The state equation of the ideal gas

$$pV = nRT \tag{5.31}$$

can be then written as

$$p(z) = \frac{n(z)}{V}RT = \frac{n_p(z)}{N_A}RT$$

or

$$p(z) = n_p(z)\,k_BT\,.$$

We substitute this equation in Eq. (5.30), obtaining

$$\frac{dn_p(z)}{dz} = -\frac{mg}{k_BT}n_p(z)\,. \tag{5.32}$$

This differential equation can easily be integrated into separating variables, namely writing it as

$$\frac{dn_p(z)}{n_p(z)} = -\frac{mg}{k_BT}dz, \tag{5.33}$$

which only contains, on the left-hand side, the function $n_p(z)$. Integrating, we have

$$\ln n_p(z) = -\frac{mgz}{k_BT} + \mathrm{const} \tag{5.34}$$

where const is the integration constant, which we do not need to specify now. Finally, taking the exponential of both sides, we get

$$n_p(z) = n_{p0}e^{-\frac{mgz}{k_BT}} \tag{5.35}$$

where n_{p0} is now the integration constant. We immediately see that its meaning is to be the numeric density at z = 0, namely $n_{p0} = n_p(0)$.

Figure 5.7 shows the numeric density distributions in elevation for two gases of different molecular mass, H_2 and O_2, at the same temperature. Both distributions are exponential. With increasing elevation, the gas of larger molecular weight rarefies sooner. The real atmosphere cannot be considered isothermal in such an elevation interval, but the ratio of hydrogen to oxygen density does diminish with increasing elevation.

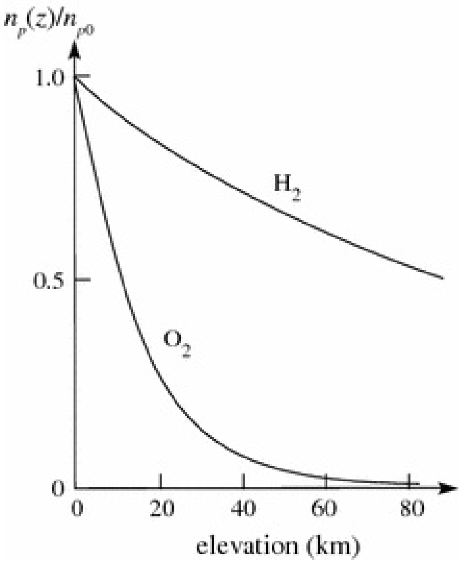


Fig. 5.7 Numeric density relative to sea level of H_2 and O_2 versus elevation in an isothermal atmosphere

We notice that the expression *mgz* in Eq. (5.35) is just the potential energy of a single molecule, in the field force of the weight. The argument we made to reach Eq. (5.35) is valid for any field of forces with potential energy $U_p(z)$. The numerical density is given by

$$n_p(z) = n_{p0} e^{-\frac{U_p(z)}{k_B T}} \tag{5.36}$$

where $n_{p0} = n_p(0)$ is the integration constant. This is a particular case of a general and very important expression, called a *Boltzmann distribution* , of position, in this case. We shall encounter more cases in the subsequent sections.

Number density cannot be easily measured, but the pressure, which is proportional to it, can. Clearly, the elevation dependence on pressure is

$$p(z) = p_0 e^{-\frac{U_p(z)}{k_B T}} \tag{5.37}$$

with $p_0 = p(0)$. We finally observe that $n_p(z)$ is proportional to the probability of finding a molecule at elevation *z*, which we call *f*(*z*), namely the position distribution function . In practice, one is often interested in the ratio of probabilities, namely, in this case, the ratio of the probabilities of finding a molecule at two different elevations , say z_2 and z_1. This is given as

$$\frac{f(z_2)}{f(z_1)} = \frac{e^{-\frac{U_p(z_2)}{k_B T}}}{e^{-\frac{U_p(z_1)}{k_B T}}} = e^{-\frac{U_p(z_2)-U_p(z_1)}{k_B T}} \tag{5.38}$$

which is the *ratio of the Boltzmann factors* or (last term) *e* raised to the power equal to the opposite of the potential energy difference divided by $k_B T$.

As we have already noted, the atmosphere is not isothermal at all. In addition, as is well known, temperature and pressure continuously change over time. However, their average values are assumed to define the so-called standard atmosphere. Table 5.2 reports the values of temperature, pressure and density of the standard atmosphere as functions of elevation.

Table 5.2 Standard atmosphere temperature, pressure and density versus elevation

Elevation (km)	**Temperature (°C)**	**Pressure (hPa)**	**Density (kg m^{-3})**
0	15	1013	1.225
1	8.5	899	1.111
2	2	795	1.007
3	−4.5	701	0.909
4	−11.0	617	0.82
5	−17.5	540	0.737
6	−24.0	472	0.66
7	−30.5	412	0.591
8	−37.0	357	0.527
9	−43.5	308	0.467
10	−50.0	264	0.414

Notice that, in the table, pressures are given in hPa. This unit is commonly used in meteorology, because it is rather close to the torr = 1.33 hPa. The torr is the pressure of a mercury column 1 mm in height and is also called the “mm Hg”. It is an old unit, recalling when pressures were measured with mercury barometers. It is still used, in particular, for blood pressure.

Consider now the general case in which the potential energy of the molecule is a function of the three coordinates, say $U_p(\mathbf{r})$, where **r** is the position vector. The distributions at the *x* and *y* coordinates are obtained with arguments identical to those we have just developed. We can say that the probability of finding a molecule in the infinitesimal volume *dx dy dz* at the position vector **r** is given by

$$f(\mathbf{r})\,dx\,dy\,dz = f_0 e^{-\frac{U_p(\mathbf{r})}{k_B T}}\,dx\,dy\,dz. \tag{5.39}$$

The function $f(\mathbf{r})$ is the probability per unit volume (probability density) of finding a molecule in the given position or, equivalently, with a given potential energy.

5.6 The Boltzmann Law for Kinetic Energy

In the last section, we found the molecules' position distribution function in a force field. We shall now find their velocity distribution function , which we indicate with $f(\mathbf{v})$ or, more explicitly, $f(\upsilon_x, \upsilon_y, \upsilon_z)$. This function is the probability of finding in the unit volume a molecule with an x-component of velocity between υ_x and $\upsilon_x + d\upsilon_x$, a y-component between υ_y and $\upsilon_y + d\upsilon_y$ and a z-component between υ_z and $\upsilon_z + d\upsilon_z$. Similar to the case of the potential energy, such a molecule has a definite kinetic energy , $U_k(\mathbf{v}) = \frac{1}{2}m\left(v_x^2 + v_u^2 + v_z^2\right) = \frac{1}{2}m\mathbf{v}^2$. The relationship between velocity distribution function and kinetic energy is equal to that between position distribution function and potential energy Eq. (5.39). We shall now prove the probability of finding a molecule in the infinitesimal volume in the "velocity space" between υ_x and $\upsilon_x + d\upsilon_x$, υ_y and $\upsilon_y + d\upsilon_y$ and υ_z and $\upsilon_z + d\upsilon_z$

$$f(\mathbf{v})\, dv_x dv_y dv_z = f_0 e^{-\frac{U_k(\mathbf{v})}{k_B T}} dv_x dv_y dv_z \tag{5.40}$$

where, as usual, f_0 is a constant.

Equations (5.39) and (5.40) show that both probabilities are proportional to the Boltzmann factor . Its exponent is, in both cases, the relevant energy, potential or kinetic, divided by the same quantity (with the dimensions of energy).

We continue considering an ideal gas in thermodynamic equilibrium at temperature T. We start by finding the distributions of the vertical component of velocity, υ_z. We once again consider a vertical column of gas of section S under the action of weight. Equation (5.36) gives the position (elevation) distribution $n(z)$. The motion of the molecules is completely disordered. In particular, there is no privileged direction. As is well known, the projections of the motion of a molecule on the axes are independent of one another. We then fix our attention on the motion along z and forget about the other components, which exist but are irrelevant.

We ask how many molecules per unit volume at a given height z have the z component of their velocity between υ_z and $\upsilon_z + d\upsilon_z$. This number divided by the total number per unit volume is $f(\upsilon_z)d\upsilon_z$, where $f(\upsilon_z)$ is the distribution function. An important theorem of statistical mechanics, of which we will not give a demonstration, states that the distribution function $f(\upsilon_z)$ depends only on temperature, and not, in particular, on z. At higher elevations, for example, the density of molecules diminishes, but their velocity distribution does not vary.

We now express the (obvious) fact that the probability of finding a molecule with any velocity is one. In formulae,

$$\int_{-\infty}^{+\infty} f(v_z)dv_z = 1. \tag{5.41}$$

Let us now consider a horizontal plane cutting the column at the height z. Let us fix our attention on the molecules with velocity between υ_z and $\upsilon_z + d\upsilon_z$, with $\upsilon_z > 0$ (going up). How many of these molecules will cross the plane in the time interval between t and $t + dt$? These are the molecules that, at time t, are below the plane at a distance less than or equal to that which they can cross in dt. This distance is $\upsilon_z dt$. The volume containing them is $S\upsilon_z dt$. Their number is, thus, this volume times the number per unit volume at the considered height [$n(z)$] times the fraction of them in the considered velocity interval [$f(\upsilon_z)d\upsilon_z$], namely $S\upsilon_z dt\, n(z)f(\upsilon_z)d\upsilon_z$.

Let us follow our molecules for a while. After some time, they are at a greater height, say z'. Their velocity is smaller, say, v'_z, and the velocity interval is changed to dv'_z. The number of molecules crossing upward through the horizontal plane at z' in dt, according to the argument we just made, is $S v'_z dt n(z') f(v'_z) dv'_z$. But these are just our molecules, because we are at the statistical equilibrium; how many are arrivals and how many are departures, $n_p(z) f(v_z)\, dv_z S v_z dt = n_p(z') f\left(v'_z\right) dv'_z S v'_z dt$. We can simplify and write

$$n_p(z) f(v_z)\, dv_z v_z = n_p(z') f\left(v'_z\right) dv'_z v'_z. \tag{5.42}$$

We now apply the energy conservation principle. Considering that the x and y components of the velocity do not vary, we write

$$\frac{1}{2} m v_z^2 + mgz = \frac{1}{2} m v_z'^2 + mgz'. \tag{5.43}$$

Differentiating both sides with z and z' fixed, we have

$$v_z dv_z = v'_z dv'_z. \tag{5.44}$$

Equation (5.42) becomes

$$n_p(z)\, f(\upsilon_z) = n_p(z')\, f\left(\upsilon'_z\right). \tag{5.45}$$

But we do know $n(z)$, Eq. (5.36), and can write

$$\frac{n_p(z')}{n_p(z)} = \frac{f\left(\upsilon'_z\right)}{f(\upsilon_z)} = e^{-\frac{mg(z'-z)}{k_B T}}. \tag{5.46}$$

But, for the energy conservation, the potential energy difference is equal and opposite to the kinetic energy difference, and we can write

$$\frac{f\left(\upsilon'_z\right)}{f(\upsilon_z)} = e^{-\frac{\frac{1}{2}m\upsilon_z^2 - \frac{1}{2}m\upsilon_z'^2}{k_B T}} \tag{5.47}$$

and finally

$$f\left(\upsilon_z\right) = f_0 e^{-\frac{\frac{1}{2}m\upsilon_z^2}{k_B T}} \tag{5.48}$$

where, as usual, f_0 is a constant that we do not need to determine now.

Notice that gravity acceleration does not appear in the final result. Indeed, the velocity distributions are completely independent of the force field, which might not exist at all. In our argument, the weight field played a purely ancillary role in helping us to establish a relation between the elevation distribution, which we knew, and the velocity distribution, which we wanted to find. The physical mechanism determining the velocity distribution at equilibrium has nothing to do with the field, being the result of the collisions between molecules.

Also note that we implicitly assumed that molecules passing from z to z' do not collide. This hypothesis can be held because, the motion being disordered, if a molecule changes velocity, there is, on average, another one taking its place.

Clearly, the distributions of the other velocity components are identical, namely

$$f(\upsilon_x) = f_0 e^{-\frac{\frac{1}{2}m\upsilon_x^2}{k_B T}}, \quad f\left(\upsilon_y\right) = f_0 e^{-\frac{\frac{1}{2}m\upsilon_y^2}{k_B T}}. \tag{5.49}$$

We see that the velocity component distributions are Gaussian with zero mean value. Indeed, the number of molecules in one direction is equal to those in the opposite one.

We can now, finally, determine the (composed) probability of finding a molecule with an x-component of velocity between υ_x and $\upsilon_x + d\upsilon_x$, a y-component between υ_y and $\upsilon_y + d\upsilon_y$ and a z-component between υ_z and $\upsilon_z + d\upsilon_z$ in the unit volume. It is simply the product of the three, independent probabilities:

$$
\begin{aligned}
f\left(v_x, v_y, v_z\right) dv_x dv_y dv_z &= F_0 e^{-\frac{mv_x^2}{2k_BT}} e^{-\frac{mv_y^2}{2k_BT}} e^{-\frac{mv_z^2}{2k_BT}} dv_x dv_y dv_z \\
&= F_0 e^{-\frac{mv^2}{2k_BT}} dv_x dv_y dv_z .
\end{aligned}
\tag{5.50}
$$

Notice, in particular, that the velocity component distribution is independent of the direction, as it should be. Otherwise, if some direction were more probable, we should observe a collective motion in that direction.

5.7 Velocity Magnitude Distribution of Molecules

We shall now find the distribution of the velocity magnitude $f(\upsilon)$ of our gas molecules. In other words, we must look for the fraction of molecules with speeds between υ and $\upsilon + d\upsilon$ is $f(\upsilon)d\upsilon$. We work in the “velocity space”, in which the axes are the velocity components. We consider an infinitesimal element, as shown in Fig. 5.8.

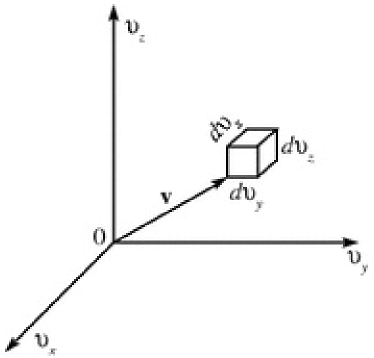


Fig. 5.8 An infinitesimal element in the velocity space

In the velocity space, the molecules having velocity in magnitude between υ are $\upsilon + d\upsilon$ are those whose representative vector has its head in the spherical shell of radiuses υ and $\upsilon + d\upsilon$. The volume of the shell is $4\pi\upsilon^2 d\upsilon$. From Eq. (5.50), the number of molecules it contains is

$$
dn_p = \alpha e^{-\frac{mv^2}{2k_BT}} v^2 dv \tag{5.51}
$$

where α is a constant, called the normalization constant , in which we have included the factor 4π. The normalization condition is the condition that the total number of molecules per unit volume must have a certain value. Let it be n_p. We must solve the equation in α

$$n_p = \int dn_p = \alpha \int_0^{+\infty} \upsilon^2 e^{-\frac{m\upsilon^2}{2k_BT}} d\upsilon.$$

The integral is a classic one, which can be found in the books in the form

$$\int_0^{+\infty} x^2 e^{-x^2} dx = \sqrt{\pi}/4. \tag{5.52}$$

Immediately, we obtain α and substitute in Eq. (5.51), obtaining

$$dn_p = 4\pi n_p \left(\frac{m}{2\pi k_B T}\right)^{3/2} \upsilon^2 e^{-\frac{m\upsilon^2}{2k_BT}} d\upsilon. \tag{5.53}$$

Dividing by the number of molecules n_p, we have the probability that a molecule has a velocity between υ and $\upsilon + d\upsilon$

$$f(\upsilon)\, d\upsilon = 4\pi \left(\frac{m}{2\pi k_B T}\right)^{3/2} \upsilon^2 e^{-\frac{m\upsilon^2}{2k_BT}} d\upsilon.$$

This corresponds to the velocity distribution function

$$f(\upsilon) = 4\pi \left(\frac{m}{2\pi k_B T}\right)^{3/2} \upsilon^2 e^{-\frac{m\upsilon^2}{2k_BT}}. \tag{5.54}$$

This is the Maxwell-Boltzmann distribution and is shown in Fig. 5.9.

