

Chapter 2

Thermodynamics for Living Systems

Abstract We start with an introductory chapter containing a concise and selected summary of the formalism of thermodynamics. Some basic concepts and the “principles” of macroscopic thermodynamics are firstly recalled, and put in an original perspective against the microscopic definitions of the same physical quantities, such as energy, temperature, pressure. The notions necessary for making the link between physics and biology are then introduced, and the prominent role of the entropy in describing the physics of living systems is highlighted. Should physics and biology be considered as two profoundly distinct camps, or rather two faces of one and a same reality? While such a hard question could hardly find an answer here, we show that even some of the most abstract conceptualisations of physics, such as the ideal gas, can be of great value in biology.

2.1 Macroscopic and Microscopic

Thermodynamics is not, strictly speaking, a physical theory on the same level as electromagnetism, gravity, quantum mechanics and so on. Besides the capability of giving a synthetic and formalised description of an ensemble of phenomena, the distinctive character of a theory is that of being able to formulate predictions. Moreover, a theory must include the possibility of confirming or refuting its own predictions. In this respect, thermodynamics is rather a powerful formalism, capable of describing diverse phenomena, allowing to unify under a rich conceptual structure experimental data from widely different domains, as well as proposing analogies between apparently disconnected phenomena.

Thermodynamics establishes relationships among physical quantities pertaining to a system, allowing to describe its behaviour under different conditions, however it needs to know nothing about the microscopic behaviour of the constituents of the system. As Steven Weinberg put it, *Thermodynamics is more like a mode of reasoning than a body of physical laws* [1]. The system of interest may be very often a machine, such as a refrigerator, a heater, an engine, a windmill. This is indeed how thermodynamics was invented, in the XIX century. However, a ‘system’ can be also any kind of experimental device, from the simplest one like a box filled of soap

bubbles, to the most complicated plasma chamber of a nuclear fusion reactor. And it could as well be a completely theoretical, idealised situation, such an ensemble of perfect mathematical points, each infinitely small but endowed with a finite mass, or an assembly of hard spheres riding on a frictionless plane, bumping into each other in straight trajectories.

Anyway, to give a thermodynamical description of our system of interest it is supposed that the system is in some way isolated from the external world, so that its behaviour can be studied in almost ideal conditions; or at least, it is arranged so as to be in contact with a reservoir (of energy, heat, matter) so big that such almost ideal conditions can be imposed from the exterior. An example of a well isolated system could be the interior of a glass balloon filled with gas; an example of a system in contact with a reservoir could be the interior of a refrigerator, kept at a constant temperature considerably lower than the surroundings.

In comparison to such ideally, or nearly-ideally controllable systems, any living organism looks like an extremely complex (and messy) system, inside which many transformations of all kind occur all the time, with a continuous exchange of energy, heat, materials, information with the surrounding environment. In this respect life would be regarded, from the point of view of thermodynamics, as not belonging to any of the above categories. Rather, almost going against the typical behaviour of thermal machines. We will do our best to cancel such a wrong impression from the mind of our reader.

Most people have (or think they have) an intuitive understanding of basic quantities such as temperature, pressure, volume. On the other hand, there are some other physical quantities in thermodynamics which require a bit more of reasoning to be understood, for example the energy, or the work. Eventually there are yet some other quantities, such as the entropy or the chemical potential, which seem rather obscure from the point of view of our daily experience, and require a deeper analysis in order to grasp their meaning and usefulness. In this Chapter we will follow a kind of hierarchical approach, introducing these various quantities starting from the most basic ones, and moving progressively up to the less intuitive ones. We will avoid as much as possible mathematical formalism if not strictly necessary, but will try to be as rigorous as possible whenever mathematics is needed.

A **macroscopic state** of a system is specified by global thermodynamic variables, or functions of state, which describe a property of the system as a whole.

The various functions of state are categorised into **extensive** and **intensive**. Extensive variables are a function of the system volume V : in other words, if we double V , the value of an extensive variable is also doubled; if we halve V , the value of an extensive variable is also divided by two. Moreover, the value of an extensive variable for a system is equal to the sum of that extensive variable for all of its subsystems: for example, the number of objects in a drawer ranged into compartments is the sum of the objects in all the compartments. Examples of extensive variables are:

- the volume itself, V
- the energy, E
- the entropy, S

- the number, N
- the deformation $\varepsilon = \Delta L/L$, of a length L

On the other hand, intensive variables are not a function of the system volume: they do not change upon a variation of V . Moreover, the value of an intensive variable for a system subdivided into subsystems is not the sum of the corresponding variables: the overall temperature of a flat divided in two rooms is just the common value of temperature, not the sum of the two temperatures in each room. Examples of intensive variables are:

- the temperature, T
- the pressure, P
- the chemical potential, μ
- the stress tensor, σ

Although not necessary for the development of thermodynamic relationships between the different macroscopic functions, the thermodynamic system can also be thought of being made up of some microscopic components (e.g., cells, atoms, molecules). In order to be given a proper mathematical treatment, the basic properties of such microscopic components are that they must be all identical and interchangeable.

Any particular arrangement (i.e., the ensemble of values specifying the details) of each and every microscopic components, corresponding to given values of the macroscopic thermodynamics variables, N, P, V, T, \dots , represents a **microscopic state** of the system, or microstate. A microstate of the system is specified by a set of variables describing the arrangement of its microscopic components: for example, the exact positions and velocities of each molecule in a balloon filled with gas.

Despite such an overly detailed microscopic accounting may seem an impossible task to carry out in a real experiment, the corresponding conceptual formulation is at the basis of all the developments of *statistical mechanics*, a powerful physical theory whose objective is to give an interpretation of macroscopic physical quantities in terms of their microscopic constituents. Since a macroscopic piece of matter contains a number of atoms or molecules of the order of the Avogadro's number, $N_{Av} = 6.02 \times 10^{23}$, and even one cubic centimetre of living tissue contains millions of cells, this connection cannot be made on a one-by-one basis, but must rather exploit some statistical properties of the ensemble. Hence the denomination of 'statistical' mechanics, as opposed to the deterministic character of Newtonian mechanics. On the other hand, the modern developments of computer simulations (see "Further reading" at the end of the chapter) allow to verify the predictions of statistical mechanics, by following the detailed dynamics of a simulated system of microscopic objects, and to prove the correspondence between microscopic and macroscopic quantities to an exceptional extent.

In classical thermodynamics there is no necessary relationship between the macroscopic state of a system and its microscopic states. In fact, thermodynamics *does not need* microscopic variables in order to work. It will be immediately apparent that, for any macroscopic state, there can be a large number of equivalent microscopic

states. Think of a large ensemble of N molecules of a gas enclosed in a balloon of volume V , for each of which you know exactly the positions and velocities at a given instant t . This is a microstate. Then, think of exchanging the velocities between two molecules, while leaving each of the two molecules at its place. This second configuration represents a different microstate, however the macroscopic condition of the system has hardly changed. Such a fundamental difference between the microscopic and macroscopic states will be the basis for introducing later on the physical variable entropy.

Notably, the fact that a seemingly very large number of microstates can correspond to one same macrostate, suggests that there should be some kind of statistics, such as a probability distribution, of these microstates. Think of the thermal agitation of the molecules in our gas, at any instant of time. For any given macroscopic condition of the gas (a set of values of temperature, pressure and volume) the many microstates composing it could indeed have different probabilities of occurring.

2.1.1 *Isolated System*

In thermodynamics, an isolated system is any ensemble of N objects, enclosed in a well defined volume of space V , which does not perform any exchanges of matter or energy with the space external to the volume V . It should be stressed that, up to this point, it is not at all necessary to specify the nature, nor the internal structure of the objects making up the system, in order to use the formalism of thermodynamics to describe it. The only basic requirements is that the objects are distinguishable, and that may be grouped if necessary into subsets according to some criteria. Our system can be composed by any number of independent subsystems (for example, the moving parts of an engine), provided the appropriate thermodynamical quantities can be defined for the different subsystems.

It is important to observe that, both in an experiment and in theory, we can never measure the absolute value of either extensive or intensive variables, but only differences of such variables between a current macroscopic state and a reference condition. This statement may appear apodictic, but it can be intuitively grasped by observing that to know the absolute system's energy we should be able to measure it in every possible microstate corresponding to a macrostate, and it is practically impossible to enumerate all the microscopic configurations available even for a system of a few objects enclosed in a rather small volume. However, by measuring relative differences between such quantities a relative scale can be established, and variations of quantities such as energy, temperature, pressure between different systems can be measured and compared.

The **Statistical Postulate of Equilibrium** states that: any isolated system spontaneously evolves towards a state, defined as the thermodynamic equilibrium, characterised by the fact that any differences of the functions of state between its subsystems go to zero. In practice this means that, once arrived at the equilibrium condition, the macroscopic functions of the system remain constant.

2.1.2 Energy

In very general terms, the energy of a system E describes its ability to operate transformations. For example, for a purely mechanical system the energy describes its capability to realise a mechanical work, W , such as lifting a mass against the gravity:

$$E = W = mgh \quad (2.1)$$

mg being the weight of the mass m in the Earth's gravitational field with constant acceleration $g = 9.807 \text{ m s}^{-2}$, and h is the relative height with respect to the point of start. As said above, we can only define differences of functions of state with respect to a reference value. If we take as zero the energy of the mass when it is lying on the floor, any position of the mass above the ground will correspond to a positive energy difference, and thus to the possibility of performing a non-zero work. The old pendulum clocks worked exactly with this principle. In the age when there was no electricity available in the household, at any recharge of the mechanism one or more weights were raised inside the clock; during the operation of the clock, the weights, attached to the wheels of the clock by a thin metal chain, would slowly descend toward the ground, setting in motion the spheres indicating the hours and minutes; once the weights reached the bottom, the clock would have exhausted all the available potential energy and would stop, until the next recharge. This was a fine and precise mechanism, capable of converting gravitational energy into mechanical energy of motion of the clock wheels.

If a system is composed of subsystems 1, 2, 3, ..., k , the total energy of the system is given by the sum of the energies of each subsystem:

$$E_{tot} = E_1 + E_2 + E_3 + \dots + E_k \quad (2.2)$$

Since the energy is additive, it must necessarily be an extensive thermodynamical variable. To specify its difference with respect to other energy-like quantities, to be introduced later on, this form of energy is usually called the **internal energy**.

2.1.3 Heat

Believed in ancient times to be some sort of fluid pervading the bodies, heat is just another form of energy. On a microscopic level, it is associated with the thermal agitation of the atoms and molecules: the faster the thermal agitation, the hotter will appear the body.

