

Chapter 1

Prelude: playing by ear

With the arrival of your most precious wine, and with this heat, my meditation is about the foresaid heat and cooling the wine. The measurement of heat is already reduced almost to perfection, and I have made records of it for the last 15 days. . . I have also found a funnel that quickly cools wine passed through it. . . [and a glass] enabling to see how much wine has cooled, which can be use for drinking too. . . These inventions came to my mind after sipping just two glasses of the wine of Your Excellency, therefore I foresee that, before finishing one of your flasks, I shall devise sublime things.

Giovanni Francesco Sagredo, letter to Galileo, 27 July 1613

Statistical physics is a “subtle” subject, whose very basic ideas are arguably harder to master than specific applications. The quotation that opens this chapter,¹ where Sagredo acquaints his great friend Galileo with his latest discoveries in thermometry, suggests to us that starting with an aperitif may not be such a bad idea, provided, of course, we do not get too drunk. The best aperitif to prepare us for our main course is thermodynamics. The little parallel with music spanning this book may help you understand why this is a good and reasonably easy start. Even a musical semiliterate like me, who toils at reading a score, can indeed strum on a guitar, basically playing by ear. For what we plan to investigate, starting with thermodynamics is akin to playing by ear. In fact, thermodynamics is a kind of “science apart” based on its own simple principles, which accounts for many features of our day-to-day experience without requiring a strong background in physics.

This does not mean that we can go very far without technical tools. For we can surely enjoy a song even if we cannot read its score at all, but it is only by learning the richness and subtleties of musical language that we can fully appreciate a Beethoven string quartet or a Miles Davis solo. Similarly, to fully grasp the deep meaning of the statistical approach to the physical world without reducing everything to a vague blathering about complexity, chaos, and destiny of the Universe requires a little fund of technical knowledge. But, as Pete Seeger used to sing, for everything there is a season. This is our season for thermodynamics, whose basic ideas we review in this chapter. First, however, let us reveal in advance our true final target.

¹ I took some freedom in translating: the Italian original, which can be read from the Galileo Portal, <http://bibdig.museogalileo.it/Teca/Viewer?an=354814&seq=596&q=sagredo>, is surely more fascinating.

1.1 Time: an arrow or a boomerang?

The whole story, in a nutshell, is just this: there is a blatant and disquieting contradiction between our common sense and the basic laws you have learned in elementary physics courses. Our daily experience tells us that time “flows” in a precise direction, from what we call “past” to what we call “future”. This watershed between past and future events is so rooted in our experience that, while we have a term to say that something happened “elsewhere” - somewhere else in space, no matter where - the word “elsewhen” is used only in science fiction novels: we always say “before” or “after”. Where does this asymmetry between space and time come from? Let’s see. We are used to seeing cups fall from tables and break in pieces, whereas we never see a myriad of fragments getting together into a cup that spontaneously jumps from the floor on the table. Slightly more dismally, each one of us is born, grows up, ages, and eventually departs. In other words, spontaneous processes in nature are *irreversible*, and the direction along which they proceed fixes an “arrow of time”. Irreversibility is then a rather intuitive concept, but since intuitive concepts are sometimes misleading, it is better to state right from the start what we mean by that. A process is irreversible if, when filmed and then shown backward (which was probably easier with old films than with digital movies), it looks odd, funny, unlikely (like glasses jumping on a table, or corpses rising from the grave). This may sound a bit fuzzy, but I have never found a better definition.²

What is really weird is that there is absolutely *nothing* in the fundamental laws of physics that points to a direction of time. Both classical mechanics, which portrays the Universe as a kind of perfect clockwork, and quantum physics, which accounts for the rather capricious behavior of the microscopic world, are perfectly reversible in time. Namely, if we reverse, at a given instant, the velocities of all the particles of a classical system (a new, perfectly admissible initial condition), all of them would retrace the path they have followed in the past.³ This means that all the energy spent to break in pieces the cup we mentioned (and to warm both those pieces and floor, although imperceptibly) can, in principle, be conveyed back to the center of mass of the cup, which could then jump up to the height where it fell from. This sounds weird, but is not *forbidden*.

So, why don’t we see it? As a matter of fact, this is the deepest question addressed by statistical physics, and for sure it is not trivial. Anticipating the answer, we can say that statistical mechanics does not say that it is *impossible* to watch a bunch of cup fragments rejoining, or corpses getting back to life. It

² A little question, which may not be so trivial: *With what* (better, *with whom*) are you actually comparing these movies to find them odd, funny, or unlikely? In other words, ask yourself if, to quantify the *direction* of time, you need some other “reference” besides a watch.

³ As we shall see, the same would happen for a quantum system, provided that the wave-function is changed into its complex conjugate.

just says that there is a *ludicrous* chance that this happens. And when I say ludicrous, I really mean it: we might have to wait many, many more times the life of the Universe, which is arguably too much even for a Buddhist believing in reincarnation. But, given enough patience, this *will* happen. In other words, as we shall see, the arrow of time may eventually (but very eventually) be a boomerang. To clarify this statement, we have however a long way to go. This will lead us to discover that the real secret of irreversibility is that we are considering macroscopic systems, made of an exceptionally large number of particles, typically⁴ of the order of the Avogadro number $\mathcal{N}_A \simeq 6 \times 10^{23}$. Thermodynamics is just our “antipasto”, an appetizer whose main ingredient is the non-trivial concept of equilibrium.

1.1.1.1 “And yet it stops”: equilibrium, the Holy Grail of thermodynamics

Galileo might be turning over in his grave after reading this title. And yet, this is true. The earth revolves round the sun, mountains rise from its oceans, rain washes them out, new lands develop from the silt carried by streams in an apparently endless cycle. But experience tells us that change is not perennial, that things eventually come to a rest. When red-hot volcanic bombs fall into the sea, water around them stirs, boils, vaporize. Yet, when the eruption ends, the sea eventually reverts to stillness, and we can pick up the lava debris without getting burnt. Insofar as our limited senses tell us, nothing changes anymore. In a kind of suicidal rush, the arrow of time points to a timeless, unchanging state that we may be tempted to call *equilibrium*. We must however be very careful, for the idea of equilibrium is subtler than we may naively expect. We better muse a bit upon it.

1.1.1.1 Does it really stop? Equilibrium and observation time

Before stating that “nothing changes anymore” we better wait, possibly a lot. To introduce this fundamental issue, I cannot do better than quoting Shang-Keng Ma.

If we pour some boiling water from a thermos flask into a tea cup, after a few seconds the water in the cup becomes an unchanging state at rest, and this is an equilibrium state. The volume of the water in the cup can be measured from the height of the water and the cross section area of the cup. The temperature of the water can be measured by a thermometer. Within one or two seconds, these measured quantities will not change significantly, and thus during this observation time the water is in equilibrium. If the time is too long, the temperature will obviously change

⁴ Actually, we shall rather have to deal with numbers like $\exp(\mathcal{N}_A)$, which sound preposterously large even compared to $\mathcal{N}_A$. In fact, numbers so large fully baffle our intuition.

and the water cannot be regarded as being in an equilibrium state. After an hour, the temperature of the water will be equal to the room temperature. If the room temperature is constant for several hours, then the temperature of the water will accordingly remain constant. Therefore, if the observation time is within this range of several hours, the cup of water can be regarded as being in equilibrium. However, water molecules will continually evaporate from the water surface. In several hours the volume of the water will not change significantly, but if the time is long enough, such as two or three days, all the water will evaporate. So if the observation period is over two or three days, this cup of water cannot be regarded as being in an equilibrium state. After a few days, when all the water has evaporated, then again this empty cup can be regarded as being in an equilibrium state. However, strictly speaking, even this cup is not in an absolute unchanging state, because the molecules of the cup can evaporate, although in a few years time there will be no change observed.

There is a keyword in this quotation, “temperature”, that we have not properly defined yet. Nevertheless, I am sure that the customary meaning you give to this word has allowed you to grasp the main message: a system may seem to be in different equilibrium states, depending on the time scale over which we observe it. The fact is, this is not exactly what we long for. As we shall see, one of the main goals of thermodynamics is indeed finding a physical quantity that characterizes equilibrium as being maximal or minimal. This requires a system, properly defined also in terms of its relation with its “surroundings”, to have a *single* equilibrium state (possibly consisting of the coexistence of different “phases”, like water in equilibrium with vapor).

We may nevertheless wonder whether the “temporary” equilibrium states described in the example by Ma can be made permanent by just changing the way the system is linked to the surrounding. For sure, if hot water is poured in a cup hermetically sealed by a screw cap, no more vapor is lost to the external atmosphere, and water reaches an equilibrium with its own vapor within the cup. Finding containers capable to keep matter inside, without sublimating themselves, for much more than the “few years” mentioned in Ma’s passage (which is a strong understatement) is not a serious problem. Yet, water would still cool down. If the container is a thermos flask, however, this will take a much longer time. In fact, one of the key steps in developing thermodynamics is conceiving a kind of “idealized thermos”, which fully prevents a system to reach what is called *thermal* equilibrium with the surrounding. We may then tentatively assume that, over a period much shorter than the cooling time, the physical properties of the system are not too different from those of water contained in an ideal thermos. Similarly, since evaporation is slow, we may expect that, for several hours after cooling is completed, the state of the system does not differ too much from that of a fixed amount of water contained in a cup with a screw cap. Does this picture of a “time-bracketed” equilibrium always work? Often, but not always. Even if it apparently does, this “equilibrium” may be rather baffling (see below).

1.1.1.2 Lethargic glasses, napping vapor, and topsy-turvy magnets

For the cup we considered, slowness in attaining “true” equilibrium is due to the nature of the separation boundaries with the surrounding. For instance, if the cup were made of metal, thermal equilibrium with the room would be reached faster. Yet, there are systems that are intrinsically “lethargic”, because of their extremely slow *internal* time scales. The best example of these kinds of system is glasses. Undeniably, the glass of a windowpane is mechanically a solid: it neither flows as a liquid, nor surely spreads around as a gas. Still, glass has little to do with what physicists consider a true solid: microscopically, atoms and molecules are in fact arranged like in a *liquid*, which has important macroscopic consequences. For instance, true solids melt, suddenly becoming a liquid at a very well-defined temperature. By increasing temperature, glass becomes what we are used to calling a “liquid” too, but it does so by *softening* progressively, in a rather extended temperature range. The common soda-lime glass is just an example of a large class of materials, including metal and polymer glasses but also structures of great biological relevance like gels, which physicists call “arrested solids”. These “pseudo-equilibrium” (also called *metastable*) structures are very long living: it may take a glass thousands or even millions of years to reach its true equilibrium state, which is in fact a crystalline solid. Several thermodynamic properties of these disordered materials are quite different both from liquids and from crystals. In fact, the specific arrested structure in which a liquid “freezes” when it becomes a glass is just one among a myriad of distinct but almost identical arrangements. Because of this extremely complex structural “landscape”, understanding why a liquid freezes into a glass and accounting for the complex behavior of these materials are still very open subjects of research.