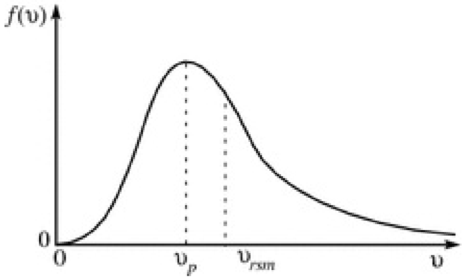


Fig. 5.9 The Maxwell-Boltzmann velocity distribution of ideal gas molecules

Let us discuss this distribution. It is the product of the square velocity υ^2 and of the Boltzmann factor of kinetic energy. The latter is the probability per unit velocity space volume, the former is proportional to the available velocity space volume for a given elementary interval of velocities $d\upsilon$. This volume is very small when its velocity is small, and grows with it. The factor υ^2 dominates at small velocities, corresponding to an almost parabolic growth in probability. At higher velocities, the decreasing exponential factor gradually takes over. The probability has a maximum, υ_p in the figure, which is the most probable velocity (it is the *mode* of distribution). Its value is obtained through the usual methods of calculus, which gives us

$$v_p = \sqrt{\frac{2k_B T}{m}}. \tag{5.55}$$

The mean values of the velocity components are obviously zero, as we have already noticed. Such is not the root mean square velocity (r.m.s. velocity, for short), which is the square root of the mean of the square velocity. This important quantity is obtained by computing the integral

$$v_{rms} = \left[\frac{1}{n_p}\int_0^{+\infty} f(v)\, v^2 dn_p\right]^{1/2} = \left[\frac{4\pi}{n_p}\left(\frac{m}{2\pi k_B T}\right)^{3/2}\int_0^{+\infty} v^4 e^{-\frac{mv^2}{2k_B T}} dv\right]^{1/2}.$$

The result is

$$v_{rms} = \sqrt{\frac{3k_B T}{m}}. \tag{5.56}$$

Comparing this with Eq. (5.55), we see that the r.m.s. velocity is larger than the most probable velocity. This is a consequence of the asymmetry of the distribution function that has a long “tail” on the higher velocity side. Notice that both averages are inversely proportional to the square root of the molecule mass. Consider two gases at the same temperature, N2 and He, for example. The smaller mass molecules (He) move faster than the heavier ones (N2). The average kinetic energies are the same for the two gases at the same temperature.

Let us consider the orders of magnitude. A relevant quantity is $k_B T$, which has the dimension of energy. Its value at room temperature, $T = 300$ K in a round figure, is

$$k_B T = 4.1 \times 10^{-21}\ \text{J}. \tag{5.57}$$

This is an atomic scale quantity and is convenient to express in electronvolts. We defined it in Sect. 4.1. We recall its value in joules:

$$1\ \text{eV} = 1.6 \times 10^{-19}\ \text{J}. \tag{5.58}$$

In this unit, at room temperature, we have

$$k_B T = 25.6\ \text{eV} \cong 1/40\ \text{eV}. \tag{5.59}$$

This is an important value to remember. Another relevant number is the temperature at which $k_B T = 1$ eV. Calling it $T(1\ \text{eV})$, it comes out as

$$T(1\ \text{eV}) = 11{,}600\ \text{K}. \tag{5.60}$$

Let us evaluate, in order of magnitude, the r.m.s. velocity at the room temperature of the lightest gas, H2. Neglecting the electron contribution, the mass of the molecule is two proton masses, Eq. (1.5), namely

$$m_{\mathrm{H}_2} \cong 3.3 \times 10^{-27}\ \mathrm{kg} \tag{5.61}$$

and thus we have $v_{rms}^{\mathrm{H}_2} = \left[3 \times 4.1 \times 10^{-21} / \left(3.3 \times 10^{-27}\right)\right]^{1/2} = 1.9\ \mathrm{km/s}$.

If we want the r.m.s. velocity for another gas at another temperature, we just have to scale it with the root of the temperature to mass ratio. For example, for silver (monoatomic molecule, *A* = 107.9) at 1000 °C (1273 K), we have $v_{rms}^{\mathrm{Ag}} = 560\ \mathrm{m/s}$.

Consider now a mixture of two or more gases in equilibrium. The average kinetic energies of the molecules of the different gases are equal, but the r.m.s. velocities are different, as our examples have just shown us. This is also true for a particulate suspended in a gas. This can be approximated with a “gas” of microscopic particles. In equilibrium, their average kinetic energy is the same as that of the molecules, but, their masses being much larger, their r.m.s. velocities are much smaller.

Question Q 5.1. Consider a particulate of equal spherical particles of 10 μm diameter and density of 2000 kg/m^3 in air at room temperature. Calculate the r.m.s. velocity.

5.8 Experimental Controls

The results of the previous sections have been reached through theoretical arguments. The conclusions can be considered valid only after having been checked by the experiment. The experimental controls have been many. We shall discuss two of them now.

As we noticed at the end of the previous section, the thermal equilibrium between two mixed “gases” may exist even if one of them is really a powder, a set of microscopic corpuscles. At equilibrium, these particles have the same average kinetic energy as the gas molecules. If the particle sizes are on the order of the micrometer, they can be seen with a microscope and some characteristics of their motion can be measured. In 1827, the botanist Robert Brown (Scotland, 1773–1858) was using a microscope to observe grains of pollen of a plant (Clarkia pulchella, to be exact) suspended in water. He observed minuscule particles, ejected by the pollen grains, executing a continuous, chaotic, jittery motion. He thought, at first, that he was observing living beings, but was soon able to exclude that possibility after observing the same phenomenon in inorganic particles of similar size. This is called Brownian motion .

Brownian motion was theoretically studied by Albert Einstein (Germany 1879—USA 1955) in 1905 and Marian Smoluchowski (Austria, 1872–1917) in 1905-1906. We summarize the interpretation as follows. The diameters of the particles in suspension, on the order of a micrometer, as we mentioned, are still enormous, four orders of magnitude larger when compared to water molecules, which are about half a nanometer across. Each of the Brown particles is continuously hit by the water molecules at the considerable rate of about 10^{20} per second. The received impulses have all the directions, but their effects do not always average out at zero if taken in short enough intervals of time. In this case, there might be more collisions on one side than on the other and the particle will jump in the unbalanced direction. These fluctuations become more important the smaller the size of the particle in suspension.

This phenomenon was quantitatively studied by Jean Baptiste Perrin (France, 1870–1942) in 1908. To this purpose, he had to prepare micrometer-size particles, all with the same diameter and the same mass. He succeeded by rubbing gamboge (a pigment extracted from the resin of tropical plants of the genus Garcinia) between his hands under water. He obtained an emulsion that, under a microscope, appeared to contain particles of different sizes. To produce a uniform emulsion (consisting of equal grains), he used a process of fractional centrifugation, profiting off the fact that the parts that settle out first are richer in larger grains. This was not at all an easy process. Indeed, after several months of careful work, Perrin was only able to obtain a few decigrams of particles of the desired size from one kilogram of gamboge. Experiments could then start. With the microscope, he observed a completely disordered motion, which never slowed down nor stopped. He defined it as eternal and spontaneous. He also observed that the smaller spheres moved faster than the larger ones.

Perrin designed and performed a series of experiments to check whether or not the system of particles behaved as foreseen by statistical mechanics . In the first experiments, he studied the equilibrium distribution of the emulsion of spherules under the action of gravity. It was exactly what was foreseen in Eq. (5.35). Knowing the mass of the spherules and the temperature, he could extract a value of the Boltzmann constant from the data and hence of the Avogadro number . Perrin repeated the experiment, changing the mass of the granules, the intergranular liquid and the temperature. In all of them, the resulting value of the Avogadro number was the same, within the experimental uncertainties.

Further tests had to be done on the velocity distribution . He knew, however, that these could not be directly measured, because the granules were changing velocity in extremely short time periods. To get around the problem, Perrin put the method theoretically proposed in 1905 by Einstein into practice. Let us fix our attention on one spherule. We fix a certain time interval Δt and measure the distances Δl between the positions of the particle in a sequence of such intervals. Notice that these are not the distances traveled by the spherule in those intervals, because, during this time, it moved in a random zigzag, rather than straight. Perrin projected the microscope image onto a sheet of paper with the so-called camera lucida technique. He worked with an assistant, one of them marking the paper with a "dot" to indicate the position of the spherule under observation every, say, 30 s whenever the other, who was looking at the chronometer, would call out for him to do so. One such "dotting" is shown in Fig. 5.10. The radius of the spherule was 0.53 µm. The positions were taken every $\Delta t = 30$ s. Einstein had given the relation between the velocity distribution and Δl distribution, and Perrin was able to show that the gamboge particles suspended in "gas" behaved as foreseen by statistical mechanics. Once more, the value of the Avogadro number extracted from the data was the correct one.

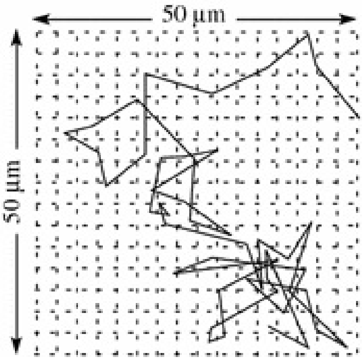


Fig. 5.10 Positions of a 0.53 µm diameter spherule in water every 30 s

Thanks to further developments in the technique, Perrin was also able to experimentally verify the energy equipartition between translational and rotational kinetic energy. The difficulty was that the Einstein's formula predicted a mean rotation of approximately 8° per 1/100 of a second for a sphere of 1 µm diameter. This is too rapid to be measurable. Perrin overcame the difficulty by producing bigger spheres of mastic of about 12 µm in diameter. They were limpid like glass spheres. The rotation was not observable. Some of them, however, contained a visible defect that could be used as a mark by which the rotational Brownian movement could be perceived. He noted at equal intervals of time the successive positions of a certain defect, from which it was possible to find the orientation of the sphere at each of these moments and to calculate its rotation from one moment to the next. He found the average rotation kinetic energy per degree of freedom to be equal to the translational one within the experimental accuracy of about 10 %.

The direct experimental verification of the gas kinetic model became possible starting in 1911, when Lois Dunoyer (France, 1880–1963) produced the first molecular rays or molecular beams . The possibility itself of producing the beams is a direct proof of one fundamental assumption of the theory, namely that, in gases, the molecules move in straight lines until they collide with another molecule or a wall of the containing vessel.

The next step, credited to Otto Stern (Germany, 1892–1969), was the direct measurement of the velocities. A typical arrangement is shown in Fig. 5.11. The oven F contains the gas under study at a high temperature T. A small aperture in the wall of the oven emits molecules in a range of directions. A second aperture, selects one direction and produces the molecular beam. The beam is in a vacuum, to avoid its molecules hitting any environmental gas molecules. After the second slit, there is a velocity selector. In its simplest form, this is made out of two disks, one separated from the other at a distance l of several centimeters. The disks are on the same axis, have two slits out of phase by an angle θ, and rotate with angular velocity ω. The detector R placed beyond the slits, can be reached only by the molecules whose velocity is such that the time taken to go through the distance l is exactly the same as that taken by the disks to rotate by the angle θ. Namely, it should be $\omega/\theta = \upsilon/l$. In other words, only the molecules with speed $\upsilon = l\omega/\theta$, within an interval determined by the width of the slits, are detected.

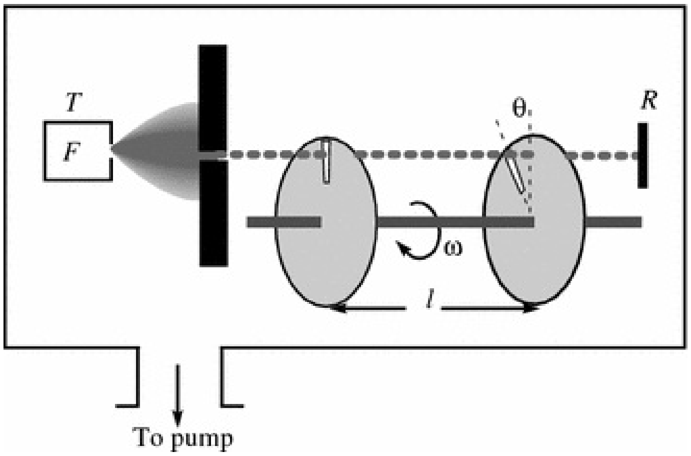


Fig. 5.11 Molecular beam experiment for measuring the molecular speeds

The experiment is done taking data with a series of angular velocity values, corresponding to the same number of molecule velocities, in the above-mentioned (small) interval. In the next section, we shall discuss a detector for a particular type of molecule, the alkali metals. In any case, the detector delivers an electric current, whose intensity is proportional to the number of molecules on the detector per unit time. This number is proportional to the number of molecules per unit volume with velocities in the selected interval, multiplied by the velocity itself. If kinetic theory is correct, the detector's current intensity I should then be

$$I(v) = \alpha v^3 e^{-\frac{mv^2}{2k_B T}}. \tag{5.62}$$

where α is a proportionality constant that we do not need to determine. Notice the extra power of velocity in front of the exponential.

Figure 5.12 shows the results of a precise measurement made by R.C. Miller and P. Kausch in 1955.[1] They developed a velocity selector based on the same principle as the one we described. In brief, instead of two disks, they used a cylinder with a helical slit, as if it were a continuous series of disks instead of two at a distance. The detector used a hot tungsten wire, based on the surface ionization effect, which we shall describe in the next section. The curve in the figure is Eq. (5.54) (normalized to the data) as a function of the selected velocity divided by its most probable calculated value υ_p. The temperature was $T = 466 \pm 2$ K, and the pressure reported in the insert in the figure is the pressure of the K gas in the oven. The calculated value of the most probable velocity was $\upsilon_p = 628 \pm 2$ m/s, while the measured one was $\upsilon_p = 630 \pm 2$ m/s. As one can see, the agreement between theory and data is excellent. However, looking carefully, one notices that, at the lowest velocities, the measured points are systematically a bit below the curve. This is an instrumental effect, due to the collisions of the potassium molecules when they are in the oven.