While remaining conserved, energy can however change of form: for example the mechanical energy of a system can be turned partly or totally into heat, Q . Every change of the form of energy bringing the system to a different macrostate with final energy E_{fin} , must be reported to the initial energy, E_{in} , which is the reference

state. Energy can be transformed into mechanical work, or any other kind of work (electrical, magnetic...), always indicated by W , while the system can transfer part of its energy to the surroundings as heat, Q , and end up in the final value E_{fin} , such that the initial energy E_{in} is conserved. Such an equivalence between different forms of energy is primarily grounded in countless experimental observations on the most various kind of systems, which are summarised by the classical equation:

$$\Delta E = E_{fin} - E_{in} = -\Delta W + \Delta Q \quad (2.3)$$

This is the so-called **First Principle** of thermodynamics, nothing but another way of expressing the universal conservation of energy, by stating that mechanical work and heat are equivalent forms of the energy.¹

It is customary to separate the energy turned into heat from the other forms of energy, in the balance of energy transformations, for the reason that, differently from other forms of energy, heat is a sort of “dead end” for the energy: once a part of energy is turned into heat, that energy is lost forever and cannot be recovered to perform any more work. This point will become more clear a bit later, once the concept of temperature is introduced.

The very important experimental observation here, for which not a single violation has ever been observed yet in the history of physics and chemistry, is that any transformation of energy is always accompanied by some loss in the form of heat.

A transformation of energy without any heat loss would be an ideal transformation, sometimes called **adiabatic**: this is defined as a transition between two states via a virtually infinite sequence of infinitesimally small and slow transformations, to make a system go from an initial to a final state without any heat loss.

2.2 Perfect Gas

A perfect gas is a very simple, idealised thermodynamical system, therefore very practical to study as a bookkeeping example of more complex situations. We consider a volume V , filled by a number N of microscopic particles, i.e. each with a negligible size compared to the volume V . Each one of these point-like particles has a mass m . We choose N and V such that the particle density:

$$\rho = N/V \quad (2.4)$$

¹It may be interesting to note that the first definite statement about this, eminently physical, principle came in fact from medicine. Around 1840, Julius Robert Mayer, then a physician in Java, deduced the energy equivalent of heat by observing differences in venous blood colour, which he attributed to different oxygen concentrations, and hence to different amounts of heat produced by the body. His empirical calculations led him to a value of 3.58 J/cal, not too far from the more accurate value measured by James Joule just a few years later, 4.16 J/cal, by means of calorimetry experiments [2, 3].

is so low that collisions between particles are practically highly unlikely. Each particle of our perfect gas is characterised by a position and a velocity \mathbf{v}_i in the volume V . Particles make perfectly elastic collisions with the walls of V , meaning that their energy is the same before and after the collision, only their direction changes because of momentum conservation. Additionally, we consider that there are no interactions whatsoever between the particles (chemical, electrical, magnetic, gravitational), therefore their position in space is not relevant to their energy. As a consequence, the only pertinent form of energy to each particle $i = 1, \dots, N$ is the kinetic energy:

$$E_i = \frac{1}{2} m v_i^2 \quad (2.5)$$

The total energy of the gas, E , is given by the sum of the kinetic energies of all the N particles:

$$E = \sum_{i=1}^N E_i = \sum_{i=1}^N \frac{p_i^2}{2m} \quad (2.6)$$

where we have introduced the particle momentum vector $\mathbf{p} = (p_x, p_y, p_z) = m\mathbf{v}$.

Every microstate of the perfect gas corresponds to a set of values of positions and momenta $\{\mathbf{r}_i, \mathbf{p}_i\}_{i \in N}$, different at every instant, provided that E remains constant since the system is isolated. In principle, the number of microstates for a given combination of $\{NVE\}$ is infinite: we may change the momentum of a particle in infinitesimal increments (taking care of changing two \mathbf{p} vectors at a time by equal and opposite amounts, so that the energy remains constant), and obtain infinitely many different microstates of the system. Instead of the absolute number of microstates for an energy E , which is clearly impossible to count, we can look at the function expressing the *density of microstates* in an energy interval dE around E . This latter can be rather easily estimated by considering the very special situation of the perfect gas. (This function can be calculated just for a few other simple systems, but it becomes a very complicate task for any realistic system, with arbitrary interactions among the particles.)

2.2.1 Counting Microstates

The peculiarity of the perfect gas is that particles do not have any mutual interaction. Therefore, their energy is given by a simple sum of squares of the particle momenta, $E = \frac{1}{2m} \sum_i p_i^2$. We can also detail this expression according to the Cartesian components of each particle's momentum vector:

$$E = \frac{1}{2m} \sum_{i=1}^N p_{ix}^2 + p_{iy}^2 + p_{iz}^2 \quad (2.7)$$

the sum containing $3N$ terms.

This formula for the total energy looks like a sort of Pythagoras' theorem for an ideal "triangle" with $3N$ sides. In fact, one could read the above formula as: $\sqrt{2mE}$ is the distance from the origin of the point \mathbf{p} , with coordinates $(\mathbf{p}_1, \dots, \mathbf{p}_N)$, in a space with $3N$ dimensions.

This is a very interesting interpretation of the energy of a perfect gas. If we imagine \mathbf{p} as a vector with $3N$ components centred at the origin, and we imagine to change arbitrarily any of its components in every possible way that give the same final value of E , the vector \mathbf{p} will describe a $3N$ -dimensional sphere centred at the origin with radius $R = \sqrt{2mE}$.

Then, it may be thought that the number of available microstates (i.e., different combinations of the \mathbf{p}_i) for a given energy, let us call it $\Omega(E)$, should be proportional to the surface of this sphere: a larger value of the energy E corresponds to a larger sphere surface, and therefore to a larger Ω . A system with assigned values of $\{NVE\}$ (the macrostate) must conserve its energy, so it will ideally "move" on this constant-energy surface, exploring all the Ω microscopic configurations whose E is compatible with the macrostate.

What is the surface of a sphere in $3N$ dimensions? We can reason by analogy:

- $2\pi R$ is the perimeter of a circle (sphere in 2 dimensions)
- $4\pi R^2$ is the surface of a sphere (3 dimensions)

The surface of dimension n appears to be proportional to the volume divided by R , $S(n) \approx V(n)/R$, therefore we may write:

$$S(3N) = cR^{3N-1} \quad (2.8)$$

for the surface of a hypersphere in $3N$ dimensions. With a little algebra, it can be shown that the exact proportionality coefficient is $c = (2\pi)^{3N/2} / \Gamma(3N/2)$ (see the Appendix A for the properties of the special Γ function). For very large values of N , one can approximate $N - 1 \simeq N$. Therefore, replace $3N - 1$ by $3N$, and $\Omega \approx R^{3N}$. Since $R = \sqrt{2mE}$, it is finally $\Omega(N, E) = (2\pi)^{3N/2} (2mE)^{3N/2} / \Gamma(3N/2)$.

Now, we note that each microstate is specified by assigning also the values of positions \mathbf{r}_i for each particle, besides their momenta or velocities: two microstates could have the same distribution of velocities, but differ in the positions of some particles. Since each particle of the perfect gas can be found anywhere in the volume V , independently of the others, the number Ω should be multiplied by a factor V contributed from each particle. The complete expression of Ω is therefore:

$$\Omega(N, V, E) = V^N \frac{(2\pi)^{3N/2} (2mE)^{3N/2}}{\Gamma(3N/2)} \quad (2.9)$$

2.3 Entropy and Disorder

Disorder in real life is conceived as opposed to order. A sequence, a pattern, a book-shelf can be quickly seen to be ordered or disordered. Both order and disorder seem easy to recognise, when looking at an array of objects, or at your children's room. However, the view of the contrast between order and disorder as just an attribute of the spatial organisation is somewhat limited. A spatially well-ordered crystalline lattice of magnetic atoms can turn out to be completely disordered, if we look at the values of magnetisation in that same crystal above some critical temperature. In physics, a macroscopic state of disorder of a system consists in the inability to formulate predictions about the behaviour of the system or, in other words, a very limited knowledge about its actual state. For the microscopic states of this same system, the condition of disorder would correspond to the microscopic variables (positions, velocities of the particles) assuming a wide spectrum of completely random values.

Strange as it may seem, this very concept of disorder will allow us to establish a conceptual link between the macroscopic system and its microscopic constituents. Moreover, it should not be thought that the condition of equilibrium is synonymous to some kind of order. Indeed, the macroscopic state of equilibrium of a system is not represented by one particular microstate, for example one having a very peculiar distribution of velocities, or one with a very regular arrangement of the positions of the particles. Microscopic equilibrium is associated to an equal probability distribution of all the admissible micro states (i.e., those corresponding to a same macrostate), *provided the disorder of the system is maximised*. Where such a quite surprising observation comes from? How come that equilibrium should correspond to a maximum disorder, in parallel to the disappearing of any differences between all the state variables of the system?

We are now ready to propose a definition of the order and disorder condition for a system, by constructing a mathematical function explicitly dedicated to this purpose. The characters of such a function should be at least: (1) that it be linked to the constant energy of the system, and (2) that it be proportional to the number Ω of microscopic realisations of the same value of macroscopic energy. The latter condition implies that such a quantity must be extensive.

Since V and E appear as multiplicative factors in the expression for Ω , Eq. (2.9), such a function must be necessarily be constructed from the logarithm of Ω , so that if we break the system into subsystems with energies E_1, \dots, E_k , the value of such a quantity can be obtained as the sum of the corresponding quantities for each subsystem k . For historical reasons, this new physical quantity is called the **entropy** of the system, and is indicated by a letter S :

$$S = k_B \ln \frac{\Omega}{N!} \quad (2.10)$$

The factor k_B is the Boltzmann constant, equal to 1.38×10^{-23} J/K, or 8.5×10^{-5} eV/K, allowing to express the entropy in energy-like units (we note that entropy may

be also used in different contexts, such as information theory, in which no energy-like units are necessary). The $1/N!$ factor was introduced by J. W. Gibbs, one of the ‘founding fathers’ of statistical physics, to discount the fact that the particles in the perfect gas are all identical. This implies that microstates with identical positions and velocities of all particles, but in which the “label” of any two particles is exchanged, are identical, therefore the number Ω must be divided by the number $N!$ of equivalent permutations of the N identical particles. On the other hand, this multiplicative factor has no physical basis for classical particles, and was introduced by Gibbs merely on empirical grounds (because without it the entropy would not be strictly additive [4, 5]). The presence of the $1/N!$ can be fully justified only in the framework of quantum mechanics, which admits the non-distinguishability to be a fundamental property of quantum particles.

Therefore, from the above semi-analytical derivation, the entropy of the perfect gas is obtained as:

$$\begin{aligned}
 S(N, V, E) &= k_B \ln \left(\frac{(2\pi)^{3N/2} V^N (2mE)^{3N/2}}{\Gamma(3N/2) N!} \right) = \\
 &= k_B [N \ln V + N \ln(2mE)^{3/2} - \ln \Gamma(3N/2) - \ln N! + c] \simeq \quad (2.11) \\
 &\simeq N k_B \left\{ \ln \left[\left(\frac{V}{N} \right) \left(\frac{2mE}{N} \right)^{3/2} \right] + c \right\}
 \end{aligned}$$

This is the Sackur-Tetrode equation for the *absolute entropy* of the perfect gas,² independently derived in 1912 by the physicists Otto Sackur in Germany, and Hugo Tetrode in the Netherlands [6, 7]. Its utility lies in the fact that it is one of the rare examples for which we can calculate explicitly the absolute value of the entropy, which is otherwise a rather elusive quantity to grasp in real life.