Lethargic glasses are not however the only kind of metastable states. For example, suppose you wish to cook spaghetti, so you have put a steel pot on the fire. Of course you know that, to cook pasta properly, you have to toss it when water is already vigorously bubbling. But you are quite hungry, and that damn water seems very reluctant to boil. Hence you do it as soon as you spot some tiny bubbles forming on the bottom surface of the pot. Surprisingly, that hesitant water suddenly starts to boil vigorously.⁵ What happened? In the pot, the temperature was actually slightly above 100°C, a value where boiling is supposed to start.⁶ Yet, for some unknown reason, water takes a little nap, remaining in a “superheated” state until the contact with pasta wakes it up. Before it boils, all water properties have values that are almost identical to those just below 100°C, yet the superheated state is not *stable*, since an external perturbation like tossing pasta - or adding salt, if you have not done it yet - tells water that there is a *better* equilibrium condition. This

⁵ You better try with a steel pot: with aluminum cookware, this would hardly happen (understanding why is not trivial at all).

⁶ At sea level: each 1000 m increase in altitude causes the boiling point to drop by about 3.5° C.

abrupt awakening is rather peculiar: if you blow on a cup, tea cools a bit faster - again, we have to inquire why - hence thermal equilibrium is attained in a shorter time: but nothing as spectacular as boiling happens.⁷ Chemical reactions are another class of phenomena where the road to equilibrium is often blocked by a little “bump”. Hydrogen and oxygen can live peacefully together in a tank: but just turn on a lighter - please, do it metaphorically - and see what happens!

After all, in elementary physics courses, you have surely encountered the concept of stability of mechanical equilibrium. Put a ball at the bottom of a hemispherical container, and it will happily stay there: this is stable equilibrium under gravity. Now turn the container upside down, and put the ball right on the top of it: this is also a state of mechanical equilibrium, but totally unstable, for any small vibrations will cause the ball to fall from a position where the potential energy is maximal. In Chap. 3, we shall see that something similar can happen for *thermal* equilibrium: take many little magnets, pointing as they should to the (magnetic) North Pole, and then suddenly turn the earth upside down, so they now point to the south. Of course, turning the earth is not easy, so one must use a more sophisticated strategy, but the result is rewarding and astonishing. For some time, the magnets do not realize that the magnetic field was overturned, and behave as though they are in a kind of equilibrium in which it is hard to guess what is hot, what is cold, and what “temperature” means. Nevertheless, these weird equilibrium states can be obtained and profitably exploited in important practical applications.

An important clue of metastability is often the following. Suppose that, starting from different initial conditions, you reach, by performing the same operations on the sample, *different* final states where nothing apparently changes. In other words, suppose that the final state depends on the previous *history* of the sample, which is for instance the case of ferromagnetic materials that we discuss in Chap. 5. This cannot be a true equilibrium state, at least if we plan to develop a science like thermodynamics that pretends to relate to each other all equilibrium properties, with no reference to what has happened before to the system. For specific values of the *external parameters*, by which we mean all physical properties of a system that are under the direct control of the experimenter,⁸ equilibrium must be *unique*.

⁷ Similarly, a slight thermal or mechanical disturbance does not usually drive a glass to equilibrium, but rather to another slightly different glassy state. Actually, if a tiny crystallite forms, the whole glass might rapidly turn into an ordered solid: yet, inducing crystallization in a glass by an external perturbation is usually much harder than promoting the formation of gas bubbles in a superheated liquid.

⁸ As we shall see, external parameters are not only geometrical properties, such as the system volume or surface, but also external mechanical or electromagnetic “fields” (including gravity, which can be eliminated if experiments are performed on the International Space Station, or “amplified” by a centrifuge).

1.1.1.3 False friends: nonequilibrium stationary states

Keeping in mind the above caveats, we could still tentatively call “equilibrium” the condition reached when nothing apparently changes, provided we frame it within a given observation period and that we test its stability. Yet, unfortunately, something may still be happening even in the absence of any visible change. Take for instance a long metal bar, placing one of its ends inside a stove and the other one in the fridge. After a very short time, the bar reaches a state where nothing apparently changes. If you check accurately, you may find that the portion of the bar close to the stove is slightly more expanded (less dense) than the region close to the fridge and, if you have a thermometer (which again I assume you have the common concept of), you will find that the temperature along the bar decreases linearly between the two ends: yet, neither the density, nor the temperature distribution change in time.

Is this a true equilibrium state? Definitely no, because something *does* take place: even if we do not see it, energy, in a form that we shall shortly call heat, flows through the bar from the stove, acting as a “source”, to the fridge, acting as a “sink”. This is what we shall call a *stationary* state: energy *must* flow to maintain a time-invariant density and temperature profile. Unfortunately, stationary states where these “invisible” flows are present lack the simple, elegant thermodynamic description of true equilibrium states. We already mentioned for instance that equilibrium is characterized by a maximum or a minimum of a suitably defined quantity: this is not the case of stationary (also called “steady”) states. One may however try to bypass the problem, assuming that each small portion of the bar, within which temperature and density are approximatively constant, is in a kind of “local” equilibrium. This approach is often very powerful, giving a correct picture of many phenomena happening at steady state. Yet, this description relies on assumptions that are surely more controversial than those of equilibrium thermodynamics. In fact, there are effects that *cannot* be explained using a local equilibrium approach. Much worse than this, when trying to give a *microscopic* interpretation of thermodynamics, stationary states turn out to be a real nightmare. Hence, when dealing with stationary states, be very careful!

Summing up, “equilibrium” is everything but a trivial concept, and care should always be taken to ensure that concepts developed for a true equilibrium state also work in a condition that may just *look* like equilibrium. Eventually, the best solution is perhaps assuming, as suggested by Herbert Callen, that a system is in equilibrium if its properties are properly described by (equilibrium) thermodynamics.⁹ It is then time to focus on the almost “magical” science of thermodynamics.

⁹ This may almost sound like a tautology, but it is not so different from what we do in other fields. For instance, the idea of probability has been given very different and often conflicting “interpretations”. Eventually, the best solution is taking probability as an

1.1.2 *Thermodynamics: a bold and possibly paradoxical science*

Thermodynamics is a rather “bold” science, and this is possibly the reason for its almost magical success. In fact, thermodynamics claims that the equilibrium state of a system, and how equilibrium changes when some constraint is released (allowing for instance the system exchange energy or matter with the surrounding), can be fully described by means of a few macroscopic physical quantities. This means that, in principle, thermodynamics does not require to know the laws obeyed by the microscopic constituents of the system. Thermodynamics is then a kind of meta-physics, namely, a science that is “beyond” physics.¹⁰ Most likely, the cardinal laws of thermodynamics would also apply if our description of the microscopic world changed completely. As a matter of fact, thermodynamics can shed light on intricate problems even when the underlying physics is still obscure. For instance, Stephen Hawking has developed a powerful and elegant description of the thermodynamics of black holes, although a full understanding of these strange objects requires a quantum theory of gravity, which is so far wishful thinking.

Definitely, trying to explain almost everything, from the operation of a car engine to the fate of a black hole, with a few general laws and a small number of physical quantities sounds rather bold. But the real key to success of thermodynamics is the peculiar nature of some of these quantities. Together with concepts like energy or pressure which are already familiar from mechanics, new quantities “emerge” from considering the *collective* behavior of the huge amount of atoms a macroscopic system is made of, quantities which have no meaning when referring to a single one of these atoms. In fact, when all is said and done, thermodynamics works just because it deals with the collective properties of these huge collections of atoms.

There is however another, apparently paradoxical aspect of thermodynamics. We said that equilibrium is, in some sense, the “graveyard” of time: at equilibrium, nothing changes. Actually, thermodynamics does describe how the macroscopic world changes, which may justify the suffix *dynamics* instead of *statics*, but only to the limited extent of comparing an initial to a final equilibrium state, without describing what happens in between. In thermodynamics, there is no need at all for a “flowing time”: all we have is a sequence of snapshots of a system in different equilibrium states, which can be related using transformation laws. How can such a timeless science pretend to explain the arrow of time? Yet, this sounds less paradoxical if we observe that most of the discrete sequences of snapshots thermodynamics deals with are still *ordered*: quite often, if frame B follows frame A in a natural sequence, a

axiomatic concept that satisfies some precise rules, and then check whether or not these abstract rules accounts for the facts we want to investigate by probabilistic methods.

¹⁰ Although, with due respect to the original “etymology” of metaphysics, we may also say that it comes “after” physics.

sequence where A follows B does not occur. After all, this is totally equivalent to viewing one of those “irreversible movies” we wish to explain, themselves sequences of frames. Whether or not we fancy these movies immersed in a continuously flowing time, with each frame labeled with the instant when it was taken, is totally irrelevant, and may be just a way to satisfy our desire of a “dynamical” view of the world. But the question of a *privileged ordering* of macroscopic phenomena is real. Fully aware that philosophers may turn up their noses, I shall then go on discussing the problem of the arrow of time starting from this apparently timeless science of thermodynamics.

1.2 Temperature

1.2.1 Thermodynamic systems

We have seen that the state reached by hot water in a cup depends on the way we “confine” it. If the cup is actually a thermos, water is clearly in a state of *mechanical* equilibrium, meaning that the hot fluid does not spread around. Of course, if instead of water the thermos contains compressed air, its walls must sustain the pressure difference with the outside: hence, mechanical equilibrium has to do with pressure, a concept we are already familiar with. Vapor remains confined in a sealed cup too, but in this case the system reaches a condition of *thermal* equilibrium, meaning that if we dip our finger in it, we cannot appreciate any physiological change. Walls like those of an ideal thermos, which prevent a system from reaching thermal equilibrium, are called, for reasons we shall shortly see, *adiabatic*, and a system fully enclosed by adiabatic walls is called *isolated*. A *closed* system is instead a system bounded by barriers that are unfit for thermal insulation, but still keep matter in, which are technically dubbed *diathermal* walls.

Let us however come back to the thermos: if we look more carefully, it actually contains both liquid water and vapor, which we may regard as *two* systems in equilibrium with each other. This is another kind of equilibrium that, to be reached, requires some liquid to evaporate or vapor to condense. Liquid and vapor are then *open* to mass exchange. Yet, here there is no physical “wall” keeping liquid and vapor apart: what keeps water to the bottom and vapor to the top is gravity, which in fact *acts* as a wall. Although they are free to exchange mass, open systems can be kept apart by an “external field” like gravity (but also, as we shall see, electric and magnetic forces). Suppose then we carry the thermos inside the International Space Station, where gravity is (almost) absent: does weightlessness prevents water and vapor to split up? Actually, it does not: the thermos would contain a single big vapor bubble surrounded by water, or conversely a big drop of water surrounded by vapor, depending on the material its walls are made of. Here,

as we shall discuss, what keeps liquid and vapor apart is the demand of minimizing their contact area, which has an energy cost. As you see, for open systems the concept of “wall” is rather fuzzy. In fact, a liquid or even a solid that can exchange mass with an *infinitely large* external environment will eventually spread all around. If a system is made of more than one molecular component, like a salt solution, finding a *semipermeable* wall that retains one of the components inside, letting the others pass more or less freely, is much easier. Your kidneys are a splendid example of these “partially open” systems, since their walls discharge water and waste products of metabolism while retaining proteins and other precious biological molecules.¹¹

What does actually fix, however, the relative amounts of liquid and vapor in the thermos, or the quantity of solvent that passes from a salt solution into an external water reservoir through a semipermeable membrane? Understanding mass-exchange equilibrium in open systems will lead us to individuate a new, extremely important thermodynamic quantity. For the moment, we shall be content to deal with thermal equilibrium, which will allow us to introduce the first of these “collective variables”: temperature.