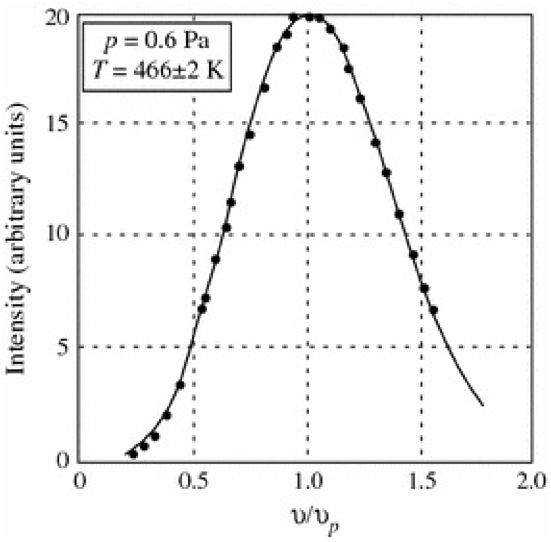


Fig. 5.12 Velocity distribution measure by R.C. Miller and P. Kaush in a K beam

---

5.9 Applications of the Boltzmann Law

The Boltzmann law states that, in a macroscopic system in equilibrium at the temperature *T*, the ratio between the number of microscopic components (molecules or others) per unit volume n_{p1} with energy U_1 and number per unit volume n_{p2} with energy with energy U_2 is

$$n_{p1}/n_{p2} = e^{-\frac{U_1 - U_2}{k_B T}}. \tag{5.63}$$

The Boltzmann law is extremely important, because it appears ubiquitously in all the statistical systems at equilibrium. In a number of circumstances, the Boltzmann factor , as the exponential on the right-hand side is called, is enough to describe the orders of magnitude of a problem. If one considers the details, this factor might be multiplied by some function of temperature. However, in practice, the exponential varies with temperature much quicker than any other function, dominating the behavior of the system. We shall now consider several examples. Some of them involve rather complicated situations, but we shall limit ourselves to rough evaluations. This type of evaluation is often extremely useful, especially in a complex problem.

Evaporation of a liquid.

Consider a liquid and its vapor in a container in equilibrium at the temperature *T*. We have already discussed the problem from the thermodynamic, hence macroscopic, point of view. Let us discuss it now from the mechanical statistical, namely microscopic, point of view. We shall find that the Clapeyron equation is a consequence of classical mechanics.

Let n_{pg} and n_{pl} be the number of molecules per unit volume in the vapor and liquid phases, respectively. In a closed system, their sum, $n_{pg} + n_{pl}$, is obviously a constant. The system is in statistical equilibrium. At any time interval, some molecules in the liquid have enough kinetic energy and are close enough to the surface to skip into the vapor (winning the attraction of their “liquid sisters”), while quite a few vapor molecules, on average, fall into the liquid.

Let W be the work necessary to take a molecule, with zero kinetic energy, from inside the liquid to outside of it. This is the energy difference between outside and inside. The Boltzmann law says that the ratio of the numbers of molecules per unit volume outside and inside is the Boltzmann factor:

$$n_{pg}/n_{pl} = e^{-\frac{W}{k_B T}}. \tag{5.64}$$

If we want to be rigorous, this expression does not tell us very much. Suppose, for example, we want to know the details of the dependence on temperature of n_{pg}, and consequently of the vapor pressure to which it is proportional. We then write

$$n_{pg}(T) = n_{pl}(T)\, e^{-\frac{W}{k_B T}}.$$

where the temperature dependence has been explicitly written. On the right-hand side, n_{pl} is the number of molecules in the liquid phase *per unit volume*. This quantity depends on temperature, because the volume of the liquid depends on temperature and the Boltzmann law does not tell us how. However, if we are far from the critical point (where gas and liquid densities are equal), then it is $n_{pg} \ll n_{pl}$, namely the exponential factor is very small. To put it another way, the exponent is large in absolute value

$$\frac{W}{k_B T} \gg 1. \tag{5.65}$$

Under these conditions, a small variation in temperature makes the Boltzmann factor vary strongly, dominating, in this way, the temperature dependence of n_g. Consider water, for example. Its vaporization molar heat is 40.5 kJ/mole. This means that 40.5 kJ are needed to take an Avogadro number of molecules outside the liquid. The work W for a single molecule is then 40.5 kJ/N_A. Expressing energy in electronvolts, we find $W = 0.4$ eV. Notice that W is substantially the depth of the van der Waals potential. The exponential factor in

Eq. (5.65) at room temperature is

$$e^{-\frac{W}{k_B T}} = e^{-17} = 4.1 \times 10^{-8}. \tag{5.66}$$

which is very small.

In practice, if we are interested in the behavior of the system in broad terms, we can consider n_l constant, at least in a limited temperature range, and write

$$n_{pg}(T) \cong \text{const}\, e^{-\frac{W}{k_B T}}. \tag{5.67}$$

Let us now look at what thermodynamics specifically says. This is the Clapeyron equation for one mole. If p is the saturated vapor pressure, V_g and V_l the molar volumes of vapor and liquid, respectively, and Q_V the molar latent heat of vaporization, the equation is

$$\frac{dp}{dT} = \frac{Q_V}{T\left(V_g - V_l\right)}. \tag{5.68}$$

Under the conditions discussed in Sect. 5.7, the equation can be (approximately) integrated, obtaining

$$p = p_0 e^{-\frac{Q_V}{RT}} = p_0 e^{-\frac{Q_V}{N_A k_B T}}. \tag{5.69}$$

Considering that the vapor pressure is proportional to the numeric density of the molecules in vapor gas n_{pg}, we see that Eqs. (5.67) and (5.69) say the same, with

$$Q_V = W N_A. \tag{5.70}$$

Namely, the molar vaporization heat is the work to be done against the van der Waals forces to take an Avogadro number of molecules (one mole) out of the liquid in the gas. Let us now compare the two approaches, thermodynamic and statistical mechanic, which led to the same result. Thermodynamics allows us to find the Clapeyron equation Eq. (5.68), which is exact. In addition, it holds for the entire phase transition, using the proper parameters. The latter, namely the molar volumes of the two phases, cannot be foreseen, not even in principle. Statistical mechanics, on the other hand, has given us, with very simple arguments starting from mechanics, the approximate Eq. (5.67). In such a way, it has shown the physical meaning of the equation. In addition, it allows us to calculate, at least for simple systems, Q_V and $V_g - V_l$ starting from knowledge of the intermolecular forces. The two approaches are complementary. Thermodynamics is more powerful, especially for complex systems, but it tends to hide the physical meaning.

Thermionic emission .

As we have already stated, metals are solid bodies having microcrystalline structure. Namely, they are aggregates of microcrystals that are invisible to the naked eye but can be seen with a microscope. The atoms of the microcrystals are not neutral, but ionized, because some (one or two, in general) of their electrons are not bound to "their" atom, but are free to move about inside the crystal. They cannot, however—at least in a large majority as we shall now see—abandon crystal entirely, being globally attracted by the ions. All these electrons form a sort of gas, which makes the metal a good conductor, both of heat and electricity. They are, consequently, called conduction electrons .

The work that is necessary to take an electron from inside to outside the metal, at zero kinetic energy, is the *work function*, which we shall again call W . It is characteristic of the metal. Its magnitude is on the order of the electronvolt.

Consider the metal at temperature T. This is also the temperature of the conduction electrons' gas. Some of these, those hotter and closer to the surface, may be able to jump out. To have equilibrium, as in the case of evaporation, the system must be closed inside a container and no other gas must be present. We can make a wire of the metal, close it inside a glass envelope and evacuate it. A gas of free electrons, namely those not bound to the metal, forms inside the envelope. Their density is extremely small at normal temperatures but increases as the temperature increases, as we shall see. In every time interval, there are electrons coming out of and electrons falling inside the metal. The two rates are equal at the statistical equilibrium. This situation is very similar to evaporation. We can say that the number density of electrons in the "gas" n_{pg} depends upon temperature according to the Boltzmann law

$$n_{pg}(T) \cong n_{p0}\, e^{-\frac{W}{k_B T}} \tag{5.71}$$

where n_{p0} is the number of electrons in the metal per unit volume. Once more, it is approximately constant, namely its variations in regard to temperature are much smaller than those of the Boltzmann factor. To be precise, n_{p0} is not rigorously constant because the volume depends on temperature. The effect we are considering is called the *thermionic effect* . It is exploited in the electronic valves.

We can do an experiment to control Eq. (5.71) as follows. We make a thin wire of the metal under study, a few centimeters long and a couple of millimeters in diameter, we fold it and position it on the axis of the structure, as in Fig. 5.13. We place the folded wire inside of a metallic cylinder. We enclose the structure in a glass envelope, having two conductors joined to the wire and cylinder coming out of the glass container, as shown in Fig. 5.13. We evacuate the air and seal the envelope hermetically. We have thus built a thermionic diode. We get an electric current through the wire to heat it at a high temperature, on the order of 1000 K. The work functions are typical, on the order of 1 eV corresponding to $k_B T$ = 11,600 eV. The Boltzmann factor is on the order of e^{-10}. Consequently, in this case as well, it dominates the temperature dependence.

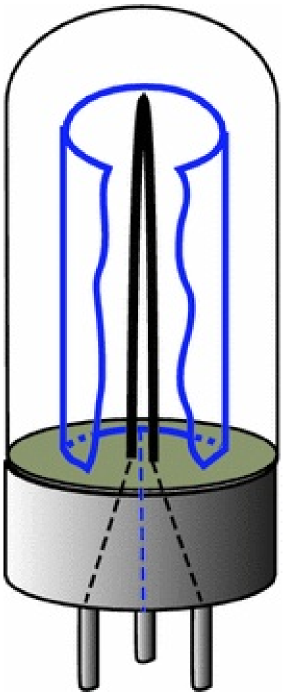


Fig. 5.13 Thermionic diode

In our experiment, we do not establish the equilibrium. Instead, we apply a potential difference between the wire (the cathode) and the small tube (the anode), generating an electric field in the space between them. The corresponding force accelerates the electrons as soon as they exit the wire and brings them to the anode. We measure the current leaving the cathode. In this cylindrical geometry, which is used in practice, the electron density in the gas n_{pg} is not the same as in the entire volume, but is a decreasing function of the distance from the cathode. To make the argument simpler, let us assume a plane geometry, namely that the cathode and anode are two parallel planes of area A. Assume also that all electrons have the same velocity υ. The number of electrons leaving the cathode in one second is their number in a parallelepiped of base A and height υ, namely $n_{pg}\upsilon A$. The current intensity is this number times the electron charge q_e. In reality, electrons have different velocities, as we know, and we must take a suitable average ⟨υ⟩, which we do not need to specify. Finally, we can write for the current intensity

$$I = \text{const}\, q_e \langle \upsilon \rangle A e^{-\frac{W}{k_B T}}. \tag{5.72}$$

Once more, the equation does not tell us everything, because $< \upsilon >$ is an unknown function of temperature. The corresponding experimental law was found by Owen Williams Richardson (UK, 1879–1959) in 1901. This was only four years after the discovery of the electron by Joseph John Thomson (UK, 1856–1940). The Richardson law is

$$I = \text{const}\, A T^2 e^{-\frac{W}{k_B T}} \tag{5.73}$$

which contains an extra factor T^2. This factor is important if we are looking at the details, but its temperature dependence is much slower than that of the Boltzmann factor.

Surface ionization .

The detection of an atomic or molecular beam (such as the ones we have mentioned in Sect. 5.8) presents some difficulties. These are due not only to the fact that molecules are neutral, but also to the fact that the characteristics of the beam molecules do not differ much from those of the residual gas that is always present, even in an evacuated apparatus. To understand the orders of magnitude, consider that the beam densities might be typically on the order of 10^{14} atoms/m^3 and that velocities are on the order of 500 m/s, which is on the same order of the residual gas molecules' velocity. If the vacuum residual pressure is, for example, 10^{-9} Pa, the molecule density is 3×10^{15} molecules/m^3, which is on the same order as the beam.

A particularly lucky case is that of alkali metals , which can be detected with efficiency close to 100 %. This is due to the circumstance that, for them, one of the electrons is only weakly bound to the atom (for this reason, they are monovalent, as we already mentioned). This is called valence electron. The binding energy, namely the work to be done to take the electron off, is of a few electronvolts. This is the *ionization energy I* (for example, $I = 5.2$ eV for Li, 5.1 eV for Na, 4.3 eV for K).

On the other hand, the work function W of certain metals, like tungsten and platinum, is particularly large ($W = 6$ eV for W), larger than the ionization energy of the alkalis. As a consequence, if an alkali atom is very close on the atomic scale, namely by a fraction of nanometers, to the tungsten surface, it becomes energetically favorable for its valence electron to jump into the metal.

This happens with a certain probability, which is, once more, given by the Boltzmann factor . In other words, the ratio between the number of atoms that ionizes n_{p1} and the atoms that remain neutral n_{p0} is

$$n_{p1}/n_{p0} = e^{-\frac{I-W}{k_B T}}. \tag{5.74}$$

Notice that, in this case, $W > I$; hence, the exponent is positive and the larger part of the electrons goes to the tungsten. This is the *surface ionization* phenomenon.

To detect an alkali atom beam, we prepare a tungsten wire with a very clean surface and bring it to a high temperature (otherwise, the atoms get absorbed). We position a small metallic plate near to the wire and give it a negative electric potential. The beam atoms hit the wire, a large majority leaving an electron and bunching back as positive ions, and are attracted by the plate. The effect is an electric current that we can measure. Let us evaluate the order of magnitude of its intensity. Assuming 100 % efficiency, the current intensity is the number of ions of the beam per unit volume times the velocity of the ions times the area of the detecting wire times the charge of the ion. With the above mentioned values of density and velocity and a wire of 0.1 mm diameter 1 cm long (area 10^{-6} m^2), we have $I = 10^{14}$ atoms/m^3 × 5 × 10^2 m/s × 10^{-6} m^2 × 1.6 × 10^{-19} = 8 nA, which, as a current intensity, is small but not difficult to measure.

Gas molar heats

We conclude the section by giving a few hints as to the quantum mechanical explanation of the *freezing out of the degrees of freedom* and consequently of the molar (or specific) heats we discussed in Sect. 5.2. Let us fix our attention on the diatomic molecule, specifically on its oscillations. We represent the molecule as a harmonic oscillator. Classically, with continuity, the oscillation energy can have any value from a certain minimum, say E_0, and 0. The minimum energy occurs when the two atoms stand with zero velocity at the distance at which the potential energy is a minimum, and the maximum energy, zero, occurs when they break apart. Quantum mechanics tells us that this is not true. Energy is “quantized”. The total energy of the oscillator can only have discrete values differing one from the next by the well-defined quantity ΔU. In other words, the energies the quantum harmonic oscillator can have are in the sequence

$$U_0, \quad U_1 = U_0 + \Delta U, \quad U_2 = U_0 + 2\Delta U, \quad \ldots \tag{5.75}$$

The Boltzmann law, however, still holds. Let us consider the state of minimum energy U_0 and that of energy U_1 immediately following it. The ratio of the probabilities of a molecule being in the two states is, once more, given by the Boltzmann factor, namely

$$\frac{f_1}{f_0} = \frac{e^{-\frac{U_0+\Delta U}{k_B T}}}{e^{-\frac{U_0}{k_B T}}} = e^{-\frac{\Delta U}{k_B T}}. \tag{5.76}$$

If the temperature is low enough, such that

$$\Delta U \gg k_B T \tag{5.77}$$

then the absolute value of the negative exponent on the right-hand side of Eq. (5.76) is large and the exponential is extremely small. Consequently, the probability that the oscillator can have the smallest energy larger than the minimum is extremely small. Clearly, this is even truer for the higher energies. If we now consider a set of a large number of identical oscillators, we see that practically all of them must be in the lowest energy state. The degrees of freedom internal to the molecule cannot be excited, because they are “frozen”.