It is an experimental fact that the entropy of an isolated system can only increase with any spontaneous transformation. Thermodynamics was born in the early XIX century, with the purpose of knowing better how thermal machines worked [8, 9]. In the old experiments designed to understand the relationships between energy and heat, the French physicist Sadi Carnot defined around the year 1824 the concept of a cyclic thermal engine, for which he established that it was necessary to have at least two parts of the engine at different temperatures in order to extract useful work. About twenty years later, the German physicist Rudolf Clausius realised that any transformation in which only heat was exchanged, between a “hot” and a “warm” body, would be *irreversible*. Please notice the “...” in the previous sentence. Indeed, this is a bit of a circular definition, since we do not know exactly what is a hot or a warm

²By using arguments from quantum mechanics, it is shown that the constant appearing in the equation must be $c = \frac{5}{2} + \frac{3}{2} \ln \frac{2\pi}{3h^2}$, with h the Planck’s constant.

body, unless we measure a difference between two bodies in contact.³ The definition of a quantity to be used as temperature, in fact, is intimately linked to the definition of the entropy (although the practical notion of temperature, and tools to measure it, were already known since at least two centuries earlier). What the experiments actually measured was the fact that the flow of heat is always unidirectional, and this direction of the flow allows to establish which body is “hotter” than the other, thus making for a temperature scale.

Sometime during the first half of the XIX century, all the experimental observations about the loss of energy into heat, expansion, and irreversible phase changes, were summarised in the so-called **Second Principle** of thermodynamics:

$$\Delta S \geq 0 \quad (2.12)$$

which states that in any spontaneous transformation the entropy of the system must increase. The function entropy was introduced just to give this principle a formal mathematical statement. Besides, we note that the second principle was in fact the first to be established, in 1824 by Carnot. The ‘disorder’ interpretation of entropy were to be laid out only about 60 years later, by Ludwig Boltzmann. According to our construction of the entropy function of state, in terms of the size Ω of the available space of microstates, an increase in S corresponds to an ever larger number of microstates, all becoming equally probable at equilibrium. Therefore, as far as the entropy increases, the information about the microscopic state of the system spreads out onto a larger and larger ensemble. We can interpret this as a net loss of information, or an increase of the disorder of the system.

2.3.1 Irreversibility and Probability

Let us consider a perfect gas initially confined in a volume V . If we let the gas expand so as to double its volume in such a way that the gas does not perform any work, the only term that changes in the Eq. (2.11) above is V^N . The corresponding entropy change is:

$$\Delta S = S_{fin} - S_{in} = k_B [\ln(V)^N - \ln(V/2)^N] = Nk_B \ln 2 \quad (2.13)$$

always positive. One may ask why we implicitly assumed that the approach to equilibrium implies the *spontaneous* expansion of the gas from $V/2$ to V . The idea of a spontaneous transformation is linked to the concept of **irreversibility**. In a classical definition, irreversible means that the system could never go back spontaneously to its initial state. However, a more subtle interpretation can be posed in terms of the

³Already in the middle of XVIII century the English natural philosopher John Locke, in his *Essay on the human intellect*, had defined the temperature as a relative concept, stating that a body could be hot or warm only in relation to another body.

probability of occurrence of the accessible system states. This is not just an academic disquisition, since the macroscopic irreversibility contrasts with the perfect reversibility (called *time invariance* in physics) of the microscopic events.

It may be considered that the N microscopic particles, in their continuous thermal agitation, should explore any possible microscopic configuration. As a result, a microstate with $N/2$ particles in the half-volume $V/2$ and $N/2$ particles in the other half-volume $V/2$, should have the same probability as another microstate in which all the N particles are in the half-volume $V/2$, and zero in the other half. That may be true, provided the entropy of the system is increased in the transformation, according to the Second Principle. It is easy to see that the entropy of the gas in a doubled volume increases. On the other hand, the probability that starting from a volume V , one mole of gas ($N \sim 10^{23}$ particles) goes back spontaneously by random fluctuations inside a volume $V/2$ is:

$$prob = \left[\frac{(V/2)}{V} \right]^{10^{23}} \approx 10^{-8,240,000,000,000,000,000,000} \quad (2.14)$$

Evidently, for a system with a very small number of particles, at the limit of one single atom or molecule, the theoretical reversibility (originating from the fact that both the classical and quantum mechanics equations are unchanged if the time t is exchanged with $-t$) should be possible also in practice. In the recent years, it has indeed become feasible to perform experiments in which one single molecule is tracked in time and, after accurate measurements, local violations of the Second Principle have been found, when the number of degrees of freedom is reduced to a minimum [10].

Such an observation is common to any systems with a reduced number of degrees of freedom. If we flip ten coins, the probability of obtaining the same face on all of them is small but non negligible: this probability is simply $(1/2)^{10} = 0.001$. However, if we flip an Avogadro's number of coins, the probability of having the same face on all coins is $(1/2)^{10^{23}}$: in this sense, there are no strictly irreversible transformations, only very much improbable ones!

We will come back on this relationship between entropy, reversibility and disorder in Chap. 3, when discussing the probability of assembling complex proteins and nucleic acids starting from the simpler molecular building blocks.

2.4 Closed Systems

A system which can exchange energy but not matter with its environment, is called a **closed system**. The ideal example of a closed system is given by considering a box divided into two parts by a wall taken to be infinitely rigid, except for a small ideally elastic portion of negligible size compared to the whole wall surface. When two particles coming from the opposite parts of the box collide against this flexible wall, which is said to be ideally elastic, they can exchange their respective kinetic

energies as if undergoing a perfectly elastic collision, however without the possibility of jumping in the opposing part of the box. As required by the above definition, the two parts of the box can exchange some energy but not matter.

Such a highly idealised system will now allow us to introduce the concept of **temperature**. Let us imagine that in the two parts of the box, which we call A and B , there are N_A and N_B particles, respectively, with energies E_A and E_B . The energy of the combined system is:

$$E_{tot} = E_A + E_B \quad (2.15)$$

The total energy must be constant. However, the subdivision into E_A and E_B can take any combinations resulting in the same value of E_{tot} .

The total entropy of the combined system is:

$$S_{tot} = S_A(E_A) + S_B(E_B) \quad (2.16)$$

But once E_{tot} is fixed, the two values E_A and E_B are no longer independent. Therefore:

$$S_{tot} = S_A(E_A) + S_B(E_{tot} - E_A) = S_{tot}(E_A) \quad (2.17)$$

a writing which underscores the dependence of S on a single energy variable. The entropy equation gives for the two gases:

$$S_{tot}(E_A) = k_B \left[N_A \left(\frac{3}{2} \ln E_A + \ln V_A \right) + N_B \left(\frac{3}{2} \ln(E_{tot} - E_A) + \ln V_B \right) \right] + const \quad (2.18)$$

Now we ask: which is the most probable value for E_A ? The statistical postulate seems to suggest that any value should be equally probable at equilibrium. On the other hand, since the number of microstates increases with energy, we should have the maximum possible energy on each side, in order to simultaneously maximise the entropy. To find the maximum of entropy from the above equation, it is sufficient to take its derivative with respect to the unique variable E_A and set it equal to zero:

$$\frac{dS_{tot}}{dE_A} = \frac{3}{2} k_B \left[\frac{N_A}{E_A} - \frac{N_B}{(E_{tot} - E_A)} \right] = 0 \quad (2.19)$$

This clearly shows that the above conditions can be satisfied only if the energy is shared between A and B in a way proportional to the number of particles in each subsystem (actually, proportionally to the number of degrees of freedom of the particles in each subsystem, which is $3 \times$ the number of particles, one for each cartesian coordinate). In particular, if the number of particles in the two subsystems is equal, the energy is also equally shared, $E_A = E_B$. If this is not the case, the maximum of the entropy corresponds to:

$$\frac{E_A}{N_A} = \frac{E_B}{N_B} \quad (2.20)$$

This result was experimentally known well before the introduction of the microscopic ideas about the constitution of matter. It is the **Equipartition law**, which was based on the experimental observations performed by the French chemists A.-T. Dulong and P.-L. Petit (1819) on the constancy of the specific heat of solids at high temperature.

2.4.1 Temperature

Indeed, it is just on the basis of this result that we can introduce the concept of temperature, a quantity for which most of us have an intuitive perception, but which is just a bit more complicate to put on formal grounds.

A formal definition, eventually valid for any system and not just for a perfect gas, is the following. Starting from the entropy equation, we want to find the variation of entropy as a function of an infinitesimal change of energy from E to $E + dE$, all the other thermodynamic functions remaining unchanged. We have:

$$\begin{aligned} dS(E) &= Nk_B d\left(\frac{3}{2} \ln E + \ln V\right) = \frac{3}{2} Nk_B (\ln(E + dE) - \ln E) = \\ &= \frac{3}{2} Nk_B \ln(1 + dE/E) \simeq \frac{3}{2} Nk_B \frac{dE}{E} \end{aligned} \quad (2.21)$$

The temperature T is identified as the numerical coefficient in the right-hand side of the equation:

$$\frac{dS}{dE} = \frac{1}{T} \quad (2.22)$$

This is the fundamental definition of the temperature, which we postponed up to this point (although the practical notion of temperature would come first, historically) in order to firstly have well clear in mind the concepts of energy, heat and entropy. From this definition, we see that a small variation of the energy makes a large variation of entropy when the system is at low temperature; on the other hand, as the temperature is increased, the increase in entropy as a function of energy gets smaller and smaller. (For the perfect gas, this should be already evident from the relationship $S \propto \ln E$, hence $T \propto E$.)

By the last definition, the law of equipartition reads:

$$\frac{E}{N} = \frac{3}{2} k_B T \quad (2.23)$$

This equation says that the average energy per particle is equal to $\frac{3}{2} k_B T$. By considering that each point particle is defined by three degrees of freedom (v_x , v_y , v_z), it also says that each degree of freedom in the system contributes $\frac{1}{2} k_B T$, which was exactly the experimental deduction of Dulong and Petit in their study of solids.

For a perfect gas the temperature coincides with its kinetic energy, apart from a numerical coefficient with appropriate dimensions:

$$\frac{E}{N} = \frac{3}{2}k_B T = \frac{1}{N} \sum_{i=1}^N \frac{1}{2} m v_i^2 \quad (2.24)$$

This is an important consideration, in that the above equation puts into correspondence the macroscopic vision of thermodynamics, represented by the temperature of the perfect gas, with the statistical properties of its microscopic constituents, that is the kinetic energy of an ensemble of particles.

Let us consider two systems A and B at two different temperatures, $T_A \neq T_B$, both at their respective equilibrium. By the moment we put them in contact, the particles from the two systems will start colliding, and redistributing their energy until no further exchanges are capable of changing the distribution of velocities. The new equilibrium state is therefore characterised by the condition $T_A = T_B$. This is sometimes called the **Zeroth Principle** of thermodynamics. In simple words: *temperature is the quantity that becomes equal when two macroscopic closed systems, freely exchanging energy, attain a common equilibrium.*

2.4.2 Caloric Definition of the Entropy

Up to now we have been constantly looking for microscopic definitions of known macroscopic thermodynamic quantities. However, as we noticed above, thermodynamics does not need such microscopic definitions in order to work. Therefore, it may be surprising that at this point we have only a microscopic definition of the entropy. Entropy has been defined in terms of the number Ω of microscopic realisations of a given macroscopic state. Is there a macroscopic function which corresponds to such a microscopic definition?