1.2.2 The Zeroth Law

A simple assumption we can make about thermal equilibrium is the following:

If a system A is in thermal equilibrium with a system B, and B is in thermal equilibrium with a system C, then A is also in thermal equilibrium with C

which, using the symbol $\sim$ for “is in thermal equilibrium with” (and $\wedge$ for “and”), can be formally written

$$(A \sim B) \wedge (B \sim C) \Rightarrow (A \sim C). \quad (1.1)$$

This may sound obvious, but it is not: what we just stated is the so-called “Zeroth Law” of thermodynamics, embodying the idea that the equilibrium state reached when a system is in thermal contact with another one is unique, which is actually crucial to give a precise meaning to the concept of temperature. In fact, Eq. (1.1) states that “being in thermal equilibrium” is a transitive relation, and since it is of course reflexive ($A \sim A$) and symmetric too ($A \sim B \Rightarrow B \sim A$), it is mathematically an equivalence relation, which partitions the thermodynamical systems into classes with each class consisting of all and only those systems that are in mutual thermal equilibrium. We can then label each class with the value of a “property”, which is in fact *defined* through this equivalence relation. Assuming that this property takes

¹¹ Admittedly, these are amazingly complex walls, for the whole “traffic” is scrupulously regulated: so much that we are still far from being able to simulate them artificially.

on continuous values, each label can be a real number, which we call temperature. What two systems in thermal equilibrium share is then the value of temperature.

But how do we *measure* temperature? The basic idea is using as reference a system that has some physical property that depends on T according to a known law. Systems A and C will then be at the same temperature if, when put in thermal contact with a system B , which plays the role of a “thermometer”, the latter indicates the same value on a suitable scale. This however allows us to set just an *empirical* temperature scale, like the temperature in degrees Celsius or Fahrenheit we commonly use. As we shall see, a much more meaningful scale can be defined by investigating the operation of heat engines.

♡ In fact, making a “good” thermometer is everything but trivial. Thermometry has a long history, which dates back to the first instruments developed by Galileo and Santorio Santorio¹² at the beginning of the XVII century, although some methods to quantify the “degree of heat” (or cold) were already known to Hellenistic scientists. The conceptually simplest thermometers exploit the observation that gases, when heated, expand consistently. In fact, gas thermometers are in principle quite reliable, because their volumetric expansion is *linear* in T , if they are sufficiently dilute (provided, of course, that the expansion is made at constant pressure, because unavoidably a gas thermometer is a barometer too, which is the main practical properties). Moreover, how do we know that at constant pressure the volume of a dilute gas increases *linearly* with T ? Because arguably all of us were told about the experiments originally done by Dalton and later confirmed by Gay-Lussac (who however credited the discovery to his friend Jacques Charles). But wait a moment: how did Dalton measure temperature? In fact, with a mercury thermometer.

Thermometers based on the thermal expansion of a liquid are much less sensitive to pressure variation than gas thermometers. Yet, even supposing that we have fixed two points on the scale, like the freezing and boiling point of water,¹³ how do we know that the thermometer gives the correct temperature *in between* these points? In fact it does not, because the coefficient of thermal expansion of liquids depends itself on temperature. Thus, if we have calibrated the thermometer by placing two marks on the scale at, say, 0°C and 100°C , the level of the liquid at 50°C will not be *exactly* halfway between the marks. For liquids like water or alcohol, this effect is not negligible, while the thermal expansion of liquid metals like mercury is almost linear in a wide temperature range.

However, how do we know that mercury is better than water or alcohol? Henry-Victor Regnault had a very interesting proposal to overcome these conceptual difficulties in calibrating thermometers. If we take exactly the *same* amounts of freezing and boiling water and mix them in an adiabatic calorimeter preventing the dispersion of heat, he claimed, the resulting mixture will necessarily be at the average temperature of 50°C . This condi-

¹² This is not a misprint: sometimes Italians have curious names.

¹³ Rather curiously, it took quite a long time to realize that the best way to establish a thermometric scale is by fixing *two* points. And, by the way, the boiling point of water is surely not a very accurate one (surely much worse than the body temperature of a healthy person).

tion can then be used to check the accuracy of thermometers. In fact, while the readings for water and alcohol thermometers were rather wide of the mark, the results for mercury were very close to 50°C . Yet, even Regnault's method is not free from suspicion, since, to state that the temperature of the mixture is *exactly* 50°C , we have to assume that the specific heat of water does not change with temperature. Today we know that this is quite a good assumption, since the specific heat of water increases by less than 0.05% between the freezing and the boiling point but, again, how could Regnault be so sure about it? Besides, while it is not easy to prepare water at the freezing point by having it in equilibrium with ice, having “boiling” water at exactly 100°C is experimentally much harder. Notice however that Regnault made a big step forward. He made *two* separate assumptions, that mercury expands linearly with T and that the specific heat of water does not depend on T , and found that they are experimentally *consistent*, which gives much more confidence in the validity of *both* of them.

This short story of thermometry, of which we gave only a very partial account, tells us something more general about the way science actually works: very often, we trust a measurement because we trust the *law* on which it is based. In fact, we shall see that the thermodynamics of a “black body” will provide us with a new way to measure temperature from the spectrum of the radiation emitted by an heated body. You probably know that the whole universe is pervaded by a cosmic background radiation (CBR) that corresponds to that emitted by a black body at about 2.73 K, which provides the most important confirmation of the Big Bang theory (which is not just a TV show). Today, we trust the law of black body radiation so much that even deviations from complete isotropy of the CBR of the order of 1 part over 10^5 are investigated to test the origin and evolution of cosmological structures. ♡

Summing up, temperature is the first of those “collective” properties that specifically pertains to a thermodynamics system as a whole, basically because it quantifies mutual thermal equilibrium. To give it an absolute meaning, however, we need to introduce other physical quantities that characterize the system and, more than that, find how these quantities vary when we change its equilibrium state. We better start from concepts that are already familiar from mechanics.

1.3 Work and internal energy

We know from mechanics that the energy of a system of particles can be changed by doing work on it: to this aim, internal forces are totally ineffetual and, in the absence of *external* forces, the total energy of the system is conserved. Energy is surely not the only quantity that does not change in time: in fact, there are many, many more “integrals of motion”.¹⁴ Three of

¹⁴ A system with f degrees of freedom, a concept we shall revisit in the next chapter, has actually $2f - 1$, integrals of motion, most of them corresponding to quantities which can hardly be given a physical meaning, except in special cases.

these constants of motion are however very special: energy, linear momentum, and angular momentum. In the absence of external forces and torques, their conservation is in fact a consequence of the invariance of the laws of physics with respect to translations and rotations in space and time.¹⁵ However, if we just consider systems whose center of mass is kept fixed, and which are not in a state of rigid rotation around the center of mass, conservation of linear and angular momentum are trivial, hence energy takes on a central role

Knowing, as we do now, that a macroscopic body is nothing but a collection of individual particles, it is then natural for us to guess that a thermodynamic system has an “internal” energy. However, this was not obvious at all at the time thermodynamics was developed, when even the very existence of those atoms and molecules that most chemists already took for granted was questioned by several leading physicists. In fact, the concept of internal energy was introduced in thermodynamics to answer a very basic question: what of that fraction of the mechanical work done on a system which does not lead to an overall motion of the body? Since energy conservation is such a useful concept in mechanics, to avoid “emptying the baby out with the bath” we might conceive that energy is stored in the system in some unknown form. Then, performing/extracting work on/from a system would amount to moving it to a new, well-defined equilibrium state that differs by the value of the internal energy.

Yet, things are not so easy, because the new state reached by a thermodynamic system depends in general on the *specific way* work is performed, namely, on the kind of transformation the system undergoes. Luckily, there is a way to circumvent the problem: if the system is thermally *isolated*, that is, fully enclosed in those adiabatic walls that do not allow thermal equilibrium with the external environment to be reached,¹⁶ the final equilibrium state is unique. Then, if an infinitesimal amount of work δW is done on the system, the internal energy increases as $dE = \delta W$, where dE is a true mathematical differential because the internal energy is a well-defined *state function*, whose change depends *only* on the initial and final thermodynamic conditions of the system. Before discussing what happens in a generic transformation, it is however useful to quantify the work made on a system in adiabatic conditions by changing its “geometry” or, more generally, some of the external parameters whose values determine the state of the system.

¹⁵ This is because of a general theorem by Emmy Noether, stating in simple words that when the laws of physics are invariant with respect to changes of a continuous parameter, there is a physical property associated with this “symmetry” which is conserved.

¹⁶ Again, we are not requiring the walls to be *perfectly* adiabatic. What we mean is thermal equilibrium is reached over time scales which are much longer than those of the processes we are interested in.

1.3.1 Work by compression or expansion

The simplest way to make work on a system is changing its volume V . This is in fact the only way to transfer mechanical energy to a simple gas enclosed in a hermetic container provided with a movable piston. From elementary physics we already know that in this case the work done on¹⁷ the gas when its volume changes by dV is simply

$$\delta W = -PdV, \quad (1.2)$$

where, in general, P is the *external* pressure, namely, the force per unit area applied to the system from the outside. In fact, P coincides with the thermodynamic pressure of the system only when compression or expansion take place *quasi-statically*, by which we mean very slowly compared to the internal equilibration times of the system. If volume changes take place very rapidly, indeed, there might not even be a well-defined value of pressure for the whole gas: think for instance to what happens when air is blown in an organ pipe. For liquids and solids, which are barely compressible, there are however other geometrical parameters that can be varied besides the volume.

1.3.2 Elastic and plastic deformation

If a wire of length ℓ , one of whose ends is fixed to a wall, is stretched by $d\ell$ applying a force T to the other end, the work done on the system is $\delta W = +Td\ell$, which in this case has the simple structure of a force times a displacement. If the force is not too large, the wire reacts elastically: provided that the stretching is not too rapid (which would generate a longitudinal wave propagating along the wire), the wire is in uniform tension, and this tensile force coincides with T . If released, the wire comes back to its initial length and the internal tension vanishes. Tension is therefore a negative pressure: while positive pressure resists compression, tension opposes elongation. Yet, any wire, whatever the material is made of, has a yield stress, namely, a maximum value of the applied force beyond which the wire cannot stand the applied stress anymore, and stretches irreversibly without being able to recover its rest length. In other words, the internal tension relaxes, and the effect of the work done by T is a *plastic* deformation of the wire.

This simple example is easily generalized. Every solid, subjected not only to an external traction or compression force, but also to a shear stress, resists elastically up to a point, but then yields plastically. Usually, plastic deformation entails a volume change too, but this does not account for the whole amount of work performed by the external forces, which to a large extent

¹⁷ Note that we take the work as positive when done *on* the system, while in engineering the opposite convention is often used.

is used to change the *shape* of the material. In fact, the volume change is negligible for those materials, the elastomers (rubbers), which display an extremely large elasticity. The volume change of a solid is quantified by the *Poisson's ratio* ν , which is negative ratio of transverse to axial strain (with respect to the direction along which the compressional or tensile stress acts). For $\nu \simeq 0.5$ there is practically no volume change, which is the case of elastomers. Conversely, when $\nu \simeq 0$, the solid compresses along the direction of the applied force without expanding in the other directions, or vice versa: cork is arguably the only natural material that has this property, which is of course very useful when bottling wine. Recently, curious materials that, when squeezed along a direction, do the same along the transversal ones too (hence with a *negative* Poisson's ratio), have even been devised.