Let us look more closely at the processes. The energy can also “equipart” to the internal motion only if energy can be transferred in the collisions from translation or rotation energy to vibrations. But, for quantum mechanics, there is a minimum energy transfer, a quantum of energy , which is ΔU. Clearly, if the mean kinetic energies of the translational and rotational motions, i.e., $k_B T$, are much smaller than the energy quantum, this, in practice, cannot happen. Classically, on the contrary, energy can be transferred in any quantity, even minuscule, and equipartition always happens. We see how the problem is in the very nature of classical physics, in which natural quantities are continuous. Some of them, including energy, are not. According to a Latin proverb, *Natura non facit saltus* = Nature does not make jumps. But Nature does what she likes and indeed, *facit saltus*.

On the other hand, if the temperature is high enough, such that

$$\Delta U \ll k_B T \tag{5.78}$$

then $k_B T$ is much larger than the energy quantum ΔU and quantum mechanics tends toward classical mechanics. The latter then gives correct predictions. We understand now how classical mechanics predictions work better at high enough temperatures. We profit from a very general statement. Quantum mechanics does not show that classical mechanics is wrong. Quantum mechanics is valid over a much wider range, but it includes classical mechanics, tending toward it under well-defined conditions.

We have seen that the temperature limits of classical mechanics, the Debye temperature, are different for different gases. The reason is that the energy quantum ΔU is different for different gas molecules. This quantity is proportional (by a fundamental constant called the Planck constant) to the proper oscillation frequency of the molecule, given by Eq. (5.17) for a diatomic molecule. As a matter of fact, the interatomic forces, and consequently the effective spring constant κ, do not differ very much from one gas to another. However, the atomic masses, and consequently the reduced mass μ, change considerably. As a consequence, the proper frequency is lower for the heavier gases and classical mechanics gives correct predictions for them down to lower temperatures. The explanation of the limits of classical mechanics for the molar heats of solids (Sect. 5.3) is similar.

5.10 Nature of Irreversibility

All thermodynamic processes can ultimately be reduced to the motions of their molecules and atoms, namely to mechanical processes. We have learnt that, at the microscopic level, all forces are conservative. As a consequence, all mechanical processes at the molecular level are reversible. This seems to stand in contradiction to the fact that the real thermodynamic processes are irreversible. We shall now see that the contradiction is only apparent. Let us start by discussing two examples.

We build a simple pendulum, attaching a small sphere to a thin wire that we fix to the ceiling of a box with adiabatic walls, as in Fig. 5.14a. The box, which is closed off by insulating walls, also contains air. The system is in mechanical and thermodynamic equilibrium. In particular, pendulum and air have the same temperature. We take the pendulum out of its mechanical equilibrium position and let it go. The system is no longer in equilibrium. The oscillations continue for a while, but their amplitude gradually decreases, and finally, the pendulum comes to rest. The system is back in mechanical and thermodynamic equilibrium. Measuring with a sensitive thermometer, we find that the final temperature of the system is higher than when we began.

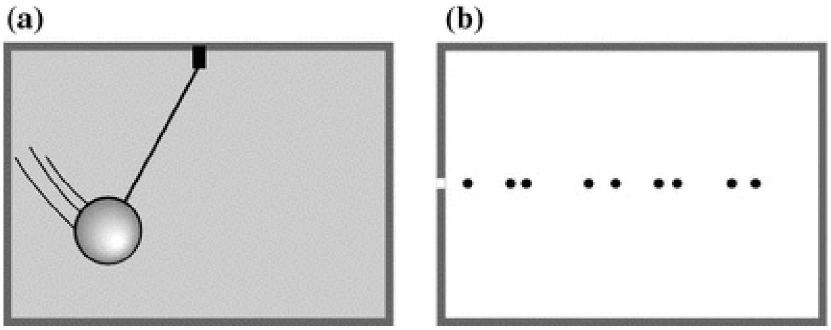


Fig. 5.14 Statistical non-equilibrium states of systems of molecules. **a** Partially ordered motions in a

pendulum, **b** ordered motion in a box

Let us now consider the pendulum as a system of molecules, and compare its status at its first passage through the equilibrium position, when all the macroscopic energy is kinetic, with that at the final position. The macroscopic kinetic energy of the body has disappeared, having transformed in the kinetic energy of the disordered thermal motion of its molecules, as certified by the increase in temperature. If we only consider the (macroscopic) mechanical energy, we think that the process has dissipated energy, but when we consider the microscopic motion as well, we see that energy was conserved. This is the first law of thermodynamics. Let us look at the process from the point of view, so to speak, of the pendulum molecules. Initially, their motion has two components, an ordered one in which all of them have the same velocity due to the motion of the pendulum and a disordered one corresponding to the internal kinetic energy. In the final state, the "ordered" kinetic energy has transformed into kinetic energy of the disordered motion of the pendulum and air molecules. Clearly, the initial energy can be distributed amongst the huge number of molecules in an incredibly large number of equivalent ways. In other words, the final state in which the motion of the pendulum has ceased can be realized in a number of molecular states much, much larger than the initial state in which a large fraction of energy is in the ordered motion.

The second example is a thought experiment, shown in Fig. 5.14b. A beam of molecules enters through a hole in the wall of a container in which we had created a complete vacuum. This is impossible in practice, but we are dealing with a thought experiment. In the initial state, all the molecules move orderly in the same direction, one after the other. The state is not of thermodynamic or statistical equilibrium. After a certain number of molecules have entered, we close the hole. The molecules will repeatedly hit the walls of the container. At the nanometric scale, any real "plane" wall is not a plane at all, but shows bumps and hollows of different sizes, shapes and directions. Consequently, each collision changes the direction of the incoming molecule casually and unpredictably. Soon, the motion of the molecules is completely disordered. They move around the entire container in all directions in the chaotic motion of the ideal gas. If we repeat the experiment, we always obtain the same result. The initial state can only be realized in a few molecular arrangements, the final one in an enormous number of molecular positions and velocities.

The described processes are completely compatible with the mechanical reversibility. Suppose we can operate on the system at a certain instant by leaving all the molecules in their position and inverting all their velocity vectors.

The subsequent motion will be the exact original one running inversely. We might, ideally, have shot a movie of the first process. If we were to play it backwards, we would see the second process. Molecules will hit one another and the walls, but after the exact same time that was passed since we closed the hole, all the molecules would be moving in order in the same direction, exiting the hole one after the other.

The conclusion, a theorem proved by Boltzmann , is that the transition from a thermodynamic non-equilibrium to the equilibrium state corresponds to the transition from a state that can be realized at the microscopic level in a small number of ways to one that can be realized in an enormous number of ways. The thermodynamic equilibrium state is the state with the maximum possible microscopic realizations. We see that the second law of thermodynamics has a probabilistic nature. In principle, the passage of heat from a colder to a hotter body and the spontaneous transformation of internal energy into mechanical macroscopic energy (and no other effects in both cases), are not rigorously impossible, but they are extremely improbable. The irreversibility of the thermodynamic processes is ultimately due to the enormous number of molecules.

The number of molecules in the macroscopic bodies is *really enormous*. As a consequence, the words “extremely improbable” do effectively mean “impossible”. Let us look at an example: the free expansion of an ideal gas. We have an adiabatic container divided into two halves by a wall. Initially, the gas is in one half, while the other half is empty. We open a small hole in the separating wall and the gas expands to occupy the entire container. Its volume has doubled and, as we know, the temperature has not changed. In the final state, each molecule in its disordered motion, on average, spends half of the time in the half volume on the left, half of the time on the right. As a consequence, the probability of it being in the initial volume is ½. The gas being ideal, its molecules are independent of one another. Hence, the probability of finding N molecules on the same side is $(1/2)^N = 2^{-N}$. The number of molecules in a mole is the Avogadro number. Its order of magnitude is 10^{23}. The probability of finding all of the molecules on one side is then $2^{-10^{23}}$, an inconceivably small number. This means that we might be able to see all of the gas in half of the container once in $2^{+10^{23}}$ experiments. If we were able to do such an experiment in one second, this would happen every one million times the age of the universe.

However, if the number of molecules in the system is not so huge, spontaneous displacements from equilibrium are observable. These are small and last for only brief intervals of time. They are called *fluctuations* . As a matter of fact, we have already seen an example; the Brownian motion . Consider one of the Perrin spherules a few micrometers in diameter suspended in water. The number of collisions per second of a water molecule with a particle is huge, on the order of 10^{20}. Sometimes, the momentum that the particle receives from the collisions, for example, on its left are larger than that from those on its right. It then starts moving to the right. A bit of thermal energy has transformed into kinetic energy of an ordered motion of a micrometer-size body. Locally, a small transformation from disorder into order has taken place. At this scale the phenomenon is not impossibly improbable. However, for particles a little larger, the process is not observable.

Other examples are the temperature and density fluctuations. In any small region of a body, these quantities are not rigorously constant, but change a bit, going both up and down. However, these fluctuations are extremely small. For example, in a milligram of water at room temperature, the temperature fluctuations are on the order of 10^{-8} K (or 10 nK).

Boltzmann noticed, with reference to these phenomena, that

> …the existence of such cases does not disprove our theorem. On the contrary the theory of probability itself shows that the probability of such cases is not mathematically zero, only extremely small.

5.11 Entropy and Thermodynamic Probability

We shall now translate the arguments of the preceding section into mathematical equations. The second law of thermodynamics says that the non-equilibrium states of any isolated thermodynamic system evolve spontaneously towards states of larger entropy. This corresponds, from the statistical point of view, to the transition to a state that can be realized in a much larger number of ways. Clearly, the thermodynamic state function entropy is connected to the number of possible microscopic realizations. To express this connection, we must distinguish, for a given system, its thermodynamic state, called its *macrostate*, and its microscopic state, called its *microstate* . A macrostate is defined by the values of the thermodynamic variables (volume, temperature, pressure, concentration, etc.). A microstate is a mechanical state of molecules, defined by the coordinates and velocities of all the molecules ($6N$ variables in total, if N is the number of molecules). Obviously, if we specify the macrostate of the system, we do not define its microstate. In other words, a given macrostate corresponds to a certain number of microstates that realize it. This number is called the *statistical weight* or *thermodynamic probability* of the macrostate. The adjective "thermodynamic" is used because the quantity is a number larger than one, while the mathematical probability is normalized to 1, namely it is between 0 and 1. We shall indicate thermodynamic probability with Γ. The macrostates of isolated systems spontaneously evolve towards macrostates of larger entropy S and larger Γ. The relationship between the two quantities is

$$S = k_B \ln \Gamma. \tag{5.79}$$

This very important equation in physics was found by Boltzmann, starting from the laws of classical mechanics and probability theory. It has been engraved on the Boltzmann tombstone. We cannot prove the equation here, but we can justify the logarithmic dependence. Consider a system made of two parts. The number of its realizations Γ is equal (assuming the two parts to be independent) to the product of the numbers of realizations of each part separately, say Γ_1 and Γ_2. Hence, $\Gamma = \Gamma_1 \Gamma_2$, and Eq. (5.79) gives us

$$S = k_B \ln \Gamma = k_B \ln (\Gamma_1 \Gamma_2) = k_B \ln (\Gamma_1) + k_B \ln (\Gamma_2) = S_1 + S_2. \tag{5.80}$$

Indeed, the entropy of a system is the sum of those of its (independent) parts, the thermodynamic probability is the product of the probabilities of its parts, and the logarithm of a product is equal to the sum of the logarithms of the factors. Logarithm is the only function with this property.

The definition we gave of the thermodynamic probability of a macrostate being the number of microstates that realize it presents a difficulty. Consider a certain microstate. If we vary the position or velocity of a molecule a little, the macrostate does not vary. But coordinates and velocity are continuous variables, which can assume an infinite number of values. It looks like the thermodynamic probability is infinite. We avoid the difficulty as follows.

To define a microstate, we need $6N$ quantities. We can then represent the microstate with a point in a $6N$-dimensional space. This is called the *phase space* . We arbitrarily define an elementary cell as a portion of the phase space of small, but arbitrary, size and we define as being indistinguishable from the microstates whose representative points fall within the same cell. When we calculate the thermodynamic probability, we count them as one. The arbitrary choice of the elementary cell does not have consequences, as long as we consider entropy *differences*, as we did in our study of thermodynamics. Indeed, if we change the volume of the cell by a factor, the number of microstates corresponding to a given macrostate changes by the same factor. As a consequence of the logarithmic dependence, entropy changes by an additive constant while entropy differences remain unaltered. In other words, the consequence of the arbitrary size of the elementary cell is that entropy is defined up to an additive constant. We will only mention here that quantum mechanics gives a precise prescription for the size of the elementary cell and, consequently, of the entropy constant.

We have seen how thermodynamics defines the infinitesimal entropy variation as $dS = \delta Q/T$, where δQ is the heat reversibly received by the system and T is the temperature of the source (which is also that of the system, because the process is reversible). Let us now look at the physical reason for that. When we give heat to a system, we increase the thermal motion of its molecules. Namely, we increase the disorder or the number of microstates that realize the new macrostate. It is also clear that, for the same received heat, the disorder increase is inversely proportional to the thermal energy that already exists, i.e., the absolute temperature. This explains the thermodynamic definition.

We conclude by calculating the entropy variation in the free expansion of an ideal gas by the factor of 2 that we considered above, using the Boltzmann equation Eq. (5.79). We have already found that the ratio between the numbers of microstates corresponding to the final and initial macrostates is 2^N, if N is the number of molecules. The entropy variation is then

$$\Delta S = k_B \ln \frac{\Gamma_2}{\Gamma_1} = k_B \ln 2^N = N k_B \ln 2 = nR \ln 2 \tag{5.81}$$

where, in the last term, n is the number of moles. We have found the relation known from thermodynamics.

In Sect. 4.5, we found the expression of the entropy difference between two states f and i of a van der Waals gas , namely

$$S_f - S_i = nC_V \ln\left(\frac{T_f}{T_i}\right) + nR \ln\left(\frac{V_f - b}{V_i - b}\right). \tag{5.82}$$

Then, we observed that this expression differs from the one valid for an ideal gas only for the volumes being diminished by the covolume rather than the total volumes. We observed that, in both cases, we deal with the volume available to the molecules. Finally, we noticed that the term a/V^2, representing the molecular interactions in the van der Waals equation, does not appear. We can now understand the reason for this. The thermodynamic probability of a macrostate, at a given temperature, depends only on the volume available to the molecules, not on the presence or not of interactions amongst them.

Problems

5.1. Evaluate the number of molecules in a glass of water.

5.2. What is the mass of a mole of electrons?

5.3. How can one find the molecule number density n_p of a substance of known density ρ and molecular mass μ?

5.4. How many molecules are there in a cubic meter of air at 0 °C at one atmosphere pressure (take it in round numbers 100 kPa)?

5.5. A gas flow is made of equal molecules of mass *m*, with a numeric density *N*/*V*, that move at the same velocity **v**. The flow hits a wall of unit area at the angle θ with the normal. Evaluate: (a) the number of collisions per unit time; (b) the pressure exerted by the gas.

5.6. The average kinetic energy of the molecules of an ideal monoatomic gas is $\langle k_B T \rangle = 6 \times 10^{-21}$ J, the pressure is 100 kPa. Find: (a) temperature. (b) the molecules' numerical density.