In fact, the most ancient definitions of entropy given by Clausius (1854) and others (Maxwell 1867, Gibbs 1902, etc.), which largely preceded the microscopic formulation above, were suggested by experiments of transformation of energy into work. As already discussed, it was (and still is) experimentally observed that in every transformation some part of the energy was irreversibly transformed into heat, dispersed in the environment. For an energy transformation at constant temperature, Clausius defined the quantity:

$$\Delta S = \frac{\Delta Q}{T} \quad (2.25)$$

as the fraction of initial energy lost into heat, and introduced the term entropy for the quantity labelled S (with dimensions of [Energy]/[Temperature]). Therefore, by going back to Eq. (2.3), we can write:

$$\Delta E = E_{fin} - E_{in} = -\Delta W + \Delta Q = -\Delta W + T \Delta S \quad (2.26)$$

Mechanical equilibrium of a perfect gas under pressure

A wall of the box of volume V is free to move against a fixed spring, with constant k , which holds it in place at equilibrium (see Fig. 2.1a). The lateral section of the box is S , and the equilibrium length of the box is L , in the direction parallel to the force $F = -kL$ applied by the spring. Therefore, $V = SL$.

The total energy of the perfect gas is the sum of the kinetic energy of the molecules, plus the elastic potential energy $U = kL^2/2 = FL$ stored in the elastic spring:

$$E_{tot} = E_{kin}^{GP} + FL \quad (2.27)$$

If the wall is moved from L to L' , with a change of volume from $V = SL$ to $V' = SL'$, the entropy of the gas changes by a quantity:

$$\begin{aligned} \Delta S &= Nk_B \Delta(\ln(E_{kin})^{3/2} + \ln V) = Nk_B \left(\frac{3}{2} \frac{\Delta E_{kin}}{E_{kin}} + \frac{\Delta V}{V} \right) = \\ &= \frac{1}{T} \left(\Delta E_{kin} + \frac{Nk_B T}{L} \right) (L - L') \end{aligned} \quad (2.28)$$

since $\Delta(\ln x) \simeq \Delta x/x$, and $E_{kin}/N = 3/2(k_B T)$. Energy conservation implies:

$$E_{kin}^{GP} + FL = E_{kin}'^{GP} + FL' \quad (2.29)$$

or:

$$\Delta E_{kin} = E_{kin}' - E_{kin} = -F(L' - L) \quad (2.30)$$

Therefore:

$$\Delta S = \frac{1}{T} \left(-F + \frac{Nk_B T}{L} \right) \Delta L \quad (2.31)$$

At equilibrium the variation of entropy stops, $\Delta S = 0$. The new equilibrium position of the wall is:

$$L' = \frac{Nk_B T}{F} \quad (2.32)$$

The force of the spring divided by the surface S is the pressure, $P = F/S$ (a force per unit surface). If we multiply Eq. (2.26) by S , with $V = SL$, we find:

$$V = \frac{S}{F} Nk_B T \quad (2.33)$$

or:

$$PV = Nk_B T \quad (2.34)$$

This is just the *equation of state* of the perfect gas. It should be noted that, in a isolated system, every volume variation implies a variation of the temperature (since the latter represents in a perfect gas a variation of the kinetic energy, $\Delta E_{kin} = -P \Delta V$).

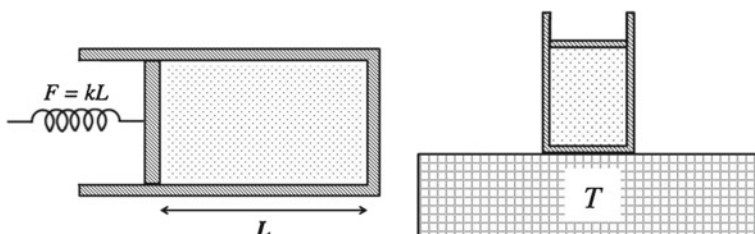


Fig. 2.1 *Left* Schematic of a gas-containing box of cross section S and variable length L . The total volume is $V = SL$. The walls are assumed to be infinitely rigid so that the gas molecules perform only perfectly elastic collisions. The *left-side* wall can move without friction, to adjust against the force $F = kx$ imposed by the mobile spring. The corresponding internal gas pressure is $P = F/S$. *Right* The gas box in contact with a thermostat at temperature T . By definition, the volume of the thermostat is so much bigger than the volume V of the gas, to be unperturbed by whatever happens inside the volume V

This ‘caloric’ definition of the entropy was the one chemists and physicists had been familiar with for the whole XIX-th century, and still is the most useful definition for engineers and whoever works with macroscopic thermal machines. It should be absolutely remarkable, then, that we can reobtain all the well known properties of a perfect gas, such as the equipartition, or the equation of state (see next greybox on p. 24), starting from Boltzmann’s statistical formulation of the entropy. Boltzmann had already used a similar equation in his early work on the mechanical interpretation of the Second Principle, yet without thinking of it as a probability. The identification of the phase space density Ω with a microscopic probability came a few years later, in his work of 1871 on the kinetic theory of gases [11]. Either way, the theoretical definition of entropy written as Eq. (2.10) was a purely conceptual endeavour, compared to the fully experimentally-grounded caloric definition above. In his time, Boltzmann had long disputes with his colleagues scientists, and even with journal editors, who refused to take his assumptions about the microscopic behaviour of atoms as nothing more than practical speculations. His pioneering views about the microscopic connection with the macroscopic world would be confirmed only many years later, by J. Perrin’s experiments on colloidal suspension that allowed to measure with high precision the values of both the Avogadro’s number and of Boltzmann’s own k_B constant [12], and were finally vindicated by the revolution of quantum physics.

2.5 Free Energy

Now we move on, to considering our perfect gas enclosed in a volume V and in contact with a surface at constant temperature T . It may be imagined that the volume V is placed over a very large block of material, as in Fig. 2.1b, so large to not be affected by the transformations eventually occurring inside the much smaller volume V . The perfect gas can exchange kinetic energy across the contact surface

with the block, however according to our definition the temperature of the latter will remain unchanged. The sum of the subsystem represented by the volume V plus the whole block is again an isolated system. The block is called a thermal reservoir, or **thermostat**.

2.5.1 Exchanges of Energy at Constant Volume

The greybox on p. 24 shows that the temperature of the perfect gas is linked to the changes of volume for an isolated system. Since $\Delta E = T \Delta S - P \Delta V$, we should have $T \Delta S = P \Delta V$ for a transformation occurring at constant internal energy, $\Delta E = 0$. This is a well known result of thermodynamics: the spontaneous expansion of a gas ($\Delta V > 0$) implies an increase in the entropy ($\Delta S > 0$), which also implies that it should be impossible to observe a spontaneous contraction of the volume ($\Delta V < 0$), since this would violate the Second Principle. A compression of the gas must be accompanied by a parallel variation in the energy, $\Delta E < 0$, the negative sign of the energy meaning that this is supplied to the system from an external source, in an amount such that $\Delta S = (P \Delta V - \Delta E)/T > 0$.

Let us now imagine that the volume V is a closed (sub)system, exchanging energy with the block at constant temperature T (a *thermostat*, Fig. 2.1b). In this way, we can imagine a transformation, for example an expansion or a compression of the volume, during which the temperature of the subsystem V also remains constant, the difference in energy being compensated by the (positive or negative) exchange with the block. It would seem that in this way the Second Principle could be violated: a volume compression could lead to a diminution of the entropy. In fact, we must not fool ourselves and look at the overall entropy variation. During the transformation, the block transferred a quantity of energy $-\Delta E$ to the volume V , to keep its temperature equal to T . Therefore, the entropy of the block changed by $\Delta S_{block} = -\Delta E/T$, so that the total entropy variation is:

$$T \Delta S_{tot} = T \Delta S_V + T \Delta S_{block} = T \Delta S_V - \Delta E \quad (2.35)$$

For a closed system, this is the quantity that has to be maximised during a transformation. In the place of the internal energy, we can introduce a more complete function of state, F , which is called (Helmoltz) **free energy**:

$$F = E - TS \quad (2.36)$$

which is minimised during the approach to equilibrium at constant V and T .

The concept of a “free” energy is connected with the idea that in thermodynamics we always think of the capability of our system to develop some work (mechanical, or other). The difference between E and TS means that when equilibrium is reached at constant temperature, E is not yet at its minimum, and therefore we could think of using this residual energy to do some more work. Instead, in such conditions, the quota of work actually available is the difference $E - TS$.

2.5.2 Exchanges of Energy at Constant Pressure

Still in the same conditions of contact with the surface of the large block, we could think that our perfect gas could use the energy supplied to perform an expansion at constant pressure P . In this case we would let the volume V change so that the pressure can remain at a constant value. The energy supplied by the block in this case would be $\Delta E + \Delta W = \Delta E + P\Delta V$, namely a part necessary to keep $T = \text{const}$ as before, plus a part to keep $P = \text{const}$. The corresponding change in entropy is:

$$T\Delta S_{\text{tot}} = T\Delta S - \Delta E - P\Delta V \quad (2.37)$$

This leads to the introduction of yet another function of state, which is also called (Gibbs) free energy:

$$G = E + PV - TS \quad (2.38)$$

to be minimised in the approach to equilibrium at constant P and T (actually, this is just another way of saying that entropy has always to be maximised, no matter what you do to your system).

The quantity $H = E + PV$ is called the **enthalpy** of the system. It is a quantity often measured in experiments, since it takes into account all of the energy traded in the transformation of a sample (for example, the melting of an ice cube into water, accompanied by a reduction of volume at constant pressure).

At equilibrium, when $\Delta G = 0$, the system pressure can be formally defined as:

$$P = T \left(\frac{dS}{dV} \right)_E \quad (2.39)$$

(the subscript E indicating to calculate the derivative at constant total energy) in good analogy with the formal definition of temperature above.

2.6 Open Systems

A thermodynamic system capable of exchanging both energy and matter with its surroundings is called an **open system**. This is evidently the most general situation, and it is the typical context of living systems, as opposed to either isolated or closed systems, which are typical representations of a sealed laboratory experiment or a well isolated thermal machine, such as an oven or a refrigerator. An open system is characterised by the possibility of varying the number of constituents N (atoms, molecules, cells...), or their concentration $c = N/V$, according to a given value of **chemical potential** μ .

The Gibbs-Duhem equation

The derivation of the equivalence between free energy and chemical potential offers a nice example about how the thermodynamical relationships are derived. Sometimes this may seem a blind mathematical procedure which must just be followed strictly. Even if the mathematical steps may have not much of a meaning in itself, the final result may often show unsuspected correlations between different physical variables, or experimental measurements.