1.3.3 Work and surface tension

Fluids, which have no definite shape, cannot of course undergo elastic and plastic deformations. However, liquids have a free *surface*, which is another geometrical parameter that can be modified. In fact, there is an energy cost associated with the formation of a liquid surface. If you take a close look at those little insects known as “water striders”, “pond skaters”, or even “jesus bugs”, which live right *on* the surface of still and clean ponds, you may realize that, to sustain them, the water surface has to behave like an inflated balloon in tension. Similarly, if you observe a drop forming at the mouth of a slowly-dripping faucet, you can notice that the drop, even if subjected to its own weight, remains attached to the faucet for some time without falling, as if it were “glued” to the tap. Furthermore, when a brush is dipped in water and then taken out, its bristles stick together because of the water film that wets them. In all these cases, the surface of water behaves as an elastic membrane. To quantify the tension of this membrane, suppose we cut the liquid surface with an ideal “molecular lancet”, which breaks once and forever the hydrogen bonds between the water molecules lying on the opposite sides of the cut (see Fig. 1.1). Similarly to what happens to an inflated balloon, the two edges of the cut would rip apart, pulled by a total force that is obviously stronger the longer a cut is made. We can then write

$$F = \sigma \ell,$$

where ℓ is the length of the cut, and σ is called the *surface tension* of water, or, more generally, of the investigated liquid.

With this definition, σ is a force per unit length. However, it is physically more meaningful to regard surface tension as an energy per unit area. In fact, consider a film of liquid, attached as in Fig. (1.1) to a frame having

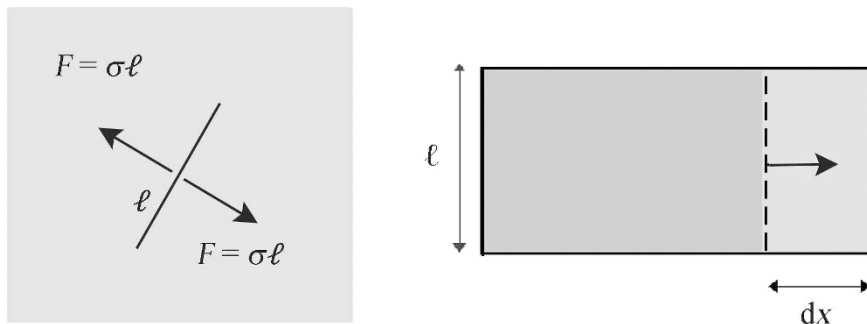


Fig. 1.1 Surface tension as force per unit length, or energy per unit surface

one movable barrier.¹⁸ To stretch out the film, we have to pull the movable barrier with a minimal force $\sigma\ell$, where ℓ is the side length of the film. Hence, the work needed to stretch the film by an amount dx is $\delta W = \sigma\ell dx$, namely,

$$\delta W = \sigma dA, \quad (1.3)$$

where dA is the increase of the film surface area. Therefore, surface tension may be regarded as the energy required to create a unit surface of the liquid, and can therefore be indifferently measured in N/m or in J/m². To be more accurate, σ is the energy per unit area of the boundary between a liquid and its saturated vapor, but its value does not appreciably change if we consider the boundary between the liquid and the surrounding air. In standard temperature and pressure conditions, water has for instance a surface tension with air of about 72×10^{-3} N/m, whereas σ is 3-4 times less for hydrocarbons, because the forces keeping together the molecules of non-polar liquids are much weaker. For the opposite reason, the values of σ for liquid metals are much larger: mercury for instance has $\sigma \simeq 0.5$ N/m. It is useful to point out that these values refer to very clean interfaces: the presence of even small amounts of adsorbed impurities can markedly reduced σ (as pointed out in footnote 18, this is particularly true in the case of adsorbed surfactants).

We have defined surface tension as the energy cost per unit area to make an interface between a liquid and a gas. But, of course, there also be a cost to make an interface between two different liquids. Or, surely, there is also an

¹⁸ Of course this is a kind of ideal experiment. Making a free-standing liquid film is actually possible, but this requires using special “protecting” molecules called surfactants. This is what happens for instance when we blow soap bubbles, which are in fact thin spherical liquid films stabilized on both sides by surfactant (soap) molecules. Soap not only reduces the surface tension of water, but also acts as a “healer” that rapidly dresses the wounds that start to form in the film.

energy cost to make a solid surface.¹⁹ In general, indeed, atoms or molecules that lie at the interface have different interactions with the surrounding than those in the bulk of a material. For instance, those molecules lying on the surface of water in contact with air are attracted just by the molecules lying below the surface, namely, by half the molecules they interact with when they are in bulk.²⁰ These molecules are more “costly” because they give a lesser contribution to the cohesive energy of the system. In general, we speak more properly of the *interfacial tension* between two condensed media, reserving the expression “surface tension” only for the interface between a liquid (or a solid) and the gas. Two very important effects due to surface and interfacial tensions are discussed below.

Laplace pressure. With what we said about surface tension, it is not too hard to understand why, provided they rise slowly, air bubbles in champagne are spherical: for a fixed volume, this is of course the shape that minimizes the contact area between water and air. What is arguably less familiar is that pressure in the bubble is slightly larger than in the surrounding liquid. This is easier to grasp thinking of our elastic analogy: to inflate a balloon, you must blow inside it, which just means increasing the pressure inside. Let us quantify this pressure difference $\Delta P = p_{in} - p_{out} > 0$, called the *Laplace pressure*, by noticing that ΔP must counterbalance surface tension, which forces the bubble to shrink in order to reduce its interface with water. The bubble must then be in mechanical equilibrium under the action of these two forces, which implies that the bubble energy $E(r)$ must be minimal with respect to the bubble radius r . Therefore, the work required to increase the surface by $dS = 8\pi r dr$ must be equal to the work $-\Delta P dV = \Delta P(-4\pi r^2 dr)$ done by the pressure forces. From $\partial E(r)/\partial r = 0$ we have $-4\pi r^2 \Delta P + 8\pi r \sigma = 0$, which gives

$$\Delta P = \frac{2\sigma}{r}. \quad (1.4)$$

Note that the larger the Laplace pressure, the smaller the size of the bubble. For instance, while the internal pressure of an air bubble with a radius of 1 mm is larger than the pressure of the surrounding water by only about $140 \text{ N/m}^2 = 140 \text{ Pa}$, which is just 0.14 % of the atmospheric pressure, the Laplace pressure for a (hypothetical) micro-bubble with $r = 1 \mu\text{m}$ would be approximately equal to the pressure your eardrums feel 14 m under water! The expression for the Laplace pressure can be easily generalized to the case of a non spherical surface, but Eq. (1.4) is all we shall need to discuss problems like the onset of water boiling.

¹⁹ In this case, the energy cost usually depends also on the orientation of the surface with respect to the crystalline axes of the solid.

²⁰ A simple reasoning should then make you understand why the surface of a liquid (in gravity) is horizontal.

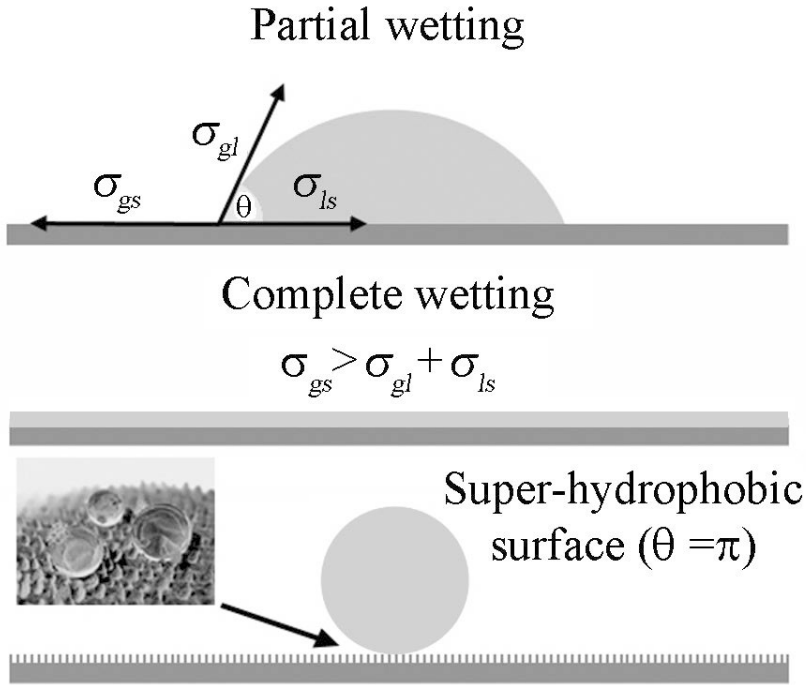


Fig. 1.2 Wetting phenomena

Wetting. As mentioned in footnote 19, in the case of solids the concept of surface energy is more complicated. Although in the following we shall not deal very much with solid surfaces, it is nevertheless useful to mention a very important effect involving the interface between a liquid and a solid: *wetting*. By saying that a liquid “wets” a solid surface, we mean that the fluid spreads on the surface as a thin coating film. Conversely, if the liquid arranges on the surface as separate droplets, we speak of *partial wetting*. By just thinking that this is what we do *not* want to happen when we paint walls, yarns, or car bodies, you can realize how important wetting is in practical applications.

Let us then see under what conditions a liquid wets a solid by finding the shape of a drop on an ideally flat (and clean) surface in partial wetting conditions. A droplet lying on a horizontal surface is a spherical dome, with a height that depends on the contact angle θ that the liquid makes with the plane (see Fig. 1.2).²¹ What fixes the contact angle? The contact rim of the droplet is subjected to three forces per unit length of the rim. First, we have the interfacial tension σ_{ls} , which acts to decrease the contact area between liquid and solid. This force is contrasted by the interfacial tension σ_{gs} , which

²¹ We suppose that the droplet is small enough to neglect gravity effects, which lead to some “flattening” of the shape.

tends to reduce the gas/solid interfacial area. Finally, we also have σ_{gl} , practically coincident with the surface tension of the liquid, which is directed along the tangent to the droplet at the contact rim. Mechanical equilibrium requires the sum of the components along the horizontal to vanish. Hence, we must have $\sigma_{ls} + \sigma_{gl} \cos \theta = \sigma_{gs}$, yielding a contact angle at equilibrium

$$\cos \theta = \frac{\sigma_{gs} - \sigma_{ls}}{\sigma_{gl}}. \quad (1.5)$$

By increasing the term at the r.h.s., θ decreases, until it vanishes for $\sigma_{gs} - \sigma_{ls} = \sigma_{gl}$. When $\sigma_{ls} + \sigma_{gl} \leq \sigma_{gs}$, the energy cost to form *two* interfaces, one between solid and liquid and the other between air and liquid, becomes lower than the cost of forming the original air/solid interface, hence the liquid spreads on the surface as a film that fully wets the solid. The more the film spreads, the thinner it gets, till its thickness becomes in principle of the order of the molecular size. In practice, however, the film spreading gets slower and slower, and is strongly influenced by the roughness of the surface and by the presence of impurities. The *dynamics* of wetting is actually a very complex phenomenon, and still an active research subject.