5.7. Consider the probability density functions in Fig. 5.15 and determine, in both cases, the value of *A* to have *f*(*x*) normalized and the root mean square value of *x*.

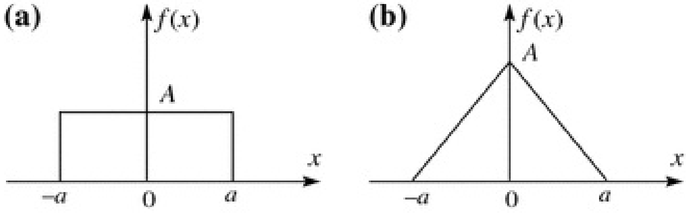


Fig. 5.15 Two probability density functions

5.8. Calculate the most probable and the root mean square velocities of the N2 molecules at 20 °C.

5.9. Calculate the root mean square rotation angular velocity of N2 at 20°C. The distance between nuclei is $a = 0.37$ nm. Use classical physics.

5.10. An ideal gas is in a conservative and central force field. The potential energy of a molecule at the distance r from the center is $U_p(r)$. Write the expression of the number dN of molecules per unit volume between r and $r + dr$ as a function of r, knowing that the numerical density at r_0 is n_0.

5.11. In a gas in thermal equilibrium, what are the fractions of the molecules with kinetic energy larger than (a) the mean, and (b) three times the mean?

5.12. Consider the experiment by Perrin on the height distribution of equal spherical particles suspended in water. The density of his particles was $\rho_p = 1250 \text{ kg m}^{-3}$ and their radius $r = 0.21$ µm. The temperature of the suspension was $T = 293$ K. Looking through the microscope, he counted the number of particles per unit volume at the heights z_1 and z_2 separated by $z_2 - z_1 = 30$ µm. He found that, at the higher level, the number was smaller than at the lower by a factor of 2.1. Find NA.

5.13. Considering an isothermal atmosphere at $T = 20$ °C with a sea-level pressure p_0, find the pressure at altitudes: (a) $h = 5000$ m (about the top of Mont Blanc), (b) $h = 9000$ m (about Mount Everest); (c) in a mine 2000 m deep. Take the molar mass of air equal to 29.

5.14. A container divided into two equal parts contains a gas of 6 molecules. Find: (a) the thermodynamic Γ and mathematic P probabilities of the states with the following numbers of molecules in the two halves: (0, 6); (1, 5); (2, 4), (3, 3); (b) the probabilities of finding the system in the states (2, 4), (3, 3), (4, 2).

5.15. A thermodynamic system passes from one state to another, having twice the thermodynamic probability. What is the entropy variation?

5.16. Calculate the entropy variation ΔS and the ratio of the thermodynamic probabilities Γ_v / Γ_l for the evaporation of a liter of water at $T = 373$ K. The vaporization heat is $Q_{lv} = 2256$ kJ/kg.

5.17. A milligram of water at $T = 293$ K passes into a new thermodynamic state with a thermodynamic probability 1000 times larger. (a) Calculate the entropy variation ΔS. (b) Can we detect the variation measuring ΔT?

5.18. What is the behavior of the thermodynamic probability of a system making an adiabatic process?

Footnotes

[1] R.C. Miller and P. Kausch, Phys. Rev. **99** (1955) p. 1314.

Alessandro Bettini, A Course in Classical Physics 2—Fluids and Thermodynamics, Undergraduate Lecture Notes in Physics, DOI 10.1007/978-3-319-30686-5_6

6. Transport Phenomena

Alessandro Bettini[1] ✉

(1) Dipartimento di Fisica e Astronomia, Università di Padova, Padova, Italy

✉ **Alessandro Bettini**
Email: alessandro.bettini@pd.infn.it

Up to now, we have mainly studied the properties of thermodynamic systems in equilibrium. In this chapter, we shall consider important examples of processes that, starting from non-equilibrium, lead to the establishment of equilibrium conditions. These are called kinetic processes and are spontaneous and, obviously, irreversible.

Suppose, for example, we have a pot of water in thermal equilibrium with the environment. We insert an electric immersion heater in the middle of the pot, and after a few moments, we take it back out. The temperature in the central zone is now higher than in the peripheral parts. The system is no longer in thermal equilibrium. It will evolve towards a new equilibrium state spontaneously. Thermal energy, or heat, will diffuse from the hotter to the cooler parts until the temperature is uniform everywhere. This is a thermal energy diffusion process (in absence of convective motions).

As another example, let us drop a sugar cube into a hot cup of tea. Just after the sugar has completely melted, the sugar concentration in the tea is much higher in the region of tea where the cube was dropped, decreasing with increasing distance. The system is not in equilibrium. The sugar diffuses all through the liquid, until its concentration is uniform, i.e., equal throughout the teapot. Similarly, if we introduce a quantity of gas into one portion of a room, for example, by spraying a perfume, with time, its molecules will diffuse throughout all the air in the room.

These are simple examples of two transport phenomena, the propagation by conduction of heat, or of thermal energy, and the diffusion. In the first case, the transported physical quantity is kinetic energy of molecules, in the second, the molecules themselves (of sugar or perfume). We shall also consider a third case, the transport of a vector quantity, the linear momentum (or, in an equivalent manner, velocity). Let us consider an example.

Suppose we can work on a canal in which water flows with velocity constant in time and uniform, namely equal at all the points of the fluid. The speed is not high, a few centimeters per second. We place a wood tablet on the surface of the liquid and give it a different velocity, say higher, from that of the liquid. After a while, we raise the tablet. In a zone near where the tablet was, water has a velocity, and a linear momentum, different from that of the rest of the liquid. In the previous example, we had injected thermal energy or an extraneous gas; now we inject linear momentum. Also in this case, the injected quantity, which is the *excess* of linear momentum, gradually diffuses throughout the system until the linear momentum is uniform. The phenomenon, as we shall see, is directly connected to viscosity.

The three diffusion phenomena that we have just exemplified will be treated in the first three sections. All of them are due to the chaotic motion of the molecules and their collisions. In Sect. 6.4, we shall define two important correlated quantities needed for a description of collisions, the cross-section, which measures the probability of a collision, and the mean free path, the mean distance crossed by a molecule between two collisions. Having these concepts, we shall come back, in Sect. 6.5, to a further discussion of the transport phenomena and see their similarity when considered at the molecular level.

6.1 Heat Conduction

Consider a body isolated from heat sources, with temperature different from point to point. The system spontaneously evolves towards a state of uniform temperature. Heat flows, or is transferred, from the hotter to the colder regions until temperature is uniform, independent of position. If two (or more) regions of the body are in contact with heat sources at different temperatures, heat flows from the higher temperature to the lower temperature source. The processes are different, but the underlying physics is the same. In both cases, we have a thermal energy transfer , which is also called a heat transfer .

Thermal energy can be transferred in three different ways: by conduction , by convection and by radiation

In a solid, heat transfer is (mainly) by conduction. If we put one end of an iron rod into a flame and hold the other in our hand, after a while, we feel it burning. If we then move the bar to a table, one end will still be hotter than the other. But shortly after, the temperature will be uniform. The thermal heat has been transferred, in both cases, from the higher to the lower temperature regions by conduction. The underlying physics phenomenon is the diffusion of the internal kinetic energy.

If the system is a fluid, heat is transferred by conduction and by convection. Convection takes place because regions of different temperature also have different density, and are not in mechanical equilibrium under their weights. As a consequence, motions arise inside the system, in which hotter and colder liquid masses mix with one another. The thermal energy travels together with the fluid masses in their macroscopic motions, as opposed to molecular motions. This process is called convection, from the Latin words for "taking with". The importance of the convective motion depends on the direction of temperature gradient relative to the vertical. Consider, for example, a pot of water on a stove. The water elements on the bottom of the pot have the highest temperature and the smallest density. Consequently, they rise toward the surface while colder elements descend. The mixing action is very efficient. Contrastingly, if the heat source is above the container, those portions of liquid warming first, and decreasing in density, are already above the colder ones and do not move much.

Thermal energy transfer also takes place in a vacuum. Consider a box in which we have created a vacuum containing two metallic plates facing one another at different temperatures. As a matter of fact, the space between the plates is empty of matter but does contain something. It contains the electromagnetic radiation emitted by the hot plate and (partially) absorbed by the cold one. The process is the radiation transfer.

Having mentioned, for the sake of completeness, the three thermal energy transfer mechanisms, we shall now study only the first one, conduction. As we have already mentioned, thermal conduction means the diffusion of the internal kinetic energy.

Heat conduction phenomena, in general, take place in three-dimensional media, in which the temperature varies from point to point and as a function of time. We shall limit our discussion to the simplest situation, in which temperature depends, beyond time, on one coordinate only. Consider a plate of surface large enough for any edge effect to be negligible. Suppose the temperatures on each side of the plate to be uniform. Let them be T_1 and T_2, with $T_1 > T_2$. In general, the temperatures may vary over time.

We take a reference frame with the origin and the y and z axes on the face at higher temperature, and the *x*-axis toward the second face as in Fig. 6.1. The temperature at the points of the plate is a function of *x* and time, say *T*(*x*, *t*). Consider a surface element *dS* at a generic point of the system, taken normal to the direction of the heat flow, which is the *x*-axis in our case.

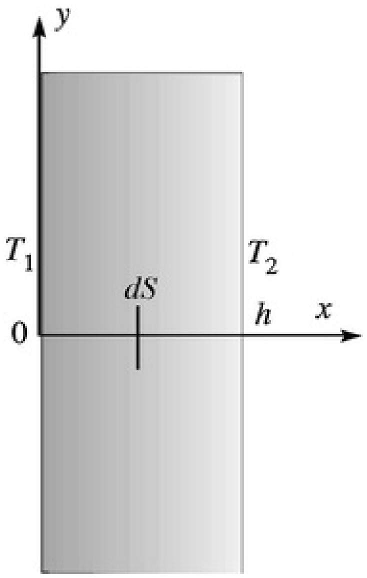


Fig. 6.1 Heat transfer by conduction in one dimension

Let *dQ* be the heat flowing through *dS* in the elementary time interval *dt*. It is evident that *dQ* is proportional to both *dS* and *dt*. The proportionality coefficient depends on the material. In 1807 (published in 1822), Jean Baptiste Joseph Fourier (France, 1768–1830) theoretically found the following law that was subsequently experimentally verified. It is called Fourier’s law

$$dQ = -\kappa \frac{\partial T(x,t)}{\partial x} dS\, dt, \tag{6.1}$$

where *κ* is the *thermal conductivity* of the substance. Thermal conductivities of different substances range over several orders of magnitude with continuity. At the two extremes, as already mentioned in Sect. 2.4, we have the good thermal conductors and the good thermal insulators . We shall give some values at the end of the section. We see that *dQ* is also proportional to the space derivative, namely the gradient, of temperature. The more quickly the temperature varies along the *x*-axis, the larger the heat transmission. The negative sign means that the heat flow is in the direction of decreasing temperature.

The heat crossing a given surface taken perpendicularly to the flow in the unit time is called *heat flux* or *thermal flux* . The *heat flux density* , or *thermal* *flux density* , is the flux per unit area. We shall call it Φ_Q . For Eq. (6.1), it is obviously

$$\Phi_Q = -\kappa \frac{\partial T(x, t)}{\partial x}. \tag{6.2}$$

Heat conductivity is then the proportionality constant between heat flux density and temperature gradient. The physical dimensions of heat flux density are

$$[\Phi_Q] = \mathrm{Js^{-1}m^{-2}} = \mathrm{Wm^{-2}} \tag{6.3}$$

and those of thermal conductivity are

$$[\kappa] = \mathrm{Js^{-1}m^{-1}K^{-1}} = \mathrm{Wm^{-1}K^{-1}}. \tag{6.4}$$

Consider a body initially at non-uniform temperature in absence of external heat sources. The temperature tends to become uniform in any case, but the time needed for that depends on thermal conductivity. It will be shorter the higher the conductivity. In addition, the temperature variation for a given received heat is inversely proportional to the specific heat and the density. Such is also the temperature leveling speed. Let us see this in formulae.

Let us take two surfaces dS_1 and dS_2 of the same area dS perpendicular to the x-axis at the same y and z. One of them is at x, the other at $x + dx$, as in Fig. 6.2. Consider the element between the two surfaces (in evidence in the figure). It receives and gives off heat through the two faces. The heat through the first face (in the positive x direction) in the time interval dt is $dQ_2 = -\kappa\,(\partial T/\partial x)_x\,dS\,dt$, where we have indicated with $(\partial T/\partial x)_x$ the partial derivative of the temperature with respect to x at the position x. Similarly, the heat through the second face (always in the positive x direction) in dt is $dQ_2 = -\kappa\,(\partial T/\partial x)_{x+dx}\,dS\,dt$. The first heat enters, the second exits; hence, the heat absorbed by the element is

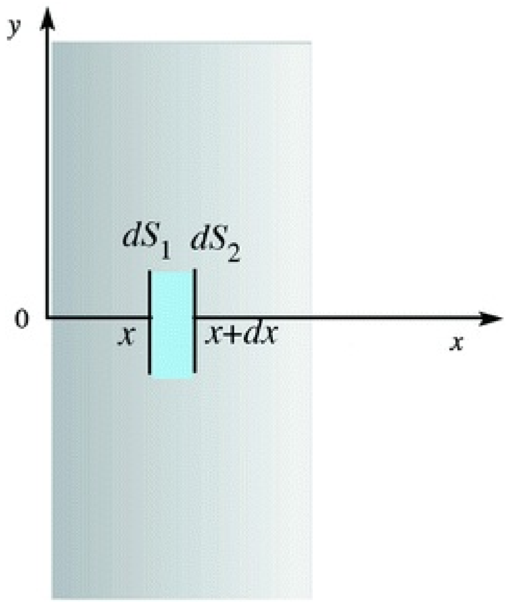


Fig. 6.2 Heat transfer through an infinitesimal element

$$dQ = dQ_1 - dQ_2 = \kappa \left[\left(\frac{\partial T}{\partial x} \right)_{x+dx} - \left(\frac{\partial T}{\partial x} \right)_x \right] dS\ dt = \kappa \frac{\partial^2 T}{\partial x^2} dx\, dS\ dt.$$

Upon absorption of *dQ*, the temperature of the elements increases by, say, *dT*. If ρ is the density and c_p the specific heat (we operate at constant pressure), we write

$$dQ = c_p \rho\ dS\ dx\, dT.$$

Putting together the two equations just found, we get

$$\kappa \frac{\partial^2 T}{\partial x^2} dt = c_p \rho\ dT,$$

which, collecting the constants together, can be written as

$$\frac{\partial T}{\partial t} = \chi \frac{\partial^2 T}{\partial x^2}, \tag{6.5}$$

where the constant χ is

$$\chi = \frac{\kappa}{c_p \rho}. \tag{6.6}$$

The differential equation Eq. (6.5) tells us that the rate of change in time of temperature is proportional to its second spatial derivative, namely the rate of change with position of the temperature gradient. The proportionality constant χ tells how quick the process of temperature equalization is. This is called *thermal diffusivity* and is equal (we had anticipated a proportionality) to the thermal conductivity divided by the product of specific heat and density. The measurement units are

$$[\chi] = \mathrm{m}^2\mathrm{s}^{-1}. \tag{6.7}$$

We shall come back to its physical meaning in Sect. 6.3. We simply anticipate here that the physical mechanism of heat transmission is the following. If in a zone of the body, the temperature is higher than in the adjacent zones, the molecules in that zone have larger on average kinetic energies than those in the other. Part of this kinetic energy in excess is transferred to the "colder" molecules through collisions. The excess kinetic energy diffuses in this way throughout the body. Equation (6.5) quantitatively describes how the molecular mean kinetic energy diffusion mechanism takes place.