Let us start with the equation expressing the First Principle for an open system in the most general way, i.e. by including also the possible variation of the number of particles. For even more generality, let us also include the possibility that there are k different families of microscopic particles in the system, each with its own chemical potential μ_k and concentration $c_k = N_k/N$:

$$E(S, V, N) = TS - PV + \sum_{i=1}^k \mu_i N_i \quad (2.40)$$

The total differential of the energy with respect to all of its independent variables is formally written as:

$$dE(S, V, N) = TdS - PdV + \sum_{i=1}^k \mu_i dN_i \quad (2.41)$$

The total differential of each term composed by two variables is, in fact, $d(TS) = TdS + SdT$, and $d(PV) = PdV + VdP$, so that the previous equation becomes:

$$dE(S, V, N) = d(TS) - SdT - d(PV) + VdP + \sum_{i=1}^k \mu_i dN_i \quad (2.42)$$

Now, let us regroup all the total differentials at the left-hand side:

$$d(E - TS + PV) = -SdT + VdP + \sum_{i=1}^k \mu_i dN_i \quad (2.43)$$

However, by looking at the definition of Gibbs free energy, Eq. (2.38), it is also:

$$dG = -SdT + VdP + \sum_{i=1}^k \mu_i dN_i \quad (2.44)$$

This is the famous *Gibbs-Duhem equation*, which is at the basis of all the chemical thermodynamics. Of particular interest for our purposes is the case of constant temperature and pressure, which is typically realised in living systems: for that condition, the dT and dP terms are equal to zero, and the following relationship holds:

$$(dG)_{T, P} = \sum_{i=1}^k \mu_i dN_i \quad (2.45)$$

The above equation is also very useful to study equilibrium concentrations of multi-component systems. For example, the relative concentration of a binary mixture of A and B at equilibrium ($dG = 0$) is: $N_A/N_B = -(\mu_B/\mu_A)$.

The chemical potential can be thought of representing a sort of external field, like a gravitational potential, but concerning concentration adjustments. In the case of gravity, a mass would be attracted towards the center of gravity of another body. By analogy, chemical entities like atoms or molecules are attracted from a region of high concentration to a region of low concentration, until a condition of equal concentration is reached. At that point, the chemical potential has the same value everywhere, and the attraction stops. In microscopic terms, the role of chemical potential in defining the particle flow direction will become very clear when describing the process of diffusion across a membrane (Chap. 5).

From the point of view of the formal theory of thermodynamics, it is worth noting that, for constant pressure and temperature, the chemical potential at equilibrium equals the partial molar Gibbs free energy (see the greybox on p. 28).⁴

2.6.1 Entropy of a Mixture

Let us consider a system with a mixture of different elements, for example particles of two different types, A and B . The volume V is divided in two parts, V_A and V_B , such that $V_A + V_B = V$. The two groups of particles, N_A and N_B , are initially confined in the two subsystems V_A and V_B . The total number of particles is $N = N_A + N_B$, correspondingly we can define the concentrations $c_A = N_A/N$ and $c_B = N_B/N$, so that $c_A + c_B = 1$. We imagine that the two volumes can be connected, for example by a mobile wall. Removing the wall corresponds to an expansion at constant temperature, each of the two gases of particles now occupying simultaneously the entire volume V . Once the equilibrium is reached, the change in free energy of the mixture is:

$$\Delta G = \Delta E + P \Delta V - T \Delta S = 0 \quad (2.46)$$

On the other hand, the two non interacting gases are described only by their kinetic energy. At constant T the kinetic energy does not change ($\Delta E = 0$), therefore:

$$\Delta S = \left(\frac{P}{T} \right) \Delta V = N k_B \frac{\Delta V}{V} = N k_B \ln \left(\frac{V + \Delta V}{V} \right) \quad (2.47)$$

Now we write the entropy for each component of the gas separately:

$$\Delta S_A = N_A k_B \ln \left(\frac{V_A + V_B}{V_A} \right)$$

$$\Delta S_B = N_B k_B \ln \left(\frac{V_A + V_B}{V_B} \right)$$

⁴Nearly all the concepts and mathematical functions for the study of systems at constant pressure or temperature, including the enthalpy, free energy, chemical potential, were developed and formalised in the monumental paper by J.W. Gibbs [13], which by common consensus marks the beginning of the modern vision of thermodynamics and its connections with chemical and electrical phenomena.

and sum the two to obtain the total entropy:

$$\begin{aligned}
 \Delta S_{tot} &= \Delta S_A + \Delta S_B = N_A k_B \ln \left(\frac{V_A + V_B}{V_A} \right) + N_B k_B \ln \left(\frac{V_A + V_B}{V_B} \right) = \\
 &= -k_B \left[N_A \ln \left(\frac{V_A}{V_A + V_B} \right) + N_B \ln \left(\frac{V_B}{V_A + V_B} \right) \right] = \\
 &= -k_B \left[N_A \ln \left(\frac{N_A}{N_A + N_B} \right) + N_B \ln \left(\frac{N_B}{N_A + N_B} \right) \right]
 \end{aligned} \tag{2.48}$$

since for either component of the gas it is $V = N(k_B T/P)$. Remembering the definition of concentration c :

$$\begin{aligned}
 \Delta S_{tot} &= -k_B (N_A \ln c_A + N_B \ln c_B) = -N k_B (c_A \ln c_A + c_B \ln c_B) = \\
 &= -N k_B [c_A \ln c_A + (1 - c_A) \ln(1 - c_A)]
 \end{aligned} \tag{2.49}$$

In the limit of small concentrations, either c_A or $c_B \ll 1$, the limit of the previous equation for the minority component is:

$$\Delta S \simeq -N k_B c \ln c \tag{2.50}$$

This equation will be very useful to study the osmotic behaviour of ions and molecules in small concentrations inside the cellular fluid.

2.7 The Biosphere as a Thermal Engine

The **biosphere** is defined as the ensemble of the Earth's matter that is actively employed in the fabrication of living organisms. A part of this matter is found directly in the organisms (animals, plants, bacteria...) and the remaining part is found in the recycled material. Liquid and solid water from oceans, as well as the air of the atmosphere, are part of the biosphere, in that they are necessary components of living systems. The environment comprises all the remaining matter of the Earth (the solid lithosphere), and the Universe, including the radiation. In thermodynamic terms, the biosphere is an open system, receiving a constant flux of energy and entropy from the outside environment (it also receives a small flux of matter, from the cosmic radiation). If we exclude the incoming flux, the biosphere can be considered a closed system. We will now delve a bit deeper in the thermodynamics of the energy and entropy flowing from the environment into the Earth's biosphere. To this purpose, we must firstly introduce the concept of **thermal engine**.

Transformation of free energy into mechanical work

Let us consider again a perfect gas inside a box, closed by a mobile upper surface on which some weights (for example w_1, w_2) can exert their gravitational force. The lower wall of the box is in contact with a thermostat at temperature T (see the schematic in Fig. 2.2a, b). The pressure is:

$$P_{in} = \frac{(w_1 + w_2)}{S} \quad (2.51)$$

To keep things simple, we take that there is no air surrounding and therefore absence of friction, and moreover that the weight of the moving wall is negligible compared to the w_i 's. At the initial equilibrium, the mobile wall is found at a position h_{in} (Fig. 2.2a).

Now the weight w_2 , for example by sliding horizontally without friction. Under the reduced weight, the internal pressure will raise the wall to the new position h_{fin} (Fig. 2.2b), with the pressure decreasing to:

$$P_{fin} = \frac{w_1}{S} \quad (2.52)$$

At constant temperature, and with no change of internal energy (perfectly reversible transformation), the free energy of the perfect gas equals the entropy. From the Sackur-Tetrode equation:

$$\Delta G = S_{fin} - S_{in} = -Nk_B T \ln \left(\frac{V_{fin}}{V_{in}} \right) = -Nk_B T \ln \left(\frac{h_{fin}}{h_{in}} \right) \quad (2.53)$$

On the other hand, the perfect gas equation of state gives for the final state pressure:

$$P_{fin} S h_{fin} = w_1 h_{fin} = Nk_B T \quad (2.54)$$

By setting $x = (h_{fin} - h_{in})/h_{fin}$, $0 < x < 1$, we can compare the mechanical work performed by the perfect gas, to the corresponding free energy variation. The work is just:

$$P_{fin} \Delta V = \left(\frac{w_1}{S} \right) S (h_{fin} - h_{in}) = w_1 (h_{fin} - h_{in}) \quad (2.55)$$

Therefore we have:

$$w_1 (h_{fin} - h_{in}) = -Nk_B T \ln(1 - x) \quad (2.56)$$

for the work, and

$$\Delta G = x Nk_B T \quad (2.57)$$

for the free energy. Since for $0 < x < 1$ it is always $x < -\ln(1 - x)$: as predicted, the quantity of work performed by the perfect gas is always less than the available free energy.

The ratio between the two:

$$r = \frac{-x}{\ln(1 - x)} \quad (2.58)$$

is bigger the smaller is x , i.e. for small variations $(h_{fin} - h_{in})$ or, again, for small values of w_2 compared to w_1 . We can imagine to split w_2 into many smaller weights $w_i \ll w_2$, to increase the efficiency (quantity of work extracted for a given available energy). This is an interesting conclusion: the transformation of energy into work is more effective if performed in many small steps, rather than in one big step. At the limit of a virtually infinite number of infinitesimal steps, this approaches the definition of a quasi-static, or **adiabatic** process, in which no heat is exchanged at all, and all the energy is turned into work.

By pursuing the reasoning from the situation displayed in the greybox on p. 31, consider the right half of Fig. 2.2. Let us imagine to slide the box V above another thermostat at a temperature $T' < T$ (always horizontally, without friction and without doing work against the gravity). The perfect gas will cool down, and therefore its volume will decrease until getting back to the height h_{in} (Fig. 2.2c). At this point, we put back the weight w_2 into place (Fig. 2.2d), and slide back the box above the first thermostat at $T > T'$ (Fig. 2.2a). The cycle can start over in this weird machine, which is in fact an idealised representation of a cyclic thermal engine. At every cycle, a fraction of thermal energy is turned into mechanical work (lifting the weight). Every cycle increases a fraction of the entropy of the surrounding environment (the thermostats), with a transfer of heat from the reservoir at high temperature T , to the reservoir at lower temperature T' . In the real world, without an external intervention, such a process cannot continue indefinitely: the entropy of the two thermostats will at some point become identical, $\Delta S = 0$, and so will their temperature.