When does a liquid wet a (flat and clean) solid surface? This primarily depends on the nature of the solid material. *All* liquids fully wet those solids that are kept together by ionic, metallic, or strong covalent bonds (the so-called “high-energy” solids), while wetting is usually partial on “low-energy” solids such as molecular crystals or plastics.²² In several technological applications, it is however useful to engineer surfaces that are *not* wetted by a fluid, which, if the fluid of interest is water, means making surfaces that are as water repellent as possible. For materials like plastics, the contact angle with water can often exceed 90° , a situation where the drop is hemispherical. These are called *hydrophobic* surfaces, while in the opposite situation, $\theta < 90^\circ$, surfaces are dubbed *hydrophilic*. Although surfaces with $\theta \simeq 180^\circ$ may in principle exist, even on very hydrophobic plastics like polytetrafluoroethylene (PTFE) contact angles do not usually exceed 120° . Yet, this does happen in nature for the leaves of some aquatic plants like lotus, on which rain forms perfectly spherical drops. The secret of these apparently sleek leaves is that they are not smooth at all. Their surface is in fact covered by a very fine “barb” of hairs having the size of a few micrometers, where the raindrops rest without touching the underlying surface. In the past few years, noticeable efforts have been made to obtain artificial “superhydrophobic” surfaces, mimicking the microstructure of lotus leaves.

²² Whether or not a liquid wets is related to the difference in polarizability between the solid and the liquid. For a simple discussion, see the book by P.-G. de Gennes *et al.* cited in the selected readings.

1.3.4 Electric and magnetic work on solids

So far, we have dealt with different kinds of *mechanical* work, made by changing a geometrical parameter of the system like its volume, surface, or length. External forces of a different nature can however change the energy of the system by varying some *internal* physical properties. For instance, work can be made on dielectric solids by applying an electric field that *polarizes* them. The external field can polarize the dielectric either by aligning the intrinsic dipole moments of polar molecules, or by inducing a dipole on atoms and molecules that, in the absence of applied field, do not possess any. Consider then a slab of dielectric material sandwiched between the metallic plates of a capacitor. Calling ℓ the thickness of the slab, A its surface area parallel to the plates, and $\Delta V = E\ell$ the potential difference applied to the capacitor, the work done to charge the plates of dQ is $\delta W = E\ell dQ$. On the other hand, $Q = \mathcal{D}A$, where $\mathcal{D} = \epsilon_0 E + \mathcal{P}$ and $\mathcal{P}$ are the magnitudes of the “electric displacement” and polarization vectors appearing in Maxwell’s equations. Since $\mathcal{P}$ is the polarizability per unit volume of the dielectric, introducing the *total* polarization $P = V\mathcal{P} = A\ell\mathcal{P}$ of the dielectric, namely, the sum of all dipole moments in the dielectric slab, we have

$$dQ = \left(\epsilon_0 dE + \frac{dP}{V} \right) A \implies \delta W = V\epsilon_0 E dE + E dP.$$

The first term at the r.h.s. is just the work made on the capacitor in the absence of the dielectric slab (namely, when nothing is placed between the plates). Hence, the work properly done on the dielectric is just given by the second term

$$\delta W = E dP. \quad (1.6)$$

In the following chapters, where we extensively deal with the magnetic properties of matter, we shall see that many materials can be magnetized by an external field, sometimes retaining a finite magnetization even when the external magnetic field is switched off.²³ Like electric fields, also magnetic fields do work on a system by changing its magnetization. Consider an electromagnet, made of a coil of N loops wrapped around a magnetizable cylindrical core of section A and length L , and connected to a battery. If the electric current in the coil is altered, for instance by means of a rheostat, the magnitude of the magnetic field (or “magnetic induction”) B generated by the solenoid changes, which in turns produces an additional electromotive force (e.m.f) in the coil $\mathcal{E} = -NA(dB/dt)$. The work done *by* this self-induced e.m.f. to transport a charge dQ along the circuit is a $\mathcal{E}dQ$. Hence, the work done *on* the electromagnet is

²³ If you are not very familiar with magnetic fields (as it often happens with students in engineering), you will find an introduction to magnetic forces in Chap. 3, where the order of magnitude of typical magnetic fields is also discussed. So, just wait a bit!

$$\delta W = -\mathcal{E}dQ = NA \frac{dB}{dt}dQ = NAidB,$$

where i is the self-induced current. Let us then introduce the magnetic field $H = (N/L)i$ generated by the solenoid, related to B by $B = \mu_0(H + \mathcal{M})$, where $\mathcal{M}$ is the magnitude of the magnetization vector. The latter is given by $\mathcal{M} = M/V$, where M is the *total* magnetic moment of the material and $V = AL$ the volume of the core. We have then

$$\delta W = V\mu_0 HdH + \mu_0 HdM.$$

The first term is again the work done on the coil in the absence of the magnetic core. Since $\mu_0 H = B_0$ is the magnetic field outside the core, the work done on the magnetic material constituting the core is

$$\delta W = B_0 dM. \quad (1.7)$$

1.3.5 Generalized forces and displacements

All the previous examples show that the work done on a system has the general form

$$\delta W = \sum_i X_i dx_i, \quad (1.8)$$

where the differentials dx_i play the role of *generalized displacements*, and the quantities X_i that of *generalized forces*. For an isolated system, doing or extracting work is the only way to change the internal energy E of the system, so we can write $dE = \delta W$. We can then identify

$$X_i = \left(\frac{\partial E}{\partial x_i} \right)_{x_{j \neq i}} \quad (1.9)$$

where the subscript means that all the parameters x_j with $j \neq i$ are kept constant. Note that, while parameters x_i like volume, surface, polarization, magnetization are *extensive* variables, namely, they are proportional to the number N of molecules in the system, the generalized forces are *intensive* quantities, which do not depend on N (for instance, if we double the size of the system, its pressure does not change). Not all intensive variables can however be regarded as generalized forces. Examples of the latter are for instance those quantities obtained by taking the ratio of an extensive parameter to the volume of the system, like the number of particles per unit volume, or *number density*, $n = N/V$.

1.4 Entropy, the movie star

We have already encountered temperature as one of those collective variables which are distinctive of thermodynamics. In this section, we make acquaintance with a new character of this kind, which is bound to become our “movie star”, or, if you prefer, the clef that will allow us to read the score where the notes and chords of statistical mechanics are written: entropy. We have a long way to go before we can fully appreciate its deep meaning, but this first brief encounter with entropy will already allow us to look at temperature in a new perspective.

1.4.1 *The nature of heat*

The concept of entropy in thermodynamics originates from considering systems that, instead of being isolated, can exchange energy with the surrounding through diathermal walls. This “invisible” energy exchange, taking place without any work done by or on the system, is what we call heat, and a system that can exchange heat (but not matter) with the surrounding is called a closed system. Heat exchanges are what allow a closed system to attain equilibrium with the surrounding, usually regarded as a reservoir, namely, as a much larger thermodynamic system at fixed temperature T . As we already mentioned, in the presence of heat exchanges the work done on (or performed by) a closed system usually depends on the specific transformation leading from an initial state A to a final state B , and cannot anymore be evaluated as the difference of the values that a state function attains in B and A .²⁴

Giving up energy conservation would however be very drastic and inconvenient, so we better retain the idea that the system does have a property we still call internal energy, whose value is completely determined by the thermodynamic state of the system. This can be done by assuming that E is exchanged with the surroundings either through work performed on or by the system, or by the transfer of heat. Writing an infinitesimal heat exchange by δQ (which will not be an exact differential either), this means that

$$dE = \delta W + \delta Q,$$

where once again we take δQ as positive when absorbed by the system. Extending this result to finite exchanged of work and heat, we obtain the First Law of thermodynamics,

$$\Delta E = W + Q, \tag{1.10}$$

²⁴ In other words, δW is not an exact differential: this is why, right from the start, we have not written it as dW

which in fact is basically a *definition* of heat: to avoid some internal energy to appear or disappear mysteriously, we introduce a kind of “crutch”, heat exchange, which balances the budget.²⁵

1.4.2 The Second Law: an offense to equal opportunities

The First Law basically therefore states that heat, like work, is a form of energy exchange. The Second Law that we are going to introduce, however, leads to the rather surprising but unavoidable conclusion that, for what concerns energy exchanges, heat is then a less valuable “currency” than work: work can be *entirely* converted into heat, but the reverse is not true. This became clear by investigating the operation of heat engines, which are meant to transform heat into mechanical work by performing a thermodynamic cycle.

♡ Reaching this conclusion, however, took a long time. We do not find hard to believe that work can be completely converted into heat: the simplest way to do this is by dissipating all work through friction. Yet, this is only because the idea that heat is microscopic energy transfer and the principle of energy conservation are nowadays unquestionable. Up to the second half of the XIX century, however, the nature of heat was controversial: in fact, many outstanding scientists like Lavoisier and Laplace considered heat as a “subtle fluid”, the *caloric*, that passes from hot to cold bodies. In fact, although smart observations by Benjamin Thompson (Count Rumford) had already cast relevant doubts on the conjecture of heat as a substance that can be just transferred, but not generated or lost, the theory of caloric was rather successful. For instance, it led Laplace to correct Newton’s expression for the velocity of sound, and to develop the equation bearing his name, that describes the temperature field associated with heat transport.

A strong supporter of the theory of caloric was Sadi Carnot, a physicist and military engineer often regarded as the “father of thermodynamics” because of his fundamental work, published in 1824, that provided the bases for an abstract theory of the operation of heat engines. In fact, Carnot believed that a heat engine, like the steam engine developed by James Watt almost half a century before, worked more or less as a waterwheel that uses water falling from a higher to a lower level, which for a heat engine corresponds to higher and lower temperatures, to produce mechanical work. In this perspective, however, the *whole* amount of heat absorbed by the engine from the hot source is transferred to the colder reservoir, exactly like it happens for falling water: generation of work is only the “side effect” of the *passage* of heat through the machine.

The celebrated experiments by James Prescott Joule, which allowed for the definition of a “mechanical equivalent” of heat, drove William Thomson (Lord Kelvin) to drastically modify Carnot’s interpretation. Eventually, Kelvin’s investigation led to a full success of

²⁵ Note that at the left-hand side we have the internal energy *difference*. In thermodynamics (but not, as we shall see, in statistical mechanics), the energy of a system is indeed defined up to an arbitrary constant.

the “mechanical” theory of heat, and to the general principle of energy conservation (the First Law), enunciated by Rudolf Clausius in 1850. ♡

Reflecting on Carnot’s ideas, Kelvin reached a first important conclusion: the operation of heat engines allows to define an *absolute* temperature, which is free from most of the conceptual problems associated with empirical temperature scales. In fact, the operation of the “ideal” heat engine devised by Carnot depended only on the amount of heat $|Q_h|$ absorbed from a hot source and on the amount $|Q_c|$ released to a colder reservoir, which may simply be the surrounding environment, and not on the thermodynamic system acting as an engine.²⁶ Thus, Kelvin proposed to *define* the ratio of the absolute temperatures of the hot and the cold source as

$$\frac{T_h}{T_c} = \frac{|Q_h|}{|Q_c|}. \quad (1.11)$$

A very interesting consequence of this assumption is that T_h and T_c must be of the same sign, hence the absolute temperature must be taken as a number $T \geq 0$. This introduces the concept of an *absolute zero* of temperature, which empiric thermometric scales have not. The question is, how can we *measure* absolute temperature? The solution to this crucial problem resulted from a fruitful theoretical and experimental collaboration between Kelvin and Joule, including a seminal study of the free expansion of a gas, which lead them to conclude that absolute zero very arguably coincides with the value suggested in 1702 by Guillaume Amontons, who fixed the zero point by extrapolating the linear dependence of P on T for gases, which he had discovered, to the point where pressure vanishes. At the time of the Kelvin and Joule investigation, that point was already known to be approximately -273°C , and taking a degree as in the Celsius scale, the absolute temperature scale is fixed.