Table 6.1 reports the values of thermal conductivity at ambient temperature for several substances. Metals have typically high conductivity. (349 $\mathrm{Wm^{-1}K^{-1}}$ for Cu, 209 $\mathrm{Wm^{-1}K^{-1}}$ for Al, 58 $\mathrm{Wm^{-1}K^{-1}}$ for Fe). Rocks have generally low conductivity, typically a few $\mathrm{Wm^{-1}K^{-1}}$. A good thermal insulator like polyurethane foam has a thermal conductivity of 0.026 $\mathrm{Wm^{-1}K^{-1}}$, which is four orders of magnitude smaller than aluminum.

Table 6.1 Thermal conductivity of different materials at room temperature

Material	**κ ($\mathrm{Wm^{-1}K^{-1}}$)**	**Material**	**κ ($\mathrm{Wm^{-1}K^{-1}}$)**	**Material**	**κ ($\mathrm{Wm^{-1}K^{-1}}$)**
Metals		***Various***		***Insulators***	
Aluminum	209	Slate	1.98	Asbestos	0.2
Iron	58	Ceramic	1.1	Rockwool	0.04
Brass	99	Granite	3.14	Perlite	0.04
Copper	349	Plasters	0.8-1.5	Polystyrene	0.035
Zinc	110	Bricks	0.8	Polyurethane	0.026
		Terracotta	0.9		
		Dry ground	0.8		
		Wet ground	2.3		

The physical reason for the high thermal conductivity of metals is that heat in them is transmitted through the thermal motion of electrons, rather than of molecules, as in the largest fraction of other materials. As we have already mentioned, metals are aggregates of microcrystals made of ions of the element. In a microcrystal, each atom has lost one or two electrons, depending on the metal. These electrons, called conduction electrons , are free to move, like a gas, inside the crystal. They are responsible for the heat conduction. Their kinetic energy is proportional to the absolute temperature. At a given temperature, the mean kinetic energy of the electron gas is equal to that of the molecules of a "normal" monoatomic gas. Consequently, the ratio between the root mean square velocities in the two cases is equal to the reciprocal of the square root of the ratio of the masses. The molecular masses are typically four orders of magnitude larger than the electron mass. As a consequence, the root mean square speed of electrons is two orders of magnitude larger than that of molecules at the same temperature. This is why the heat transmission is much faster.

Consider again the plate in Fig. 6.1 in the particular case in which the temperatures of the two sides T_1 and T_2($T_1 > T_2$) are constant in time, namely in stationary conditions . This is the case, for example, for a wall in a house. In winter, the temperature outside is lower and heat flows from inside to outside. Equation (6.5) simplifies it as

$$\frac{\partial^2 T}{\partial x^2} = 0. \tag{6.8}$$

Integrating it twice, we find the solution $T = Ax + B$, where A and B are the integration constants. We determine the constants by imposing the boundary conditions. At $x = 0$, it must be $T = T_1$, and at $x = h$, it must be $T = T_2$. We immediately find $A = (T_2 - T_1)/h$ and $B = T_1$, and we get

$$T(x) = T_1 - (T_1 - T_2)\,x/h. \tag{6.9}$$

The solution says that, under stationary conditions, the temperature decreases linearly from one face to the other.

The heat transmitted per unit time through a surface area S is obtained from Eq. (6.1) by integration. We get

$$\frac{dQ}{dt} = \kappa\,(T_1 - T_2)\,S/h. \tag{6.10}$$

The heat transmitted through the wall per second is proportional to the temperature difference between its two sides, directly to the surface and inversely to the thickness of the wall. The proportionality constant is the conductivity.

6.2 Diffusion

Consider now a solution, for example, sugar (solute) in water (solvent). If the concentration is not uniform, molecules of solute tend, on average, to move, in their disordered thermal motion, from sites of higher to those of lower concentration. Equilibrium is reached when concentration is uniform. The phenomenon is called (once more) *diffusion* . As a matter of fact, the phenomenon is similar to the free expansion of the gas we discussed in Chap. 5.

It is the transition between macrostates of smaller to higher thermodynamic probabilities. The difference is that diffusion is a much slower process.

Also now, we consider, for simplicity, a one-dimensional situation. We suppose that the concentration, which we call c, depends on time and on one coordinate only, say x. Let Φ_c be the solute mass passing in one second through a one meter square surface normal to the x-axis, called the mass flux density. The flux direction is from higher to lower concentration, namely the direction opposite to the derivative of concentration $\partial c/\partial x$, which is the concentration gradient. Clearly, there is no flow for uniform concentration, namely $\partial c/\partial x = 0$.

The relation between concentration flux density and gradient is the same as in the case of thermal energy, Eq. (6.2), namely

$$\Phi_c = -D\frac{\partial c\ (x,t)}{\partial x}, \tag{6.11}$$

where D is the *diffusion coefficient* , which depends on both solvent and solute and on temperature. The flowing solute quantity can be expressed as a mass, as a number of molecules or as a number of moles. The choices are obviously equivalent, but the physical dimensions of the diffusion coefficient depend on them. If we measure the concentration in moles per cubic meter, the flux density will be in mol $\mathrm{m}^{-2}\mathrm{s}^{-1}$ and the dimensions of the diffusion coefficient D are $\mathrm{m}^2\mathrm{s}^{-1}$.

What we have stated is also valid in the case of the diffusion of a gas in another gas and also of a gas in itself (different regions having different densities), which is called self-diffusion .

The differential equation describing the evolution of concentration in the solution is completely analogous to the case of thermal energy diffusion (with concentration in place of temperature) that we found in Eq. (6.5). What we said there is valid here, with the obvious changes, and we shall not discuss it any further. We observe, however, that, as in the case of thermal energy diffusion, the equation is valid only if the fluid is in mechanical equilibrium and the equalization of concentration is due to the thermal disordered motion of the molecules only. In other words, there are no convective motions.

Consider, for example, two liquids, water and alcohol, the first with a larger density than the second. We put alcohol in a container and then we slowly inject water with a small tube at the bottom of the container. If we are careful enough, we can produce a layer of water under the alcohol. The convective motion will be practically absent and water molecules will mix with alcohol molecules through their collisions, namely by diffusion. If we proceed in the opposite way, filling the container with water and then injecting the (lighter) alcohol at the bottom, the liquids will mix by convection.

The reason for the similarity of the heat conduction and matter diffusion is very clear when looked at from the microscopic point of view. In the first case, an excess of average kinetic energy of molecules, or of conduction electrons, diffuses to reach the equalization of kinetic energy. In the second case, the excess of solute molecular density in a region diffuses to reach the equalization of molecular density. Consequently, the differential equations ruling the two processes Eqs. (6.2) and (6.11) are equal.

There is, however, a small difference. While in Eq. (6.11), the flux and the gradient are both of the concentration, namely the same physical quantity, in Eq. (6.2), flux and gradient are of two different quantities. However, we can easily have the same quantity on both sides in the latter case as well. It is just a matter of dividing both sides by ρc_p. The resulting quantity on the left-hand side, $\Phi_Q/\rho c_p$, is the flux density of mean molecule kinetic energy, which we can call Φ_T. We have

$$\Phi_T = -\chi \frac{\partial T\ (x,t)}{\partial x}. \tag{6.12}$$

So, we understand why thermal diffusivity and diffusion coefficient have the same physical dimensions (m^2s^{-1}). The values of the two quantities are also similar for gases. For example, the thermal diffusivity of air at 0 °C is $\chi = 1.9 \times 10^{-5}\ m^2s^{-1}$, while the diffusion coefficient of water vapor in air is $D = 2.3 \times 10^{-5}\ m^2s^{-1}$.

The diffusion coefficients in air of a few gases at different temperatures are reported in Table 6.2.

Table 6.2 Diffusion coefficients in air

Substance	***T* (°C)**	***D* ($10^{-5}\ m^2s^{-1}$)**
Alcohol (vapor)	40	1.37
CO_2	0	1.39
CS_2	20	1.02
Ether (vapor)	20	0.89
Hydrogen	0	6.34
Oxygen	0	1.78
Water (vapor)	8	2.39

In liquids, diffusion is much slower than in gases. For example, the diffusion coefficient of sugar in water is 0.3×10^{-7} m^2s^{-1}, while that of common salt (NaCl) in water is 1.1×10^{-7} m^2s^{-1}. These are two orders of magnitude smaller than for gases.

6.3 Viscosity

In Chap. 1, we studied the viscosity phenomena from the macroscopic point of view. There, we considered a liquid (or a gas) contained between two solid plane horizontal sheets at a distance h. We considered the higher sheet moving at constant speed υ_0, and the lower one at rest. The liquid layers in contact with each sheet remain adhered to them and have their velocities. Inside the liquid, the velocity decreases linearly with increasing depth from υ_0 to 0 at the bottom, as shown in Fig. 6.3, which is equal to Fig. 1.11.

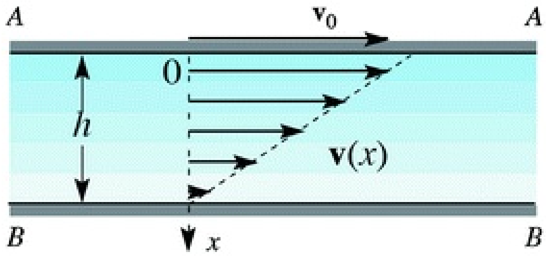


Fig. 6.3 Laminar flow between two surfaces

We have chosen the x-axis vertical downward with origin on the upper surface. Forces due to viscosity act in the direction opposite to the relative motion, between fluid layers. The shear stress , namely the force per unit surface, is given by Eq. (1.23), which we reproduce here for convenience:

$$\tau = \frac{dF^{(S)}}{dS} = \eta \frac{\partial \upsilon}{dn}. \tag{6.13}$$

We shall now see the physical meaning of viscosity . The situation we are considering is particularly simple because the velocity vector has the same direction everywhere. Only its absolute value varies and does so as a function of one coordinate only, $\upsilon = \upsilon(x)$. Indeed, as we mentioned, the system is a fluid layer between two planes, ideally of infinite extension, at the distance h between them.

The relevant physical quantities of the problem are velocity and linear momentum of the fluid elements. As we are dealing with a continuous system, we must talk of *linear momentum density* , which is the linear momentum of an elementary volume divided by the volume. Let us consider a fluid element of volume dV with mass $dm = \rho dV$. If $\mathbf{v}$ is its velocity, its momentum is $d\mathbf{p} = \mathbf{v}\, dm$. The momentum density is then

$$\frac{d\mathbf{p}}{dV} = \mathbf{v}\frac{dm}{dV} = \rho\mathbf{v}.$$

In our simple case, the momentum density varies only as a function of the height and does so linearly in absolute value between $\rho \upsilon_0$ and 0, while the direction is a constant. The situation is then completely analogous to the one considered in Sect. 6.2, where we had a layer of matter between two planes at different temperatures. In both cases, the regime is stationary. In that case, we had a thermal energy flow from the hot to the cold plate through all the sections of the layer. Similarly, in this case, we have a *flow of momentum density* from the face at higher speed to the one at lower speed.

We now want to find an expression of the linear momentum flux , which we indicate with Φ_p, and which is, by definition, the linear momentum going through the normal surface unit in the unit time. In our problem, momentum flow is in the direction of the x-axis in Fig. 6.3. Let us then take the surface element dS of Eq. (6.13) perpendicular to the x-axis. The linear momentum going through dS in a second is $\Phi_p\, dS$. This is the decrease of momentum per unit time above dS and its increase below dS. But the rates of change of momentum are equal to the forces. The forces are the shear stresses with which the fluid layers act on one another. In conclusion, the momentum flux is equal to the share stress

$$\Phi_p = \tau. \tag{6.14}$$

Notice that the regime of our fluid is laminar. If there were vortices, their effect would have been to quickly mix the fluid elements, equalizing the momentum density. The latter mechanism is similar to the heat and mass transport by convection. We shall limit our discussion to the transport, of momentum in this case, through molecular collisions .

Having found Eq. (6.14), we can write Eq. (6.13), which we already encountered in hydrodynamics, as

$$\Phi_p = -\eta\frac{\partial v}{dn}, \tag{6.15}$$

where the minus sign means, as usual, that the direction of the flow is opposite to that of the velocity gradient. We can say that the *dynamic viscosity* η measures the rapidity of the transmission of linear momentum from one place to the other.

Equation (6.15) is, in fact, very similar to Eqs. (6.11) and (6.12). It is similar but not identical, because in the latter equations, we have, on the right-hand side, the rate of change and, on the left-hand side, the flux density of the same physical quantity. We would like to have, on the left-hand side of Eq. (6.13), the velocity flux density Φ_{υ}. Velocity is momentum divided by the mass. As we are dealing with the unit volume, we should divide by the density ρ. The velocity flux density is $\Phi_{\upsilon} = \Phi_{p}/\rho$. It is now convenient using the *kinematic viscosity* to find

$$\nu = \eta/\rho \tag{6.16}$$

and Eq. (6.15) becomes

$$\Phi_v = -\nu \frac{\partial v}{dn}. \tag{6.17}$$

Equation (6.17) is now identical to Eqs. (6.11) and (6.12). The proportionality coefficient between flux density and gradient is the kinematic viscosity v; and it determines the speed at which the velocity field, left alone, becomes uniform. The physical dimensions of the kinematic viscosity are

$$[\nu] = \mathrm{m^2 s^{-1}}, \tag{6.18}$$

which are the same as those of the diffusion coefficient D and of the thermal diffusivity χ. The values of the kinematic viscosity are also similar to those of D and χ. Let us take back the example of air we have seen in the preceding section (thermal diffusion coefficient at 0 °C and normal pressure $\chi = 1.9 \times 10^{-5}\ \mathrm{m^2 s^{-1}}$; diffusion coefficient of water vapor in air is $D = 2.3 \times 10^{-5}\ \mathrm{m^2 s^{-1}}$). The kinematic viscosity of air is $v = 1.5 \times 10^{-5}\ \mathrm{m^2 s^{-1}}$. As one sees, the values are very similar.

In conclusion, a system in which temperature, concentration or velocity are not uniform is not in a thermodynamic equilibrium state. The system, left alone, will spontaneously evolve towards equilibrium. The non-uniform quantity will tend to become uniform. In all three cases, the process takes place by redistributing amongst the molecules the physical quantity of the game. The redistribution takes place through molecular (or conduction electrons) collisions. These are called *molecular transport phenomena* . Notice that even in an ideal gas, the molecules collide and redistribution takes place. The role of the weak intermolecular attraction forces present in the real and absent in the ideal gases is completely secondary. Consequently, even ideal gases are viscous, as we have just seen in the example of air. Indeed, the kinematic viscosity of air is larger than that of water, which is $v = 10^{-6}\ \mathrm{m^2 s^{-1}}$. The opposite is true for the dynamic viscosities, which are $\eta = 10^{-3}$ Pa s for water and $\eta = 1.8 \times 10^{-5}$ Pa s for air.