Such a process is completely reversible, meaning that the energy is the same at the beginning as at the end of the cycle, $\Delta E = 0$. Let us restart from the ‘caloric’ definition of the entropy, Eq. (2.25), from which we can compute the net variation of entropy between the two thermostats:

$$\Delta S_{tot} = \Delta Q \left(\frac{1}{T'} - \frac{1}{T} \right) = \frac{\Delta W}{T} \quad (2.59)$$

Here $\Delta W = P(S\Delta h)$ is the total mechanical work resulting from the displacement of the mobile wall, and $\pm \Delta Q$ is the quantity of heat taken from the reservoir at temperature T and transferred to the reservoir at temperature T' , with the appropriate algebraic sign. Since the variation of entropy must always be positive, in order to have

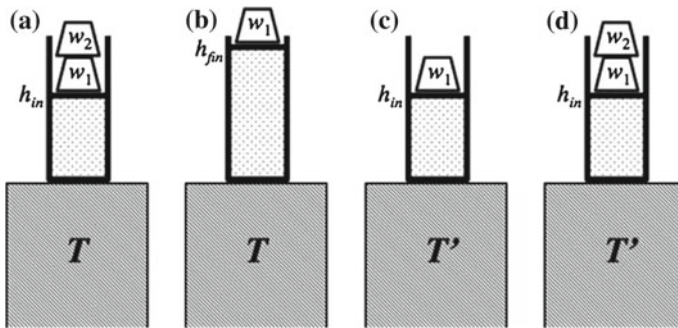


Fig. 2.2 Schematic of an ideal cyclic thermal engine, constituted by perfect gas enclosed in a box of cross section S and variable height h , with volume $V = Sh$. The top-side wall, of negligible mass compared to the weights w_i , can move without friction to adjust against the changing internal pressure of the gas. The volume V can be cyclically moved atop thermostats at different temperatures $T > T'$, steps (a) through (d)

a net work $\Delta W > 0$ it must clearly be $T' < T$. In other words, *it is the difference of temperature which keeps the thermal engine working*. This was just the principle established by Carnot in his experiments.

An ideal engine could convert all the energy into work, by an ideal (adiabatic) transformation with $\Delta Q = 0$. In all practical cases, this is forbidden by the Second Principle. However, the equation suggests that the efficiency of the engine can be increased by increasing the temperature difference between the two thermostats.

We will see in the next Chapter that the most important molecules making up the living systems often have a positive enthalpy of formation, i.e. the enthalpy of the products is less negative than that of the precursors, as well as a negative entropy of formation. As a result, also the ΔG of complex molecules such as proteins is positive, compared to the simpler building blocks (the amino acids) from which they are assembled. With a positive ΔG , such a chemical reaction would be forbidden. Therefore, to not violate the Second Principle, the formation and maintenance of living systems demands a continuous flux of free energy, as well as a supply of 'negative entropy', from the external environment, just like a refrigerator needs a continuous supply of energy to keep the temperature lower than the outside. In other words, the biosphere is not an isolated system, and is constantly maintained in a condition of non-equilibrium.

As suggested by the Eq.(2.59) above, the large temperature difference between the Earth surface (280–300 K) and the Sun (5,800 K) is at the basis of any process of energy transformation on our planet (Fig. 2.3). The radiant energy is transported by electromagnetic radiation, with its combined oscillating electric and magnetic fields. The amount of energy contained in the electromagnetic wave is directly proportional

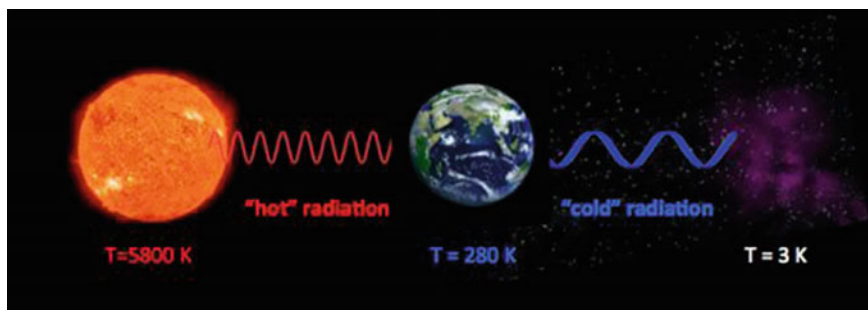


Fig. 2.3 Exchanges of radiation energy. The Sun irradiates 'hot' radiation, at a temperature of 5,800 K and short wavelength (visible and ultraviolet). A very small fraction of this energy is captured by the Earth, and reemitted at the temperature of 280–300 K and longer wavelength (infrared). This heat of radiation is dispersed towards the dark sky, which is at the temperature of 3 K. (The temperature of the sky being a remnant of the Big Bang, and constantly cooling with the expansion of the universe.)

to its frequency, ν , or inversely proportional to its wavelength λ (another result which, besides having a classical interpretation, must wait quantum mechanics for a thorough explanation).

The energy of the solar radiation is not delivered just at one frequency, but extends over a continuous distribution of intensity covering a wide range of frequencies. The maximum of the solar energy distribution (see Fig. 2.7 in Problem 2.9 below) is centred approximately around the green band of the visible portion of the electromagnetic spectrum ($\lambda \sim 550$ nm), with also an important fraction of energy delivered in the ultraviolet (UV) spectral region (λ 's of a few tens nm). On the other hand, the distribution of the radiation energy emitted by the Earth, being associated to a much lower temperature, is rather centred on the infrared (IR) band, at wavelengths around a few micrometers (see again Fig. 2.7). From the point of view of a biological system, this difference is extremely important. The visible and UV light is useful for initiating the photochemical process of chemical synthesis (and it is moreover dangerous for photo-labile proteins). On the other hand, the energy of IR waves is not useful for biological processes, other than keeping the organisms at a reasonably warm temperature. The reason is in the quantum mechanical energy-wavelength correspondence. The energy of a radiation corresponding to a temperature of 5,800 K is $k_B T = 8.6 \times 10^{-5} \times 5,800 \sim 0.5$ eV, a figure comparable with the difference between the discrete (stationary) energy levels of atoms and molecules, therefore capable of inducing chemical reactions. On the other hand, the energy of the IR waves at $T = 280$ K is about $k_B T = 8.6 \times 10^{-5} \times 280 \sim 0.03$ eV, rather comparable with the vibrational energy levels of the molecules, thus capable of producing just some waste heat.

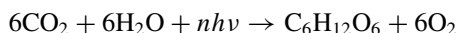
Due to the substantial temperature difference, the Sun appears as a source of negative entropy, in that the radiation arriving on the Earth allows to accommodate the apparently negative difference in the entropy balance ("apparently" only if the biosphere is considered as a closed system, remember the discussion on p. 20 about the introduction of the concept of free energy). The foremost case is that of photo-synthetic reactions, for which it can be $\Delta S_{ph} < 0$ provided this is accompanied by enough Sun-produced entropy, such that:

$$\Delta S = \Delta Q \left(\frac{1}{300} - \frac{1}{5,800} \right) > -\Delta S_{ph} \quad (2.60)$$

After the works of E. Schrödinger and others (see "Further reading" at the end of the Chapter), it has become familiar the idea that in biology it would be common to observe the transfer from a high-entropy source (the Sun) to a low-entropy drain (the Earth and its biosphere), without any violation of the Second Principle [14].

2.7.1 A Synthesis of Photosynthesis

Photosynthesis is an extremely complex mechanism, including very many chemical reactions in chain, which can be conveniently subdivided into two phases: the *light* phase, during which the photochemical reactions powered by the sunlight take place; and the *dark* phase (also called *Calvin's cycle*), whose main characteristic is the fixation of carbon. A highly simplified, synthetic writing of this chain of reactions, looking only at the start and endpoints, would be:



carbon dioxide + water + radiation \rightarrow glucose + oxygen

Once the energy of the solar radiation, $E = nh\nu$, is captured by the plant leaves, a fraction η is utilised in the photosynthetic cycles and will appear as $\Delta G_{ph} > 0$. On the other hand, the fraction $(1 - \eta)$ is reemitted at room temperature, in the form of heating of the glucose and oxygen molecules. According to Eq. (2.60), the entropy balance must accommodate the negative quantity ΔS_{ph} , by a fraction η of the solar entropy at least equal or larger:

$$\eta(n \cdot 1.986 \cdot 5,800) \left(\frac{1}{300} - \frac{1}{5,800} \right) + (-\Delta S_{ph}) > 0 \quad (2.61)$$

with $E = h\nu = RT$, and $R = N_A k_B = 1.986 \text{ cal K}^{-1} \text{ mol}^{-1}$, or $8.31 \text{ J K}^{-1} \text{ mol}^{-1}$, the universal gas constant. The factor n comes from the number of *photons* typically needed for the light phase (in quantum physics, the energy of the electromagnetic radiation is defined in terms of finite packets of energy, called photons, each carrying an energy $E = h\nu$, with $h = 6.62 \times 10^{-34} \text{ J-s}$ the Planck constant). This number is found to be somewhere between $n = 5$ and 10. An estimate of ΔS_{ph} can be given by calculating the difference between the free entropies of the (products) – (reactants), all taken in the gaseous state but water, thus obtaining $\Delta S_{ph} = (209 + 6 \cdot 205) - (6 \cdot 213.6 + 6 \cdot 69.9) = -262 \text{ J K}^{-1} \text{ mol}^{-1}$. Therefore, for the photosynthesis to be compatible with the Second Principle, the yield of photochemical reactions must be:

$$\eta > 0.35 \quad (2.62)$$

or even less if $n > 5$. In fact, in dedicated laboratory experiments it is found that the photosynthesis has a global yield of about $\eta = 0.50$. On the other hand, naturally occurring photosynthesis has a much lower efficiency because of several factors, such as the many different molecules participating in the ensemble of reactions (more complex than the simple picture above), or the losses from light-harvesting molecules, which may decay rapidly by electron-transfer reactions and other competing mechanisms, in addition to fluorescence and stimulated emission. However, the core

of our discussion is not changed: the total system entropy is globally increased, and the inflow of entropy from the Sun is indeed compensating the apparent decrease in entropy during the chemical synthesis process.

It is equally important to underscore that the simple availability of energy (and entropy) is not enough to support the living systems, in the same way that it is not enough to sit on a tank of gasoline to make a car run. It is necessary to have a converter, which transforms the available energy in useful work (mechanical, chemical, electromagnetic), the equivalent of a combustion engine for a car, which turns the chemical energy of gasoline into rotating motion of the wheels of the car. For the case of biological systems, the appropriate use of the “negative” entropy flux is done through encoding and decoding the information in the genetic code, stored in the DNA of every living being.

Terrestrial organisms have learned, during the evolution, to extract with a good efficiency the energy and entropy from the solar blackbody radiation. From the point of view of the thermodynamics balance, life on Earth is organised in the form of a pyramid. At the bottom of the pyramid we find the species capable of synthesising the base organic compounds, such as the carbohydrates, by directly using the solar radiation via the photosynthesis. Without such organisms (plants), life would not exist as we know it. Notably, the energy stored in the biomass by the photosynthesis is a ridiculous 0.023 % of the total energy received on the Earth’s surface. Nevertheless, this small amount is enough to sustain the growth and development of all the living organisms in the upper levels of the pyramid (including the energy consumption of all the human-made machines, thanks to the energy stored in fossil fuels throughout the ages).

Animals of the upper levels of the pyramid cannot directly use the solar energy and entropy, as they depend on the photosynthesising species. During the assimilation and digestion of food, energy and entropy stored by the plants in the base organic compounds are extracted and used by the herbivores, and from these they pass on to the other animal species (carnivores and omnivores). During the geological ages, natural selection has been operating in such a way to favour the species which are most effective in the process of thermodynamic extraction of the pristine stored energy and entropy. In the next Chap. 3 we will see how entropy can be stored in DNA, and used to make the building blocks of cells and tissues. Notably, sexual recombination of the genetic material has proven to be a very effective way of protecting the ‘information’ entropy stored in the DNA, against the unavoidable information degradation due to the steady increase of entropy of the surrounding environment.