Once heat is regarded as a form of energy, there is however a much more important consequence that can be derived by contrasting the operation of heat engines with the First Law. Even in an ideal Carnot engine, which provides, as we shall shortly see, the most “efficient” way to transform heat into work, part of the energy extracted from the hot source must be given away to a source at temperature $T_c < T_h$. By the First Law, the work made by the engine is therefore $W = |Q_h| - |Q_c|$. It is then impossible to devise an engine that works cyclically and produces no other effect than transforming some heat into work, which is a first way to state the Second Law of thermodynamics.²⁷ Already in this form, the Second Law contains the germ of the arrow of time, since it states that a process like the integral transformation of heat in work, which is the exact reverse of a fully feasible one, cannot take place. This is manifest if the Second Law is recast in the equivalent form given to it

²⁶ Absolute values for Q_h and Q_c are used to avoid any confusion with the convention on the signs of exchanged heat.

²⁷ Although it basically summarizes the conclusions of Kelvin’s investigation, this statement of the Second Law is due to Max Planck.

by Rudolf Clausius in 1850, stating that there are no spontaneous processes whose sole result is the transfer of heat from a cold body to a hotter one.

To be more quantitative, we have to clarify what we really mean by an “ideal” Carnot engine. This requires introducing the concept of *reversible process* as a thermodynamic transformation between two equilibrium states A and B that, given suitable conditions, can take place *spontaneously* in both directions, $A \rightarrow B$ and $B \rightarrow A$. At first, this definition may sound baffling, since thermodynamics pretends to account for natural processes, which are mostly one-way. Yet, many processes become *almost* reversible if the transformations are performed “mildly enough”. By this we mean that work is done quasi-statically, so that along the transformation the system passes through a sequence of quasi-equilibrium states. For instance, in a quasi-static compression of a gas the applied external pressure must increase very gradually from the initial to the final value. Regarding this sequence of quasi-equilibrium states as a temporal chain, reversibility requires that external or internal parameters are changed very slowly compared to the intrinsic relaxation times of the system. Experience tells us that heat transfer between two bodies is faster the larger is their temperature difference. Hence, another condition for a process to take place reversibly is that, while a system S is exchanging heat with a reservoir R , the temperature difference between S and R must be small.

By an ideal Carnot engine we mean therefore a device operating cyclically and reversibly between only two reservoirs. It is only for this idealized machine that the equality (1.11) holds. If we define the efficiency of the energy as the ratio η between the work done by the engine and the heat absorbed from the hot reservoir, using the First Law and Eq. (1.11) we have (with T in kelvin)

$$\eta = \frac{|W|}{|Q_h|} = 1 - \frac{|Q_c|}{|Q_h|} = 1 - \frac{T_c}{T_h}, \quad (1.12)$$

1.4.3 Entropy, the scales of heat

In the case of work, we managed to relate δW to the change of some extensive geometrical parameter like volume, or of an internal quantity like polarization. Can we devise something similar for δQ ? This can be done by generalizing Eq. (1.11) to a generic reversible cycle where heat can be exchanged with any number of reservoirs

$$\oint_{rev} \frac{\delta Q}{T} = 0, \quad (1.13)$$

which is known as *Clausius theorem*, and where of course each infinitesimal heat exchange δQ is taken with the correct sign. Vanishing of the integral over a cycle means that, when all transformations are reversible, the ratio of the

exchanged heat to the (absolute) temperature is in fact an *exact differential*. Hence, in a reversible transformation connecting two equilibrium states A and B , the total amount of this “rescaled” heat

$$\int_A^B \left(\frac{\delta Q}{T} \right)_{rev} \quad (1.14)$$

can be regarded as the difference $\Delta S = S(B) - S(A)$ of a new state variable S , which is called *entropy*. The entropy change in an infinitesimal reversible transformation is then given by:

$$dS = \left(\frac{\delta Q}{T} \right)_{rev} \quad (1.15)$$

It is worth stressing again that the entropy change has the *same* value ΔS even if state B is reached from state A by a transformation which is *not* reversible, as those we shall shortly discuss (the amount of exchanged heat will instead be different): we need a reversible transformation just to *evaluate* ΔS .

Using (1.15), the First Law can be written:

$$dE(x_i, S) = \sum_i X_i(x_i, S) dx_i + T(x_i, S) dS, \quad (1.16)$$

where T and the X_i , which are *intensive* variables, are regarded as functions of S and of the x_i , which are *extensive*. Eq. (1.16) suggests that heat exchange can be accounted for by regarding dS as a new, purely “thermal” displacement, and temperature as the associated generalized force. To grasp the physical meaning of this idea, consider a gas in a container C , separated in two parts C_1 and C_2 by a movable piston. If the temperature in the two parts are equal, but the pressure in C_1 is larger than in C_2 , the gas in C_1 expands, displacing the piston until the pressure values in C_1 and C_2 become equal. Pressure difference is then exactly the force that “displaces the volume”. Similarly, if C_1 and C_2 are separated by a fixed diathermal wall, and the temperature of the gas in C_1 is higher than in C_2 , heat (and therefore entropy) will pass from C_1 to C_2 until thermal equilibrium is reached. Temperature difference is then the force that “displaces entropy”. From (1.16) we have

$$T = \left(\frac{\partial E}{\partial S} \right)_{x_i},$$

which yields a new meaning of temperature as the rate of change of internal energy with entropy. Since absolute temperature is a positive quantity, the entropy must be a monotonically increasing function of the internal energy of the system. We shall see that this has a very important microscopic meaning.

1.4.4 Irreversibility and the Second Law

The privileged status of macroscopic (work) with respect to microscopic (heat) energy exchanges entailed by the Second Law is already evidential of a time asymmetry in natural processes. Yet, it is by combining the Second Law with the observation that *spontaneous* processes are *not* reversible that we obtain the most revealing result, telling us *what* is the direction indicated by the arrow of time. Irreversibility may originate for instance from the dissipation of mechanical or electric work, from heat transport between reservoirs with a finite temperature difference, from the mixing of two molecular species, from chemical reactions. However, in spite of their different nature, all these processes share a common feature for what concerns entropy changes. Let us try and discover the hallmark of irreversibility with two “case studies”.²⁸

Mechanical irreversibility. Suppose we stir water contained in a glass which is in thermal contact with a large heat reservoir at temperature T (a pool, for instance). If we stop stirring, water will sooner or later stop turning. Experience tells us that this is for sure an irreversible transformation: we never see water getting energy from the reservoir and spontaneously start turning around! When water eventually returns to its initial state of rest at temperature T , its internal energy has not changed. Hence, from the First Law, all the work W done by the stirrer has been transferred to the reservoir as a quantity of heat $Q = W$. Entropy has not changed either, thus we can write $\Delta S_s = 0$, where the subscript s stands for “system”. However, the change of the entropy of the reservoir, ΔS_r , is surely positive, because it has *absorbed* heat. Since this heat has been absorbed at constant temperature T , we simply have $\Delta S_r = +Q/T$. Therefore, in this irreversible transformation, the *total* entropy of system plus reservoir $\Delta S = \Delta S_s + \Delta S_r = +Q/T$ has increased.

Thermal irreversibility. Imagine now a metal bar, regarded as the system, with two large masses joined at either ends playing the role of reservoirs S_1 and S_2 , which are kept at different temperatures $T_1 > T_2$. Within a very short time, the bar reaches a state characterized by a temperature profile linearly decreasing from T_1 , at the end in contact with S_1 , to T_2 at the other end. In this *stationary* (not equilibrium) state, the bar just transfers by thermal conduction to S_2 all the heat Q absorbed in a given time from S_1 . After the initial short transient, therefore, the entropy of the bar does not change anymore. Conversely, the entropy of S_1 *decreases* within the same time by $-Q/T_1$, while the entropy of S_2 *increases* by $+Q/T_2$. Since $T_1 > T_2$, for the total entropy change we have

$$\Delta S_s + \Delta S_1 + \Delta S_2 = \frac{Q}{T_2} - \frac{Q}{T_1} > 0,$$

²⁸ Two other situations, the free expansion of a gas and the mixing of two fluids, will be considered in detail in the next chapter.

In both the processes we examined, therefore, the *overall* entropy (system + surroundings) increases. This is a common feature of all irreversible transformations. Using the Second Law, it is not difficult to show that this is equivalent to state that, in any irreversible transformation, the entropy change for the system, which must be calculated along a *reversible* transformation connecting the same states, is *larger* than the sum of the heat exchanged with the surroundings, normalized to the temperature at which they take place:

$$\Delta S \geq \int_{irr} \frac{\delta Q}{T}. \quad (1.17)$$

When applied to a cycle, Eq. (1.17) gives

$$\oint \frac{\delta Q}{T} \leq 0,$$

a general form of Clausius theorem that is valid for any generic cycle containing both reversible and irreversible transformations. If we apply this result to the particular case of a system that does not exchange any heat with the surroundings, we see that in spontaneous internal processes entropy never decreases, and remains constant only if all internal transformations are reversible. This is the well-known principle of entropy increase for isolated systems, which gives up a precious hint: for an isolated system, the true equilibrium state must correspond to a *maximum* of entropy.

Therefore, the arrow of time aims in the direction of entropy increase. Namely, we can tell the future from the past (or, better, we give a meaning to these words) by spotting how the entropy of an isolated system changes. We have then found the prime suspect for the crime of murdering the reversible microscopic laws of physics, so it is worth putting ourselves in Sherlock Holmes' shoes and conduct a detailed investigation about entropy. Anticipating the verdict, we shall however judge the defendant not guilty, because no offence was in fact committed.

♠ **Speed and irreversibility.** We stated that a transformation, to be quasi-reversible, must in practice be performed very slowly, and it is in fact easy to see that speed clashes with efficiency. Suppose indeed that, like in the Carnot cycle, the engine operates between two sources, absorbing an amount $+|Q_h|$ from a reservoir at absolute temperature T_h , and giving back a part $-|Q_c|$ of this energy (which is required by the Second Law) to a colder source, for instance the environment, at temperature $T_c < T_h$. In a full cycle the entropy of the engine does not change, whereas for the two reservoirs we have

$$\frac{|Q_c|}{T_c} - \frac{|Q_h|}{T_h} \geq 0 \implies \frac{|Q_c|}{|Q_h|} \geq \frac{T_c}{T_h} \implies \eta \leq 1 - \frac{T_c}{T_h}.$$

The maximal efficiency is then obtained for a Carnot cycle where the engine, initially at temperature T_c , is first brought to T_h by a reversible adiabatic transformation, put in contact with the hotter reservoir to absorb Q_h at constant $T = T_h$, and finally brought back to T_c , where it releases Q_h isothermally to the colder reservoir. The fact is, if the

system is brought *exactly* to T_h and T_c , no heat exchange takes place, for the net heat flux between two bodies at the same temperature vanishes. To increase speed, we must then necessarily give up some efficiency.