The difference between the present case and those of diffusion and thermal conduction is due to the fact that temperature and concentration are scalar quantities, while velocity and momentum are vectors. As a matter of fact, we obtained the simple Eq. (6.17) considering the simple geometry in which the velocity has the same direction throughout the fluid. The equation is valid only in this situation, not in more general ones. We can easily see that Eq. (6.17) cannot work if **v** has different directions at different points, considering a liquid in a cylindrical container uniformly rotating around the axis of the cylinder, as in the viscometer we considered in Sect. 1.6. The velocity, and momentum, of the fluid elements increase with increasing distance from the axis. That fact notwithstanding, there is no momentum flux, or, in other words, there is no action by an internal friction force. The rotation of the liquid does not take the system out of its thermodynamic equilibrium. It might, in fact, continue forever without any equalization of velocities.

6.4 Mean Free Path

The phenomena we studied in the last section depend on the interactions between molecules. In a gas at STP, the average distance between molecules is much larger than their diameter. Consequently, for the largest fraction of time, they interact only very weakly, or not at all if the gas is considered to be ideal. They do interact during the short time intervals during which they are close enough to collide. In the aggregate phases, liquid and solid, the situation is completely different. Molecules always interact and it is impossible to talk of single collisions. In this and the next section, we shall further study some properties of gases. Two important quantities are the collision cross-section and the mean free path between collisions.

Let us start by analyzing what a collision is. We chose a reference frame in which one molecule is at rest. When an incoming molecule is still farther than the molecular action radius, say a few nanometers, it moves along a straight line. The distance of this line from the center of the target is called the *impact parameter* , b, as shown in Fig. 6.5. In other words, and from hereon in any reference frame, b is the smallest distance between the two colliding molecules that could be attained if they moved without interacting in straight lines with the velocities they had before the collision.

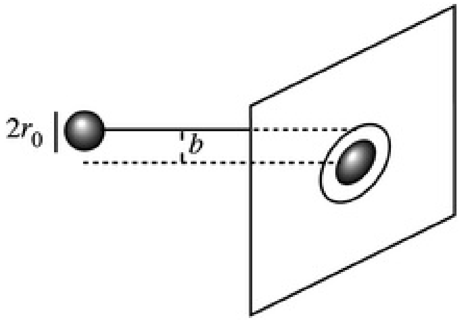


Fig. 6.4 Impact parameter and cross-section for rigid spheres

Let us consider, in an extremely rough approximation, a molecule as a rigid sphere of radius r_0. Then, the collision happens if the centers of the two molecules approach one another by at least twice the radius, namely if $b \leq 2r_0$. Let us consider a plane perpendicular to the direction of the incident molecule and draw a circle about the center of the standing one of radius $2r_0$. The collision happens if the incident molecule's direction before interacting crosses this area. This circle is called the *cross-section* , which we indicate with σ.

This is not, however, a correct description of the collision. Two molecules also interact with one another at distances between centers larger than their diameter, up to the molecular action radius r_a (see Sect. 4.1). Up to this distance, the van der Waals force changes the momenta of the two molecules. This is also considered a collision. Let us consider, as an example, an asteroid approaching a planet. When it is far away, its trajectory is almost straight, but getting close to the planet, it will describe an arc of hyperbola, to finally move away in a direction different from the original one. This, if between molecules, is also a collision. We consider any process in which the impact parameter is small enough that the velocities of the two molecules change significantly to be a collision.

To be more precise, we recall that we are dealing with a statistical process. Thus, we must think not of one, but of many incident molecules moving in the same direction but at different impact parameters. Their flux density, Φ, is their number crossing the unit area normal to their motion per unit time. The cross-section is defined as the ratio between the number of collisions on the target molecule per unit time, R, and the flux density

$$\sigma = R/\Phi. \qquad (6.19)$$

In other words, the cross-section is the collision rate on one molecule per unit incident flux density (one particle per square meter per second). The cross-section is not a geometrical quantity, as its name suggests, but rather a measure of collision probability, and it depends on the complete behavior of the van der Waals force, not only on r_0. However, the value of the cross-section obtained considering the molecules as rigid spheres gives a reasonable estimate of the order of magnitude. It is

$$\sigma = 4\pi r_0^2, \tag{6.20}$$

which is four times the area of the section of a molecule. In Sect. 4.1, we mentioned that the radiuses of the simpler diatomic molecules, like N_2 and O_2, are on the order of one tenth of a nanometer, $r_0 \approx 0.1$ nm. In correspondence, Eq. (6.20) gives $\sigma \approx 10^{-19}\ \mathrm{m}^2$. This is the correct order of magnitude. For example, the cross-section in air of those molecules is $\sigma = 5 \times 10^{-19}\ \mathrm{m}^2$.

Collisions between molecules are completely random processes. In particular, the distance between one collision and the next is a random variable. Its mean value is called the *mean free path* (between collisions). We shall indicate it with l.

Clearly, the larger the cross-section, the smaller the mean free path. To find their relation, consider a one meter long path of a molecule. Imagine the molecule sweeping a cylinder of section σ. It hits all the molecules in this cylinder. The volume of the cylinder is just σ because we took its height to be unitary. If n_p is the number density, in the cylinder, there are $n_p \sigma$ molecules. Our molecule then hits $n_p \sigma$ molecules per meter along its path, or, we can say, one collision every $1/(n_p \sigma)$ meters. This is the mean free path:

$$l = \frac{1}{n_p \sigma}. \tag{6.21}$$

In other words, the mean free path is inversely proportional to the cross-section and to the molecule number density, and consequently to the gas pressure. It does not appear to depend on temperature, namely on molecular velocities. This is true only if the cross-section is constant, independent of velocity, as we implicitly assumed. If the temperature is high, the molecular speeds are large and the velocity change in the collision is substantially due to the short distance repulsive force. If, however, the temperature is low, and consequently the velocities are smaller, the time during which two molecules are nearby is larger. The collision takes more time and molecules at larger impact parameters can deviate. We conclude that cross-sections in a gas should increase with decreasing temperature. This indeed happens, but variations are mild. For example, the collision cross-section between oxygen molecules vary by 30 % between 100 °C and −100 °C.

Let us now estimate the order of magnitude of the mean free path for the N_2 and O_2 gases at STP, for which we found the cross-section above. Recalling that a mole of gas at STP has the volume of 22.4 L, we find that the number density is $n_p = 3 \times 10^{25}\ \mathrm{m}^{-3}$. With $\sigma \approx 5 \times 10^{-19}\ \mathrm{m}^2$, the mean free path is $l \approx 70$ nm.

In conclusion, at the molecular level, there are three relevant quantities with the dimension of a length, the radius r_0, the mean distance between them d, and the mean free path. Typical values at STP are

$$r_0 \cong 0.1\ \mathrm{nm}, \quad d \cong 3\ \mathrm{nm}, \quad l \cong 70\ \mathrm{nm}. \tag{6.22}$$

The mean square velocity of, for example, air molecules is $\upsilon \simeq 500$ km/s, and consequently the mean time between collisions is $\tau \approx \upsilon\, l = 0.2\ \mathrm{ns}$.

6.5 Transport Properties in Gases

In Sect. 6.3, we saw how the diffusion coefficient, the thermal diffusivity and the kinematic viscosity have similar underlining physical processes, the same dimensions and similar values. We shall now exploit the concept of the mean free path to understand the reasons for that, to estimate their values approximately and to clarify their dependence on the gas state. We shall give order of magnitude evaluation only, focusing on the physical concepts and neglecting details.

Let us start with diffusion . Consider two gases in a container. The pressure is uniform, but the composition, namely the ratio between the densities of the gases, depends on position. For simplicity, we assume that the dependence is on one coordinate only. Let this be x. Let the two number densities be $n_{p1}(x)$ and $n_{p2}(x)$. The state is not of equilibrium. The densities tend to become uniform moving, on average, in the x direction. This is a diffusion process.

Consider the molecules of one gas, the first one, for example. Suppose that its density decreases with increasing x, say from left to right. Consider the plane perpendicular to the x-axis at the coordinate x. An average number of molecules crosses the unit area surface of this plane in a second from left to right and another number from right to left. The difference between these numbers is the diffusion flux density Φ_c. Let us evaluate it in order of magnitude.

In the case we are considering, there are more molecules crossing the plane in the positive x direction than in the opposite one, because there are more molecules per unit volume on its left that on its right. If all the molecules were to have the same velocity υ, the molecules crossing a section in the unit time would be those contained in a volume having that section as base and υ as height. As the velocities are different, υ is a suitable average, which we shall take as the root mean square υ_{rms}. We need not to be very precise, because we are evaluating orders of magnitude. The number of molecules per unit volume varies with x and we must decide where to consider it. It looks reasonable to take it at a distance on the left of the surface equal to the mean free path, because it is there that, on average, the last collision took place. The flux from left to right is then $\upsilon_{\text{rms}}\, n_{p1}(x-l)$. Similarly, the flux of molecules from right to left at $x + l$ is $\upsilon_{\text{rms}}\, n_{p1}(x + l)$.

In conclusion, the net flux density is $\Phi_c \cong \upsilon_{\text{rms}} n_{p1}(x-l) - \upsilon_{\text{rms}} n_{p1}(x+l)$. Considering that the mean free path is very small, we can write

$$\Phi_c \cong -2l\,\upsilon_{\text{rms}}\frac{dn_{p1}}{dx}.$$

Comparing it with Eq. (6.11), we can write it in the form

$$\Phi_c \cong -D\frac{dn_{p1}}{dx}.$$

We immediately see that

$$D \cong l\,\upsilon_{\text{rms}}, \tag{6.23}$$

where we have neglected the factor 2, as we are considering orders of magnitude. In other words, the diffusion coefficient is, in order of magnitude, equal to the product of the mean free path and root mean square velocity. For example, for air at STP, we have estimated a mean free path, in round numbers, of $l \approx 100$ nm, and $\upsilon_{\text{rms}} \approx 500$ m/s. With Eq. (6.23), we evaluate $D \approx 5 \times 10^{-5}\ \text{m}^2\text{s}^{-1}$, which has the correct order of magnitude (remember, we quoted diffusion of water vapor in air of $1.8 \times 10^{-5}\ \text{m}^2\text{s}^{-1}$ in Sect. 6.4). In Sect. 6.4, we stated that diffusion is slower in liquids typically by two orders of magnitude. Even if the above arguments do not rigorously hold for liquids, they are sufficient to explain the difference. In liquids, the mean free path is on the order of the molecular radii, two orders of magnitude smaller than in gases at STP.

We can write Eq. (6.23) in terms of the cross-section using Eq. (6.21)

$$D \cong \frac{\upsilon_{\text{rms}}}{n_p\,\sigma}. \tag{6.24}$$

Considering the gas as ideal, the state equation is $pV = nN_A k_B T$, where n is the number of moles. We write it as

$$p = k_B T n_p. \tag{6.25}$$

Solving for n_p and substituting in Eq. (6.24), we have

$$D \cong \frac{v_{\mathrm{rms}} k_B T}{p\sigma}. \tag{6.26}$$

We now notice that we have neglected our consideration of two different gases and have reasoned that they were equal. Consequently, it is not clear to what *l* and *σ* refer. However, these quantities are similar for different simple molecules with comparatively similar masses and sizes.

From Eq. (6.26), we learn that the diffusion coefficient in gas is inversely proportional to pressure. As for the temperature dependence, we must consider that the root mean square velocity (as any other mean) is proportional to the square root of the temperature. As a consequence, the diffusion coefficient varies as $T^{3/2}$, in the limits in which the cross-section can be considered constant.

Consider now the thermal diffusion . Considering the analogy of the diffusion processes, we can directly write

$$\chi \cong v_{\mathrm{rms}} l. \tag{6.27}$$

For Eq. (6.6), the thermal conductivity is $\kappa = \chi c_V \rho$, assuming a process at constant volume. It is convenient to have the molar heat instead $C_V = c_V N_A m$, where *m* is the mass of the molecule, and we have

$$\kappa = \chi c_V \rho = \chi c_V m n_p = \chi \left(C_p / N_A\right) n_p .$$

For Eqs. (6.27) and (6.21), we get

$$\kappa \cong \frac{v_{\mathrm{rms}}\, l\, C_V n_p}{N_A} = \frac{v_{\mathrm{rms}} C_V}{\sigma N_A}. \tag{6.28}$$

The result is, at first sight, surprising; as neither the cross-section nor the molar heat depends on density, the thermal conductivity of a gas is independent of density, hence of pressure. The reason is the following. When the density decreases, the number of molecules contributing to diffusion decreases, but their mean free path increases and with it the contribution to diffusion of each molecule. One might take this argument to the conclusion that if the conductivity is independent of pressure, it should be the same even at zero pressure, when there is no gas. This is not so, because in our arguments, we have always neglected the presence of confining walls. This is correct if, as is usually the case, the mean free path is much smaller than the distances between walls, but not in the limit above.

Equation (6.28) tells us that, in a first approximation, the thermal conductivity increases with increasing temperature as the root mean square velocity, namely as $T^{1/2}$. Actually, the increase is somewhat more rapid, because, generally, the molar heat increases and the cross-section decreases with increasing temperature.

Finally, we consider viscosity . Once more, we do not need calculation and can exploit the analogy we observed for kinematic viscosity , writing directly

$$\nu = \eta/\rho \cong v_{\mathrm{rms}}\, l. \tag{6.29}$$

For dynamic viscosity, we have

$$\eta \cong n_p m\, v_{\mathrm{rms}}\, l \cong m\, v_{\mathrm{rms}}/\sigma. \tag{6.30}$$

We see that, similar to thermal conductivity, dynamic viscosity of gases does not depend on pressure (as long as the mean free path is smaller than the size of the container). Again similarly, dynamic viscosity increases with temperature as $T^{1/2}$, if the cross-section is constant, a little more rapidly when the cross-section decreases.

As we have seen, the diffusion coefficient, thermal conductivity κ and dynamic viscosity η are strongly correlated quantities. Two (approximate) relations between them are found, one from Eqs. (6.24) and (6.30) and one from Eqs. (6.28) and (6.30). They are

$$\kappa \cong \eta c_V, \quad \eta \cong D\rho. \tag{6.31}$$

We finally observe that the viscosity of liquids generally decreases, in a different way than in gases, for increasing temperature. This is a consequence of the relative motions of the molecules being easier at higher temperatures. The viscosity decrease is small, but appreciable, in the low viscosity liquids like water, while being much larger for viscous liquids, like honey and oils.

Problems

6.1. Two equal bodies, each with thermal capacity $C = 500$ J/K, are connected by a bar of length $l = 1$ m, section $S = 5$ cm^2 and thermal conductivity $\kappa = 20$ WK^{-1}m^{-1}. The system is thermally insulated. The temperatures are initially different. Find the time τ at which the temperature difference became $1/e$ times the initial difference.

6.2. The mean free path at STP of the oxygen molecules is, in round numbers, $l = 100$ nm. What is the mean time between collisions? Use rms velocity.

6.3. A vessel initially containing air is emptied to the pressure of 1.3×10^{-4} Pa. The air temperature is $T = 290$ K. Assuming a collision cross-section of the molecules $\sigma = 5\times10^{-19}$ m^2, find, in order of magnitude, the molecule number density n_p and the mean free path l.