The energy stored in a tree

Let us take a tree of 20 years of age, which has been growing up to a mass of 500 kg, of which 400 kg in wood and the rest in leaves and circulating water. We know that burning dry wood gives about 4.05 kWh/kg, or 14,580 kJ/kg. Can we estimate from such a simple data the energy stored in the Earth's biomass?

Starting from the value of the solar constant, $C_S = 1366 \text{ W m}^{-2}$, we can calculate that a surface of 1 m^2 of leaves receives:

$$E = 1,366 \cdot (20 \cdot 365 \cdot 12) = 119.7 \times 10^6 \text{ Wh} = 430.8 \times 10^6 \text{ kJ} \quad (2.63)$$

of energy over the 20 years of its life (we multiplied the number of hours in a day only by 12, to consider that on average half of the day is actually night).

Let us take that the photosynthesis yield is about 2%, therefore the energy store would be $8.616 \times 10^6 \text{ kJ}$ per m^2 of leaves over the 20 years. A typical tree should have rather 200–300 m^2 of leaves (a gross estimate, not taking into account the tree shape, living latitude, etc.), which gives a total available energy of:

$$E = 300 \cdot 8.616 \times 10^6 \approx 2.585 \times 10^9 \text{ kJ} \quad (2.64)$$

This energy is necessary for all the vital functions of the tree including its growth, therefore our calculation will be an underestimate, when considering that all the energy is instead used only for growth. The energy accumulated in the wood is $E_{acc} = 14,580 \cdot 400 = 5.83 \times 10^6 \text{ kJ}$. This represents a fraction:

$$f = E_{acc}/E = (5.83 \times 10^6)/2.585 \times 10^9 = 0.0022 \quad (2.65)$$

i.e., 0.22 % of the total energy stored. By accounting that about 30 % of the Earth's surface is covered by forests, we get about 0.066 %, i.e., only 2–3 times bigger than the accepted value of 0.023 %. This seems a rather decent estimate, given our very rough approximations.

All this energy stored in the trees that lived on the early Earth would be found later deep in the geological layers, under the form of coal. In fact, all hydrocarbons of coal are derived from the cellulose of the biomass, fossilised during the millions of years of life on Earth, mostly during the Carboniferous age (about 300 millions of years ago). Likely, the first stage of the conversion was initiated by bacteria, which digested the organic matter by producing methane, CO_2 and oxygen; this may have been followed by stages of anaerobic decomposition, in which acids should be the waste, thus increasing the pH up to levels at which all bacteria would die. Once this proto-carbon material got buried under hundreds of meters of soil, the temperature would rise to $\approx 100^\circ\text{C}$ and the pressure to tens of atmospheres. In such conditions, carbonification could take place by turning into less and less volatile compounds, firstly forming lignin, then coal, and finally anthracite. The composition of such materials and the chemistry of their processes are still largely unknown (Fig. 2.4).

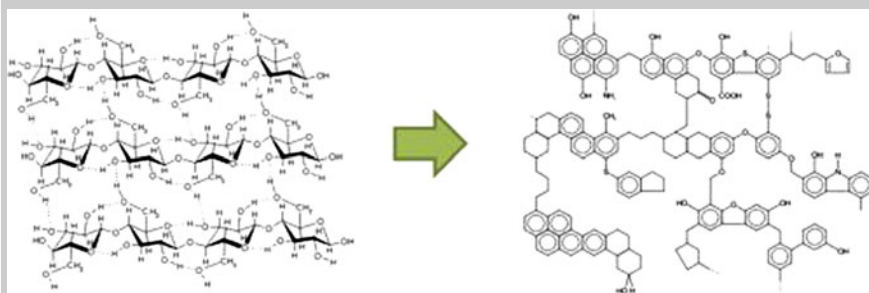


Fig. 2.4 Left From cellulose (left) to coal (right). Bonds between sugar molecules in cellulose long chains are protected against chemical attack, because cellulose tangles itself into tight microfibrils. Harsh conditions such as high temperatures and pressures are required to complete the carbonification

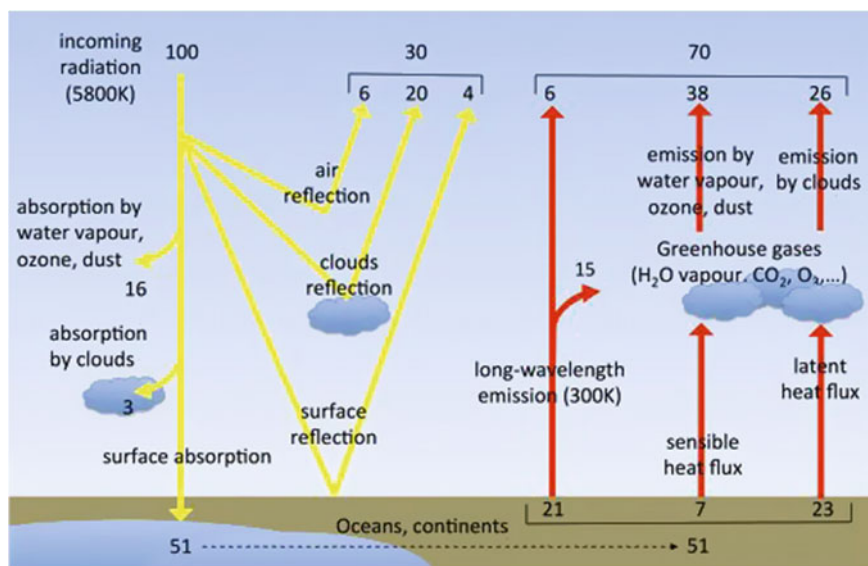


Fig. 2.5 Schematic representation of the energy balance in the greenhouse effect. An amount of solar radiation energy equal to 100 (*on the left*) is assumed to reach the Earth. In *yellow arrows*, the ways this incoming energy is reflected or stored in the atmosphere. The 51 % absorbed by the Earth's surface is re-irradiated at lower radiation temperature (indicated with *red arrows*, on the *right of the figure*). Some of this outgoing radiation is again captured and stored in the atmosphere. The sum of the *yellow* and *red* fractions stored is at the origin of the greenhouse effect

2.8 Energy from the Sun

As shown in the Fig. 2.5, the biosphere receives the solar energy in the form of a 'gas of photons' with high energy and low entropy, and emits residual thermal energy in the form of a similar gas of radiation, with lower energy and higher entropy. The transformation from short- to long-wavelength photons increases the entropy of the radiation (this could be calculated from the equations of the quantum radiation theory), and this entropy compensates the decrease of entropy in building organic molecules.

A theoretical calculation of the total energy arriving on the Earth is affected by a considerable uncertainty, since it requires a number of approximations. This total energy is termed **solar constant**. Such a quantity can be measured by a bolometer⁵ mounted on a satellite, and integrating all the radiation (IR, visible, UV, X, gamma,...) which arrives on the Earth in the direction perpendicular to the external surface of the atmosphere. Although such a quantity may have quite large fluctuations as a function

⁵A bolometer is a detector of electromagnetic radiation. Its principle is simple: it converts the absorbed light radiation into heat. By choosing an absorbing material whose electrical resistance changes with the temperature, the incident energy can be estimated by measuring the impedance variation in the detector.

of the latitude, of the time of the day, and of the period of the year, an average value for the whole Earth surface has been assessed, $C_S = 1366 \text{ W m}^{-2} = 1.366 \text{ GW km}^{-2}$. By taking the average Earth's radius $R_T = 6366 \text{ km}$, the total fluence (energy flux times the total surface, i.e., $[E][\text{T}^{-1}]$) of solar energy captured by the Earth is:

$$P_T = C_S \pi R_T^2 = 1.7 \times 10^8 \text{ GW} \quad (2.66)$$

It will be noted that, on average, the flux on the surface of our planet is only 1/4 of this value, since the flux across the circle of surface πR_T^2 must be distributed over the sphere of surface $4\pi R_T^2$. Overall, the Earth captures only a fraction equal to 2×10^{-9} of the total energy emitted from the Sun, and this is already not bad!

Moreover, we note that the energy balance is zero on average: as much energy is received from the Sun, as it is sent back from the Earth in the form of low-frequency radiation.⁶ Since the energy of emitted photons is lower, their density must be higher, however the integral of flux times energy is constant.

2.8.1 The “Greenhouse” Effect

By looking at the schematic diagram in Fig. 2.5, which represents the balance of the incoming and outgoing radiant energy, we see that the energy arriving on the Earth's surface from the Sun is reflected by about 30 % by the molecules making up the air, the water vapour of the clouds, and the white parts of the surface (snow, glaciers). This fraction represents the **albedo**, A , of the Earth's surface.

Of the remaining 70, 19 % is absorbed by the molecules of the atmosphere and clouds. Therefore, only 51 % of the incident radiation energy is delivered to the surface. This is the same quantity of energy which is reemitted, however with an energy spectrum (density of energy as a function of radiation wavelength) very different from that of the incident radiation.

Of the 51 % reemitted, 21 % is directly radiant energy. Of this, 6 % goes directly into the outer space, and 15 % is again recaptured by the molecules in the atmosphere. Another 7 % of the reemitted energy is used to heat the lower layers of the atmosphere, say the troposphere (500 m), by convection. The remaining 23 % ($21 + 7 + 23 = 51$ %) is used as latent heat in the phase transformations of the water cycle, namely evaporation and condensation of clouds.

By summing up the various contributions, we find that molecules in the atmosphere and clouds capture 64 % of the total energy (19% direct + 45 %). This fraction is reemitted in the form of infrared (IR) radiation. The energy captured is irradiated in all directions, notably both towards the upper space and the lower Earth's surface.

⁶A very small quantity of energy is contributed by the internal heat of the Earth itself, due to the primordial heat, the decay of radioactive elements in the Earth's interior, and the heat of crystallisation of the core materials. This contribution is about 1/10,000 of the amount of external energy received by the Sun.

This descendent flux of energy adds to the incident solar energy, and allows the surface temperature to attain an average value of about $+15^{\circ}\text{C}$, against the about -18°C predicted in the absence of a partially absorbing atmosphere. Moreover, it should be considered that with such lower average temperature, the extent of ice caps on the surface would be increased, with a corresponding increase in the albedo. More reasonable calculations predict an average surface temperature in the absence of the atmosphere around -100°C .

This effect of heating of the Earth's surface, originating from the reflection of part of the energy by the atmosphere, is called **greenhouse effect**, since it makes the Earth surface to resemble to a covered greenhouse. This is obviously a very beneficial and desirable effect by all the living organisms. It is due to the "greenhouse gases" contained in the terrestrial atmosphere: mainly water vapour (the principal contributor to the heating and cooling of Earth's surface), carbon dioxide CO_2 , and methane CH_4 , plus a number of minor constituents.