Of course, we might be willing to pay this price if, by doing this, we increase the *power* generated by the engine, which can be taken as the ratio between the work the engine does and the time it takes to perform a full cycle. For the power not to vanish, the temperatures of the system when put in contact with the reservoirs must then be $T'_h < T_h$ and $T'_c > T_c$, yielding an efficiency

$$\eta \leq 1 - \frac{T'_c}{T'_h} = 1 - (1 + \delta)^2 \frac{T_c}{T_h},$$

where we put $T_h/T'_h = T'_c/T_c = 1 + \delta$. Assuming $\delta \ll 1$, namely, that T'_h and T'_c are not too different from the temperatures of the reservoirs, the maximum possible efficiency is, at first order in δ

$$\eta \simeq \eta^{id} - 2\delta \frac{T_c}{T_h},$$

where η^{id} is the efficiency of the ideal cycle. Because heat fluxes are to a good approximation proportional to the temperature differences $\Delta T_h = T_h - T'_h$ and $\Delta T_c = T'_c - T_c$, the time τ required to exchange the required amounts of heat with the two sources is proportional to

$$\tau \propto \frac{|Q_h|}{\Delta T_h} + \frac{|Q_c|}{\Delta T_c}.$$

On the other hand, since in a cycle the entropy of the system does not change, we have again $|Q_h|/T'_h = |Q_c|/T'_c$, and therefore

$$\tau \propto \frac{T'_h}{\Delta T_h} + \frac{T'_c}{\Delta T_c} = \frac{1}{\delta} + \frac{1 + \delta}{\delta} = \frac{2 + \delta}{\delta}.$$

Even if we assume that the adiabatic transformations take a negligible time (which of course will not be the case, because they would hardly be reversible) the delivered power $P(\delta)$ will be proportional to

$$P(\delta) \propto \frac{\eta}{\tau} \propto \frac{\delta}{2 + \delta} \left[\eta^{id} - 2\delta \frac{T_c}{T_h} \right].$$

Still at first order in δ , the maximum power is obtained for

$$\delta \simeq \frac{\eta^{id}}{4} \frac{T_h}{T_c} \implies \eta \simeq \frac{\eta^{id}}{2},$$

namely, the efficiency approximately halves.²⁹ ♠

1.4.5 The Third Law

Entropy, as we introduced it, is defined up to an arbitrary constant, like the internal energy. This means that, in principle, we are free to choose a reference temperature at which $S = 0$, because what really matters are *differences* in entropy. This is what is usually done in engineering applications, for in-

²⁹ It is not difficult to see that, for this result to be self-consistent ($\delta \ll 1$), we must have $T_c/T_h \gg 0.2$.

stance to compile steam tables. In 1906 Walter Nernst, reflecting upon the behavior of specific heat of materials at very low temperature (which we shall investigate in Chap. 3), was however led to conclude that, in the neighborhood of 0 K, all reversible thermodynamic transformations take place with no change of entropy, i.e., that $\lim_{T \rightarrow 0} \Delta S = 0$. Nernst soon realized that this makes impossible to reach the absolute zero in a finite number of steps, a common statement of the *Third Law* of thermodynamics that gives a “special status” to the observation that, in any cooling processes, the lower the attained temperature, the harder is to further cool. Like the other principles of thermodynamics, the Third Law must of course be regarded as a postulate: yet, it is a very peculiar postulate, which can hardly be given an intuitive meaning using our ordinary experience.

Another consequence of Nernst’s statement of the Third Law, which is surely more interesting for what follows, is that S attains a finite value S_0 at $T = 0$ K, which can then be taken as a reference value with an *absolute* meaning, although thermodynamics does not provide any *specific* values for S_0 . As we shall see, what is actually hidden behind the Third Law is the quantum nature of matter. This was first suggested in 1910 by Otto Sackur, who argued for a connection between the “new law of thermodynamics” proposed by Nernst and Planck’s assumption of energy quantization. It was this suggestion by Sakur that led Max Planck to restate the Third Law in a stronger form: *The absolute value of the entropy of a pure solid or a pure liquid approaches zero at 0 K*. Namely, for the equilibrium state of a pure substance, we can strictly take $S_0 = 0$.³⁰ The origin of Planck’s formulation of the Third Law will become clear in the next chapter.

1.5 Other characters (in order of appearance)

1.5.1 Free energy

In isolated systems, the internal energy plays the role of a thermodynamic “potential”, meaning that the work the system does or receives is equal to the change in E . In a closed system, where E is also affected by heat exchanges, the role of the internal energy is taken by a new thermodynamic potential, which we are going to meet. To introduce this new thermodynamic quantity, we notice that, for generic transformations, the Second Law allows us to write Eq. (1.16) as

$$\delta W \geq dE - TdS. \quad (1.18)$$

Let us then consider the quantity

³⁰ Note that Planck carefully speaks of a *pure* substance in *equilibrium* conditions. Nothing is said about the entropy of a mixture, or of one of those metastable glassy states we mentioned in Sect.1.1.1.1.

$$F = E - TS, \quad (1.19)$$

which is called the system's *free energy*.³¹ If then the system is kept at *fixed* temperature, we have $dF = dE - TdS - SdT = dE - TdS$, and therefore

$$-(\delta W) \leq -(dF). \quad (1.20)$$

This means that the work $-(\delta W)$ done *by* a closed system in isothermal conditions is always less or equal to the change in free energy. In fact, F is dubbed “free” just because it is that limited fraction of the internal energy which can be freely used to extract work from a system, a result clearly useful for the operation of heat engines or thermal machines.

But the key role played by F in closed systems becomes manifest when considering a thermodynamic system where internal processes take place *without* exchanges of work with the surroundings. In these conditions, Eq. (1.18) reads $dF \leq 0$, where the equal sign holds only when all internal transformations are fully reversible. Because this never happens, this means that internal processes lead to a decrease of the free energy. With the same argument used for the entropy of an isolated system, we then conclude that equilibrium is the state where F is *minimal*. Hence, free energy takes on the same role of “flag” for the equilibrium state of a closed system that entropy has for an isolated system. By investigating the microscopic roots of thermodynamics, we shall discover that, concealed behind the “facade” of free energy, there is another main figure. If entropy is the queen of the realm of thermodynamics, this character is her “Minister of Finance”, or more modestly the reliable book-keeper who balances the accounts in the microscopic world.

Finally, note that for reversible processes where both temperature and volume change, we have:

$$dF = -PdV - SdT,$$

or, using the more general expression for δW

$$dF(x_i, T) = \sum_i X_i(x_i, T)dx_i - S(x_i, T)dT, \quad (1.21)$$

where we make explicit that the natural independent variables for a closed system are temperature and the generalized displacements.

³¹ Often called *Helmholtz* free energy, to distinguish F from the so-called *Gibbs* free energy, which however we shall name differently.

1.5.2 Chemical potential

In Sect. 1.1.2, by considering the equilibrium between water and vapor in a sealed container, we saw that it is often useful to consider *open* systems that, besides work or heat, can also exchange mass with the surrounding, and which therefore contain a variable number of particles. However, what actually fixes the fractional amounts of water in liquid and vapor phase was an open question. Equilibrium between coexisting phases is in fact ruled by a new character that, for open systems, actually plays a co-starring role with entropy. For a one-component open system, the internal energy will explicitly depend on the number N of molecules. Since E is extensive, the additional contribution in N must be linear, so we can write

$$dE = \sum_i X_i dx_i + TdS + \mu dN, \quad (1.22)$$

where the proportionality coefficient μ between dN and dE is the *chemical potential* of the system. Eq. (1.22) can be easily generalized to the case of a system containing several molecular species by introducing a chemical potential for each single component:

$$dE = \sum_i X_i dx_i + TdS + \sum_j \mu_j dN_j, \quad (1.23)$$

where N_j is the number of molecules of the chemical species j .³²

The qualifier “chemical”, which some physicists may find annoying, comes from considering reactive systems, like two reagents A and B yielding C, where the number of molecules of each chemical species is of course variable. In fact, we shall see how the concept of chemical potential allows us to find the conditions for equilibrium in a chemical reaction. However, the *physical* meaning of the chemical potential goes well beyond its role in chemical equilibria. In fact, from (1.23) we see that μ takes on the role of a generalized force associated with mass transfer. For instance, in the example of phase equilibrium we considered, water evaporates until the chemical potential in the vapor phase is equal to that of the liquid phase. For our purposes, the concept of chemical potential will also be extremely useful to investigate very peculiar systems like the quantum gases, and also systems where the number of “particles” is *intrinsically* variable, which is the case of the thermodynamics of electromagnetic radiation. More generally, we can state that the chemical potential, together with entropy and free energy (or that stuff hiding behind F), are the real passwords to the world of statistical mechanics.

³² In chemistry, it is more customary using as variables the number of moles ν_j of each component. Eq. (1.23) then takes the form $dE = \sum_i X_i dx_i + TdS + \sum_j \tilde{\mu}_j d\nu_j$, where the coefficients $\tilde{\mu}_j$ are molar chemical potentials.

1.5.3 Extensivity and homogeneity

If we change the number of molecules of a system from N to λN , where λ is an arbitrary multiplicative constant, both its volume and its entropy, which are extensive variables, change accordingly: $V \rightarrow \lambda V$, $S \rightarrow \lambda S$. But the internal energy is an extensive variable too. Thus, regarding E as a function of S , N , and V (taken for the sake of simplicity as the only mechanical parameter), we can write

$$E(\lambda V, \lambda S, \lambda N) = \lambda E(V, S, N),$$

namely, E is a *homogeneous* function of first degree of V , S , and N .

Homogeneous functions have an important property, discovered by Euler, which is worth recalling. Consider a generic homogeneous function $f(x_1, \dots, x_N)$ of degree n , namely, a function satisfying

$$f(\lambda x_1, \dots, \lambda x_N) = \lambda^n f(x_1, \dots, x_N).$$

Deriving both members with respect to λ , we have

$$\frac{\partial f(\lambda x_1, \dots, \lambda x_N)}{\partial \lambda} = \sum_{i=1}^N \frac{\partial f(\lambda x_1, \dots, \lambda x_N)}{\partial (\lambda x_i)} x_i = n \lambda^{n-1} f(x_1, \dots, x_N).$$

Since this expression holds for *any* λ , we are free to choose $\lambda = 1$, obtaining

$$n f(x_1, \dots, x_N) = \sum_{i=1}^N \frac{\partial f(x_1, \dots, x_N)}{\partial x_i} x_i, \quad (1.24)$$

which shows that f is just a *linear combination* of the variables with coefficients that are $1/n$ times the partial derivatives of f with respect to each variable. In the case of a homogeneous functions of the *first* degree, this means that the expression for the total differential $df = \sum_i (\partial f / \partial x_i) dx_i$ holds for the integrated terms too:

$$f = \sum_i \frac{\partial f}{\partial x_i} x_i.$$

Hence, for the internal energy we have $E = -PV + TS + \mu N$, or, if these is more than one component³³

$$E = -PV + TS + \sum_j \mu_j N_j. \quad (1.25)$$

³³ More generally, if work is also done by changing other extensive parameters,

$$E = \sum_i X_i x_i + TS + \sum_j \mu_j N_j.$$

Taking now the differential of Eq. (1.25), using Eq. (1.22) we find

$$SdT - VdP + \sum_j N_j d\mu_j = 0. \quad (1.26)$$

This *Gibbs-Duhem relation* has many practical applications, for instance to evaluate how pressure depends on temperature in conditions where two different phases, such as liquid and vapor, are in equilibrium.