6.4. A spherical vessel of one liter volume contains air. What are the maximum pressure and density at which the mean free path (assume $l = 70$ nm) is larger than the diameter of the vessel?

6.5. Find the thermal conductivity of hydrogen gas if, under the same conditions, its dynamic viscosity is $\eta = 8.6$ µPa s.

6.6. The coefficient of self-diffusion of nitrogen under normal conditions is $D = 1.4 \times 10^{-5}$ m^{2}s^{-1}. Find the values of viscosity and thermal conductivity under the same conditions.

6.7. Consider carbon dioxide (molecular mass $m(CO_2 = 44)$ and nitrogen (molecular mass $m(H_2 = 28)$ at the same temperature and pressure. Find the ratios between their self-diffusion coefficients, dynamic viscosities and thermal conductivities. Assume the ratio of collision cross-sections to be $\sigma_{CO_2}/\sigma_{N_2} = 1.5$ and the molar heats to be $C_{V,CO_2} = (6/2)R$ and $C_{V,N_2} = (5/2)R$.

6.8. The diffusion coefficient of carbon dioxide in air is $D = 1.4 \times 10^{-5}$ m^{2}s^{-1}. What is its value at 100 °C? Consider the cross-section constant.

6.9. The thermal conductivity of carbon dioxide at 0 °C is $\kappa = D = 1.45 \times 10^{-2}$ Jm^{-1}K^{-1} 1.45×10^{-2} J/(mK). What is its value at 100 °C? Consider cross-section and specific heat to be constant.

6.10. The dynamic viscosity of air at 20 °C is $\eta = 18$ µPa s. What is its value at

−20 °C? Consider the cross-section constant.

Answers

1.1. The buoyancy on any quantity of water in air is about one thousandth of the weight, because the density of water is almost one thousand times the density of air. On 10 g, this makes about 10 mg. We must do the correction.

1.2. The larger moves faster, because weight and buoyancy are proportional to the third power of the radius, while the viscous drag is proportional to its square.

1.3. 10 t.

1.4. The period decreases, as the buoyancy diminishes the weight, and we have $T = 2\pi\sqrt{\frac{I}{(mg-F_A)h}}$ instead of $T = 2\pi\sqrt{\frac{I}{mgh}}$.

1.5. They are periodic but not harmonic because the resultant force, weight plus buoyancy, is not proportional to the displacement

1.6. $\upsilon = 3$ m/s.

1.7. $\upsilon = 1.7$ m/s.

1.8. We tentatively assume the flow to be laminar. The mean velocity on the section, given by Eq. (1.42), is $\upsilon = 0.6$ m/s, corresponding to $Re \approx 600$. Hence, the hypothesis of laminar flow is correct. The volumetric flow rate is $Q = 1.9 \times 10^{-6}\,\mathrm{m^3\,s^{-1}}$.

1.9. (a) The resultant force is zero because the velocity is constant. The acting forces are: the weight, vertical downward, the drag, vertical upward, and the buoyancy. The latter is negligible, because the density of the air is much smaller than that of the particles. The force equilibrium equation is then $(4/3)\pi r^3 \rho_p g = 6\pi\left(v_a - v_p\right)\eta$. Solving for r , we have r = 15.8 μm. (b) The Reynolds number is Re = 0.19, hence, the regime is laminar and the Stokes law, which we have used, holds.

2.1. The heat released by water is cm (60°–20°), where c is the known specific heat of water, and is equal to the heat absorbed by the calorimeter c_{cal} (80°–60°). From here, we get c_{cal} . This is the standard procedure for obtaining the so-called water-equivalent of the calorimeter.

2.2. In the colder room.

2.4. $n_p = nN_A/V = pN_A/(RT) = 2.69 \times 10^{25}\ \mathrm{m}^{-3}$.

2.5. ρ = 1.29 kg m $^{-3}$.

2.6. μ = 28.9 g/mol.

2.7. The internal energy grows in processes (a) and (b), is constant in (c), and decreases in (d) and (e).

2.8. The heat exchange with the environment is zero, and the external work is zero, hence the internal energy does not vary.

2.9. $2\,Mgh = cm\Delta T$, where c is the water specific heat ($c = 4186$ J kg^{-1} K^{-1}). $\Delta T = 1.16 \times 10^{-2}$ K. This is very small. In practice, the drop of the weights must be repeated many times.

2.10. $Q = -2.1$ kJ.

2.11. The sum of process 1 and of the inverse of 2 is a cycle. Process 2 being reversible, the heat exchanged in the inverse process is $-Q_2$. The internal energy does not vary in the cycle and we have $Q_1 - Q_2 = W_1 - W_2 = 2\text{ kJ}$.

2.12. (a) and (b) $Q = -200$ kJ.

2.13. $p_3/p_1 = (V_1/V_2)^{\gamma-1} = (p_2/p_1)^{\frac{\gamma-1}{\gamma}} = 1.51.$

2.14. $\Delta U = p\,(V_2 - V_1)\,/\,(\gamma - 1)$, $\Delta U = p\,(V_2 - V_1)$, $Q = \gamma p\,(V_2 - V_1)\,/\,(\gamma - 1)$.

2.15. (a) $\gamma = Q/\Delta U = Q/\,[nC_V\,(T_2 - T_1)]$; $\gamma/\,(\gamma - 1) = Q/\,[nR\,(T_2 - T_1)] = 4.03$ and $\gamma = 1.33$. (b) $\Delta U = 5$ kJ, (c) $W = 1.65$ kJ.

2.16. (a) $\Delta U = 0$; (b) $W = 139$ kJ; (c) $Q = 139$ kJ.

3.1. $\eta = 1 - T_H/T_C = 0.40$, $\eta' = 1 - T'_H/T'_C = 0.45$. Hence, $T'_H - T_H = 42.4\,\mathrm{K}$.

3.5. The number of moles is n = 400/32 = 12.5. $W_{DA} = -\Delta U = nC_V(T_C - T_H)$. We know all but T_C . To find it, start from the efficiency. The absorbed heat is $Q_{AB} = nRT\, ln \frac{V_B}{V_A} = 36\,\mathrm{kJ}$ and $\eta = W / Q_{AB}$ = 0.4. Hence, $T_C = T_H(1 - \eta) = 300\,\mathrm{K}$ and W_{DA} = −52 kJ.

3.6. (a) and (b) increases; (c) constant.

3.7. Yes, if the process is irreversible.

3.8. (a) The entropy variation is the same; (b) No.

3.9. *A* and *C* are on the same adiabatic curve.

3.10. $S(C) - S(B) = S(A) - S(B) = \frac{Q_{AB}}{T} = \frac{W_{AB}}{T} = -10\,\mathrm{J/K}$.

3.11. (a) 8.6 J/K, (a) 14.4 J/K.

3.12. It increases by 20 J/K.

3.13. $W = T\,(S_2 - S_1)$.

3.14. $S\,(T) = \int_0^T \frac{C dT}{T} = \alpha \int_0^T T^2 dT = \frac{1}{3}\alpha T^3 = \frac{C}{3}$.

4.1. Pressure would increase.

4.2. Water would transform into an ideal gas. The number of moles is n = 1000 g/(18 mol/g) = 55.6. Hence, the pressure $p = nRT / V$ = 138 MPa.

4.3.

(a) $p_{\mathrm{id}} = 1.44 \times 10^6\,\mathrm{Pa};\ 1.36 \times 10^6\,\mathrm{Pa} = 0.95\,p_{\mathrm{id}}$

(b) $p_{\mathrm{id}} = 1.44 \times 10^7\,\mathrm{Pa};\ 1.0 \times 10^7\,\mathrm{Pa} = 0.70\,p_{\mathrm{id}};$

(c) $a = 0.54\,\mathrm{Pa\,m^6\,mol^2},\ b = 7.8 \times 10^{-5}\,\mathrm{m^3\,mol^{-1}};$
$p_{\mathrm{id}} = 2.88 \times 10^6\,\mathrm{Pa};\ p = 2.59 \times 10^6\,\mathrm{Pa} = 0.90\,p_{\mathrm{id}}.$

4.4. $W = nRT(V_f - b)/\,(V_i - b) + a\left(1/V_f - 1/V_i\right)$.

4.5. $Q = 0$ and $W = 0$ Hence, $0 = \Delta U = nC_V\Delta T - a\left(1/V_f - 1/V_i\right)$. With $C_V = (5/2)R$, $\Delta T = -5.8$ K.

4.6. $\Delta U = -a\left(1/V_f - 1/V_i\right)$; $Q = \Delta U = 86.4$ J ($a = 4\ a_{\text{mol}}$).

4.7. (a) $\Delta U = 0.26$ kJ (0 for an ideal gas); (b) 4.14 kJ (96 % of the ideal gas); (c) $Q = 4.4$ kJ; (d) $\Delta S = 25.4$ JK $^{-1}$ (1.02 than the ideal gas).

4.8. The substances with triple point pressure larger than the atmospheric pressure.

4.9. The specific volumes are the reciprocals of the densities. (a) $V_l = 10^{-3}$ m ^{3} kg $^{-1}$; (b) $V_V = 1.25$ m ^{3} kg $^{-1}$ $V_v = 1.25$ m ^{3} /kg.

4.10. $Q_V = R\frac{T_1T_2}{T_2-T_1}ln\left(p_2/p_1\right) = 61\,\text{kJ}\,\text{mol}^{-1}$.

4.11. The molar latent heat for vaporization is Q_{ev} =31.2 kJ mol $^{-1}$. $P_2 = 86$ kPa.

4.12. (a) $W_{12} = p_s\left(V_v - V_l\right)$, (b) $Q_{12} = mQ_v$, (c) $U_2 - U_1 = mQ_v - mp_s\left(V_2 - V_1\right)$, (d) $S_2 - S_1 = mQ_v/T$, (e) $H_2 - H_1 = mQ_v$.

4.13. One point on the liquid-vapor coexistence curve.

4.15. A hole can be considered a capillary. Let r be the curvature radius of the drop forming under a hole. The drop detaches when the capillary pressure $2\tau/r$ balances the hydrostatic pressure ρgh. The maximum internal pressure corresponds to the minimum possible radius, which is the radius of the hole. Then, $h_{max} = 2\tau/(\rho g a) = 60\,\mathrm{mm}$.

4.16. The applied overpressure p must be at least equal to the sum of the hydrostatic pressure ρgh and the capillary pressure $2\tau/a$. Then $p > 488$ Pa.

4.17. $h = 2\tau/(\rho g a) = 60\,\mathrm{mm}$ (OK, it is less than the length of the capillary).

4.18. The vertical forces are the weight mg downward, the buoyancy $a^2 \rho gh$ upward and the force due to the surface tension $4a\tau\cos\theta$, where θ is the contact angle. In both given cases, its magnitude is $4a\tau$; in case (a), it is downward and $h = 6.5$ mm, in case (b), upward and $h = 4.6$ mm.

5.2. $M = 0.54$ mg.

5.3. $n_p\left(\mathrm{m}^{-3}\right) = \frac{N_A\rho\left(\mathrm{kg\,m^{-3}}\right)}{\mu\left(\mathrm{g/mol}\right)1000}$

5.4.
$N/V = 2.7 \times 10^{25}\ \text{m}^{-3}$.

5.5. (a) $(N/V)\,v\cos\theta$, (b) $p = 2\,(N/V)\,mv^2\cos^2\theta$,

5.6. (a) $T = 290$ K; (b) $N/V = 2.5 \times 10^{25}\ \text{m}^{-3}$.

5.7. (a) $A = 2a/\sqrt{12}$; (b) $A = 2a/\sqrt{24}$.

5.8. $\upsilon_p = 417$ m/s; $\upsilon_{rms} = 510$ m/s.

5.9. Classically, there are two rotational degrees of freedom. Hence, $I\left\langle\omega^2\right\rangle/2 = k_BT$ and, if m_{N2} is the mass of the nitrogen molecule, we have

$$\sqrt{\omega^2} = \sqrt{\frac{8k_BT}{2m_{N2}a^2}} = 2.3 \times 10^{12}\,\text{s}^{-1}.$$

5.10. $dN(r) = n_0 e^{-[U_p(r)-U_p(r_0)]/k_BT}\,4\pi r^2\,dr$

5.11. (a) 37 %; (b) 5 %.

5.12. Use Eq. (5.35), modified to take into account the presence of water, i.e., $n_p(z_2)/n_p(z_1) = exp\left[-\frac{(4/3)\pi(\rho_p-\rho_w)r^3g(z_2-z_1)}{k_BT}\right]$. We get

$$N_A = \frac{RT}{(4/3)\pi\left(\rho_p - \rho_w\right) r^3 g\,(z_2 - z_1)} ln n_p(z_2)/n_p(z_1) = 6.2 \times 10^{23}.$$

5.13. $p = p_0\, exp\left(-\frac{mgh}{k_B T}\right) = p_0 \exp\left(-1.2 \times 10^{-4} h\,(\mathrm{m})\right)$, (a) $p = 0.55\, p_0$; (b) $p = 0.34\, p_0$; (c) $p = 01.27\, p_0$;

5.14. (a): (0, 6): $\Gamma = 1$, $P = 1/64 = 0.6$ % (64 is the sum of all probabilities Γ); (1, 5): $\Gamma = 6$; $P = 9.4$ %; (2, 4): $\Gamma = 15$; $P = 23.4$ %; (3, 3): $\Gamma = 20$; $P = 31.3$ %; (b) $P = 8.1$ %.

5.15. $\Delta S = 0.96 \times 10^{-23}\,\mathrm{J\,K^{-1}}$

5.16. $\Delta S = 6050\,\mathrm{J\,K^{-1}}$, $\Gamma_v/\Gamma_l = 10^{1.9\times10^{26}}$

5.17. (a) $\Delta S = 9.5 \times 10^{-23}\,\mathrm{J\,K^{-1}}$; (b) No, it is too small, $\Delta T = 6.7 \times 10^{-18}\,\mathrm{K}$

5.18. It does not vary if the process is reversible, but increases if it is irreversible.

6.1. $\tau = (Cl)/(2\kappa S) = 6.94$ h.

6.2. $\tau = 150$ ps.

6.3. $n_p \approxeq 4 \times 10^{16}\ \mathrm{m}^{-3}$, $l \approxeq 100$ m.

6.4. $\rho = 3 \times 10^{-7}\ \mathrm{kg\ m}^{-3}$, $p = 0.03$ Pa.

6.5. $\kappa = 0.090$ W/(m K).

6.6. $\eta = 18$ µPa s, $\kappa = 1.3 \times 10^{-2}\ \mathrm{W\ m}^{-1}\ \mathrm{K}^{-1}$.

6.7. $$D_{\mathrm{CO_2}}/D_{\mathrm{N_2}} = \left(\sigma_{\mathrm{N_2}}/\sigma_{\mathrm{CO_2}}\right)\left(v_{\mathrm{rms,CO_2}}/v_{\mathrm{rms,N_2}}\right) = (1/1.5)\ \sqrt{m_{\mathrm{N_2}}/m_{\mathrm{CO_2}}} = 0.53$$

$$\eta_{\mathrm{CO_2}}/\eta_{\mathrm{N_2}} = \left(D_{\mathrm{CO_2}}/D_{\mathrm{N_2}}\right)\left(\rho_{\mathrm{CO_2}}/\rho_{\mathrm{N_2}}\right) = 0.53\left(m_{\mathrm{CO_2}}/m_{\mathrm{N_2}}\right) = 0.83\ .$$

Index