1.5.4 Some extras to suit all tastes

We consider again, just to simplify notation, a purely hydrostatic system, where work is done only by changing volume. Moving from an isolated system, where the independent variables are V and S , to a closed system, fully specified by V and T , the new thermodynamic potential F is obtained from E by subtracting TS , which is the product between the generalized displacement S and its associated force T . Note that, since $dF = -PdV - SdT + \mu dN$, in this new description S and T basically exchange their roles: T is the generalized displacement, and $-S$ the associated force. This is just an example of a *Legendre transformation*, where an exchange of the independent variables yields a new thermodynamic potential. Let us see some more possibilities.

Enthalpy. Chemical reactions are more easily performed at constant pressure than constant volume, since the volume of the reaction products is often very different from the volume of the reagents (think for instance of a reaction between two liquids yielding a gas!). If the reaction takes place with no temperature control, for instance in an adiabatic container, we still have an isolated system, whose volume is however not fixed. A thermodynamic potential suitable to describe these conditions, known in chemistry as *enthalpy* and usually indicated with H , is obtained by putting

$$H = E + PV, \quad (1.27)$$

so that

$$dH = VdP + TdS + \sum_j \mu_j dN_j. \quad (1.28)$$

The physical meaning of enthalpy becomes apparent when we consider processes taking place at constant pressure. We have indeed

$$dH = dE + PdV + VdP = \delta W + \delta Q + PdV + 0 = \delta Q,$$

namely, the enthalpy change coincides with the exchanged heat.

There is, however, a delicate point in the definition of enthalpy. Consider a thin layer of liquid: its internal energy does not only depend on volume

and entropy, but also on the extension of its free surface area A . Namely, we can write $E = E(S, V, A)$. Like we have just done by operating a Legendre transformation on the variables $P - V$, we can also think, therefore, of a thermodynamic potential that describes heat exchanges at constant *surface tension*,³⁴ by simply defining $H' = E - \sigma A$. Similarly, we can consider the properties of a magnetic material as a function of the applied field by defining $H'' = E - B_0 M$, which, like H' , can also be regarded as an “enthalpy”, since it is obtained from the internal energy by subtracting out the product of a generalized displacement times the generalized force associated with it. Hence, if there are several ways to perform work on a system, there are as many thermodynamic potentials that can play the role of an enthalpy.

Free enthalpy (Gibbs free energy). When a process instead takes place at both constant temperature *and* constant pressure (which is the most common experimental condition), we can introduce a suitable potential

$$G = F + PV = E + PV - TS \quad (1.29)$$

The thermodynamic potential G is often called the “Gibbs” free energy, to highlight its strict relation with the “Helmholtz” free energy F . However, since G can also be directly obtained from the enthalpy H as $G = H - TS$, which is the same Legendre transformation used to introduce the free energy from E , it is probably more revealing to call it *free enthalpy*, reserving the expression “free energy”, with no further specification, for F . Nevertheless, G plays exactly the same role for a system kept at constant pressure that F plays for a system at constant volume. Regarding G as a function of P , T , and of the number of particles N_i of the different species, we have then

$$dG(P, T, N) = VdP - SdT + \sum_j \mu_j dN_j, \quad (1.30)$$

It is worth recalling that P and T are intensive variables, so N is the *only* extensive independent variable in G . It is not difficult to see that this leads to a very simple relation between G and μ , which we shall later discuss.

Like for the enthalpy, if there are several ways of performing work on the system, there are as many possible definitions of G . In what follows, however, we shall reserve the expression “free enthalpy” only for the quantity defined by (1.29), while any other thermodynamic potential obtained from a “different” enthalpy will generically be called a Gibbs’ free energy. We might go on defining several other thermodynamic potentials for an *open* system, using then μ instead of N as an independent variable. It is however worth waiting a bit, since we shall amply deal with open systems in Chap. 5. Now that we have appreciated the power of thermodynamics in accounting for

³⁴ This can be experimentally achieved by using a setup called a Langmuir-Blodgett trough.

many evidences of our daily experience, it is time to have a first look to the microscopic world. Which means meeting a Scottish botanist, Robert Brown.

1.6 Brownian motion, the *Perpetuum Mobile*

To pay Brown a visit, we must forget all what we have learnt about the arrow of time, because we actually need to jump *back* in time to 1827, to face a crucial problem with the Second Law.³⁵ So, let's peep through the shutters of a lab in London, and watch Brown while he is observing under the microscope some grains of pollen suspended in water. The trouble is that these wretched specks are not keen at all to keep still under observation, but rather seem to suffer from a kind of Saint Vitus' dance, stirring and jiggling madly about before poor Robert's eyes. Besides, the paths followed by the pollen particles are extremely irregular and look remarkably unlike one to the other. At the time, it might have seemed easy to account for this. Most naturalists, in disagreement with physicists, believed that biological stuff possessed a kind of "spirit of life" that made it superior to inanimate objects, and *Brownian motion*, as we shall call the effect observed by Brown, could have been a direct manifestation of this elusive "spirit". But Brown, who was no physicist, but no fool either, carefully avoided jumping to this conclusion. And he was right, since it was easy for him to show that even specks of humble and totally inanimate dust displayed the same frantic behavior. In fact all kinds of particles having a size lower than, say, a few micrometers, which are called *colloidal particles*,³⁶ share, to a greater or lesser extent (the smaller they are, the more frantically they move), this hectic behavior.

Just a curious oddity in the early 19th century, Brownian motion became a truly puzzling enigma a few decades later, after Clausius and Kelvin had formulated the Second Law that put an end to any dreams of designing a perpetual motion machine of the "second kind":³⁷ no way, in the macroscopic world, *all* kinds of motion eventually dissipate as heat. And yet, these tiny particles never stop, like an endless musical canon (a *Perpetuum Mobile*, indeed). The origin of Brownian motion remained a mystery till 1905, when Albert Einstein gave it a full and brilliant explanation, which is also a fundamental proof of the atomic nature of matter.³⁸ Actually, at the beginning

³⁵ If you miss a time-machine, just ask *another* Dr. Brown to lend you his own DeLorean.

³⁶ A *dispersion* in a simple fluid of particles in this size range is named a *colloid* (although you shall often find the word "colloid" improperly used for a single colloidal *particle*).

³⁷ A perpetual motion machine of the second kind does not contravene the First Law (energy is conserved), but, since it is supposed to turn heat *completely* into work, it violates the Second.

³⁸ To tell the truth, the same result was simultaneously obtained by a great Polish physicist, Marian Ritter von Smolan Smoluchowski; but he had rather too knotty a name to become popular, and certainly did not have a natural presence for TV shows (at least when

of the twentieth century, not all scientists believed in the existence of atoms and molecules. Chemists obviously did, and they already knew a remarkable amount about how to play with them. However, many distinguished physicists held the opposing view that there was no real need of these little things to explain how the world worked. That is to say, they somehow managed to stick to the old idiom “what the eye doesn’t see, the heart doesn’t grieve over”. Einstein’s explanation of Brownian motion, however, not only necessarily assumes the existence of atoms, but also provides us with a way to estimate a quantity of primary importance for statistical physics, the Avogadro number.

The only other assumption made by Einstein is that atoms and molecules are in perpetual motion too, with a kinetic energy proportional to temperature that we shall call the *thermal energy*. Hence, a colloidal particle immersed in a fluid is unceasingly bombarded by these shooting nano-bullets that transfer a little of their own energy to it, so that, in a very short time, the kinetic energy of the particle will also be equal to thermal energy (neither more, otherwise the particle would give it back to the molecules, nor less, because then it would go on absorbing energy). This means that the particle, too, must be moving. But how? Each collision is a kind of “kick” (a feeble one, but still a kick) to the particle, kicks that come from all sides. On average, all these nudges even out, so there is no a preferred direction for the particle to move. Yet, as Einstein realized, if we look at a very short time interval this is not exactly true. There will always be a little loss of balance, causing the particle to perform an extremely irregular zigzag motion, similar to the staggering walk of a drunkard, which is called a *random walk*.

A random walk is a fundamental example of a *stochastic process*, namely, of a physical process that requires a statistical description. An introduction to random walks and to their relation with Brownian motion and diffusion processes is given in Appx. D, but, to fully appreciate it, you better wait until we learn a bit of statistical physics. This brief and qualitative introduction to Brownian motion, however, already tells us something disquieting. Acting as “mesoscopic” probes of the underlying molecular agitation (tiny, but still far larger than the atoms), the colloidal particles witness that, at the atomic scale, the Second Law is at stake, at least in the “absolutist” acceptance of thermodynamics. So, we better put on a mask, flippers and a wetsuit, and dive into the depths of the microscopic world.

compared with Einstein’s hanging tongue, as reproduced on a thousand T-shirts). Nevertheless, Smoluchowski is worth being remember, not only since his approach is at the roots of the modern investigation of Brownian motion in colloidal systems, which are discussed in Chap. 6, but also because next year (2017) falls the centennial of his death.

Additional readings

★ General references

♪ Ma SK, *Statistical Mechanics*, World Scientific, Singapore, 1985.

If you really wish to understand the key ideas of statistical physics, this is, in my opinion, an unsurpassed book, from which I often draw inspiration. In spite of its colloquial style and of the minimal use of mathematical formalism, however, it is anything but an elementary text, and requires the reader a consistent effort to grasp several delicate points Shang-Keng Ma expatiates on. Nevertheless, his discussion of the basic thermodynamic concepts presented in this chapter is simply superlative.

♪ Callen HB *Thermodynamics and an Introduction to Thermostatistics*, 2nd edition, Wiley, Singapore, 1985.

An authoritative book, where the foundations of thermodynamics, presented as a self-standing science, are extensively discussed with rigor and precision. A reference text for all scholars in statistical physics, but surely a bit too advanced for our purposes (and, in my opinion, a bit too formal).

♪ Zemansky MW, Dittman RH, *Heat and Thermodynamics*, 7th edition, McGraw-Hill, Singapore, 1997.

Another “classic” textbook in thermodynamics, written at an intermediate level and with a good introduction to statistical mechanics, which is arguably more suited to the “taste” of engineers than the book by Callen.

★ History of thermometry

♪ Chang H, *Inventing Temperature: Measurement and Scientific Progress*, Oxford University Press, New York, 2004.

Hasok Chang is not a physicist, but a professor of history and philosophy of science in Cambridge. Yet, his sound background in physics (not always granted among philosophers) is evident from this interesting book, in which, besides discussing in detail the development of the concept of temperature, Chang shows how many subtle problems arise when we try to give a self-consistent definition of a basic physical quantity.

★ Surface tension and wetting

♪ De Gennes P-G, Brochard-Wyart F, Quéré D, *Capillarity and Wetting Phenomena: Drops, Bubbles, Pearls, Waves*, Springer, New York, 2004.

The best introductory book to surface tension and wetting phenomena. A must for anyone interested in interfacial effects.

★ Brownian motion

♪ Pais A, *Subtle is the Lord: The Science and the Life of Albert Einstein*. Oxford Univ. Press., Oxford (1982)

This splendid biography of Albert Einstein by Abraham Pais, which is actually a book full of “hard” science, gives a detailed description of Einstein’s theory of Brownian motion, and of all the other fundamental results obtained by Einstein that we shall discuss in this book.

Statistical Physics

A Prelude and Fugue for Engineers

Piazza, R.

2017, XV, 453 p. 59 illus., Hardcover

ISBN: 978-3-319-44536-6