The name of "greenhouse" comes from the analogy of the Earth's atmosphere with a real greenhouse. For such a kind of construction, once exposed to the sunlight, the inside air temperature is higher than the outside even in the absence of internal heat generation (which in a real greenhouse could be added). This is due to the different transparency of the glass of the walls to the radiation, that is quite good for the high-frequency light coming from the Sun, and pretty bad for the infrared radiation emitted by the Earth's surface at lower frequency. In practice, the glass behaves as a sort of energy valve, letting easily the energy to get in while being less good at letting energy out. In the analogy, the entire solid and liquid mass of the Earth's crust is the greenhouse, and the atmosphere is the equivalent of the glass. The possible problems in the equilibrium originate from the fact that the infrared transparency is even more decreased when the concentrations of greenhouse gases increase, thereby increasing to higher values the surface temperature.

Natural or artificial perturbations of the atmospheric concentrations of the greenhouse gases can alter the equilibrium of the radiation exchanges, with the effect of changing the amount of energy stored in the atmosphere. Such a disequilibrium could entail a long-term change in the atmospheric temperature, and therefore in the surface temperature. While glacial periods and warmer periods have naturally alternated on the Earth surface for millions of years, there has been in recent years a concern about the possible long-term effects of man-made alterations of the greenhouse gases concentrations, especially due to the atmospheric increase of CO_2 and CH_4 levels following the burning of large quantities of fossil fuels. This effect has been termed the **global heating** problem, since most of the indications point towards an increase of the average surface temperature, although there are also predictions based on computer models which would rather indicate a decreasing trend. This is a very complicate problem, involving the contribution of widely ranging scientific knowledge, from physics, mathematics, geography, oceanography, geology, and so on. In the following we will develop a simple model to describe some of the basic effects linked to the variation of greenhouse gases concentration.

Before going into the details of the greenhouse effect, however, a more basic question should be posed, on the basis of the thermodynamics. It was amply demonstrated in the beginning of this chapter that two parts of a system at different temperatures put in contact will evolve in the direction of attaining a common equilibrium temperature, pushed by the maximisation of the total entropy. Why the system composed by the Sun and Earth is not at the same temperature? How it is possible that the two systems, which can exchange energy without limits, should have not yet attained the equilibrium after about 10^{10} years of the age of the universe? The answer is “no way”. The radiation continuously emitted from the Sun is an expanding gas, which changes its energy density as it expands, somewhat like a perfect gas that cools down while continuously expanding. As a consequence, the system is *never* at thermodynamic equilibrium. However, it must be noted that the ‘cooling’ of the radiation cannot be explained by the concepts of classical physics, but can be understood only by making recourse to quantum mechanics.⁷

2.8.2 The Temperature of the Earth’s Surface

We consider the Sun as an ideal black body at a temperature $T_S = 5,800$ K. This is a good approximation, despite the fact that the Sun is not black at all! (That is a joke, in fact by the term *black body* in physics it is meant an object ideally capable of absorbing all the radiation it receives, therefore appearing black at low temperatures. On the other hand, a black body also radiates energy at the same temperature, since it is in equilibrium.)

For an ideal black body with spherical shape, the Stefan-Boltzmann law⁸ gives the following expression for the power emitted, as a function of the temperature T and the radius R of the emitting sphere:

$$P = 4\pi R^2 \sigma T^4 \quad (2.67)$$

This equation states that the power emitted (energy/unit time) is the product between the emitting surface (a sphere, in this case) and the fourth power of the temperature, times the Stefan-Boltzmann constant, $\sigma = 5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$.

Now let us try to estimate the Earth’s surface temperature by considering also the Earth as a black body, but at a lower temperature than the Sun. Consider that the Sun

⁷It is interesting to note that, even after a very complex mathematical treatment based on quantum mechanics, the result for the difference in entropy in cooling from a temperature T_1 to a temperature $T_2 < T_1$ is very close to the classic result for the entropy variation of a perfect gas [15]. The important difference between the two treatments is the concept of temperature, which is completely different for a classical gas and a quantum “gas of radiation”.

⁸The Stefan-Boltzmann law is physically justified only in the quantum mechanical treatment of the radiation; however, Boltzmann derived it by a fully classical argument, see Problem 2.6.

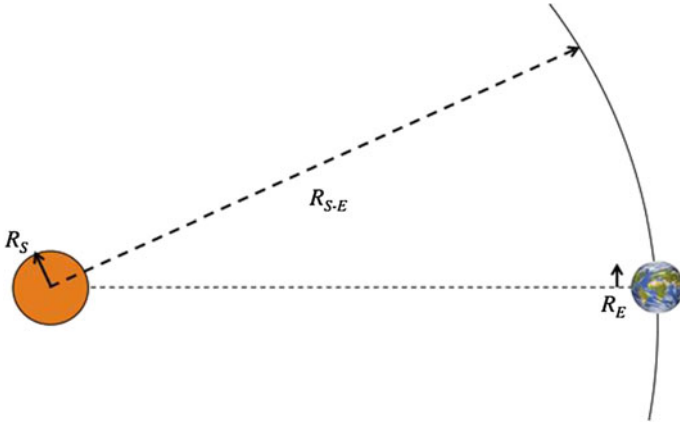


Fig. 2.6 The Sun emits all of its radiant power across its surface, which measures $(4\pi R_S^2)$ m². This same energy, traveling across the empty space, is distributed over the concentric sphere with radius R_{S-E} . The Earth intercepts a fraction of this energy, equal to the ratio of the projected circle with the same radius of the Earth, R_E , to the surface irradiated at the Sun-Earth distance. Then, this energy is distributed over the entire Earth's surface, again a sphere with radius R_E

distributes all of the power P_S irradiated across its surface, with radius $R_S = 6.9 \times 10^9$ m, in all the radial directions:

$$P_S = 4\pi R_S^2 \sigma T_S^4 \quad (2.68)$$

The Earth is located at the Sun-Earth distance, $R_{S-E} = 1.5 \times 10^{11}$ m. When the radiation reaches such a distance, it is distributed over a sphere with this same radius, concentric with the Sun (see the schematic in Fig. 2.6).

Moreover, consider that the Earth intercepts only a fraction of the Sun power, on a circle of radius $R_E = 6.366 \times 10^6$ m projected on this large sphere, and multiplied by the fraction $(1 - A)$ since the effect of the albedo is to reduce the amount of radiation P_E absorbed by the Earth:

$$P_E^{abs} = P_S \left(\frac{\pi R_E^2}{4\pi R_{S-E}^2} \right) (1 - A) \sigma T^4 \quad (2.69)$$

Finally, consider that Earth reemits the power absorbed, over all of its surface:

$$P_E^{emit} = 4\pi R_E^2 \sigma T_E^4 \quad (2.70)$$

We impose that at equilibrium, $P_E^{abs} = P_E^{emit}$. Therefore, the Earth's temperature is easily found:

$$T_E = \left(\frac{1}{4} \right)^{1/4} \left(\frac{R_S}{R_{S-E}} \right)^{1/2} (1 - A)^{1/4} T_S = 255 K \quad (2.71)$$

This shows that, to a first approximation, the observed surface temperature of the Earth is only determined by the distance and structure (temperature, size) of the Sun. As anticipated, without any account for the presence of the atmosphere we end up with an estimate of about 20° below the zero Celsius. By our standards, this would be a very inhospitable planet!

To improve the predictions of this simple model, which gives a quite cold surface temperature, we must add the effect of the atmospheric layers. For the sake of simplicity, let us take just one layer of atmosphere, at the temperature T_A , at approximately the same distance from the Sun as the Earth, R_{S-E} . This layer captures a fraction $f < 1$ of the power P_E^{emit} emitted by the Earth surface. This fraction of power sequestered by the atmosphere will be reemitted toward the space, one half back in the direction of the Earth (thus contributing to the surface heating), and the other half toward the upper sky. The balance of the power between the Earth and the atmosphere is then:

$$f(\sigma T_E^4) = 2f(\sigma T_A^4) \quad (2.72)$$

On the other hand, the balance equation between absorbed and emitted power from the Earth, $P_E^{abs} = P_E^{emit}$, is modified as:

$$P_E^{abs} = P_S \left(\frac{\pi R_E^2}{4\pi R_{S-E}^2} \right) (1 - A) = (1 - f) (4\pi R_E^2) \sigma T_E^4 + f (4\pi R_E^2) \sigma T_A^4 \quad (2.73)$$

which can be simplified, by using the Eq. (2.72), as:

$$\left(\frac{R_S}{R_{S-E}} \right)^2 \frac{1 - A}{4} T_S^4 = (1 - f) T_E^4 + f T_A^4 = \left(1 - \frac{f}{2} \right) T_E^4 \quad (2.74)$$

The corrected expression for the Earth's surface temperature is:

$$T_E = \left\{ \left(\frac{R_S}{R_{S-E}} \right)^2 \frac{(1 - A)}{4(1 - \frac{f}{2})} \right\}^{1/4} T_S \quad (2.75)$$

This equation contains f as an unknown parameter. We can use it in reverse, to obtain the average temperature at the Earth surface, $T_E = 288$ K. For this, we must set $f = 0.78$ for the fraction of power absorbed by the atmosphere. As a quality check, by imposing this condition we can derive from Eq. (2.72) the average temperature of the atmosphere layer as $T_A = 242$ K, which is actually a very good estimate for the temperature of the troposphere at the height of about 7 km above the surface.

Besides, the fitted value of $f = 0.78$ does not coincide with the fraction of solar power sequestered and reemitted by the greenhouse gases, which is rather 0.38 (see Fig. 2.5 above). However, apart from the more or less accurate numerical values, it is

interesting to note that already with such a very simple model, we can obtain a correct qualitative correlation in the temperature response: by increasing the gas concentration of the atmospheric layer, for example in methane or CO_2 , we would increase the absorption fraction f , which would entail a parallel increase of the temperature at the Earth's surface from Eq. (2.75). On the other hand, such simple correlations should not be pushed too far. The modelling of Earth's climate is a dauntingly more complex task, comprising a wealth of physical-chemical, atmospheric, oceanographic, and geophysical phenomena, which we have not even hinted at here, and commands the use of the largest computers in the world.

In particular, our very simplified model takes the Earth as a sphere with a homogeneous surface and homogeneous atmosphere layers, which is far from truth. The large-scale differences in the surface distributions of land and biomass, water, ice, clouds, are the very motors of Earth's climate, and cannot be neglected. The water vapour, which represents the major barrier to cooling by radiation emission by the T^4 law Eq. (2.67), generally is maximum at the surface near the tropics, and sharply decreases with both altitude and latitude. Because of this layer mostly opaque to infrared radiation, heat is firstly carried away from the surface by fluid convection, starting from the cloud towers of the tropics, which then carry most of the heat upward and to the poles, whence it is possible for thermal radiation emitted from these levels to escape into space.

The qualitative variations of the Earth surface also have great implications for the definition of the average surface temperature, a concept that has been recently popularised also by the media. Climate cannot be associated to a single temperature, rather the differences of temperatures drive the climate processes and create the storms, winds, sea currents, and everything that makes up the climate. The Earth surface has a large number of interacting components, which one cannot just add up and average: it would be as meaningless as calculating the "average phone number" in the phone book. If temperature decreases at one point and increases at another, the "average temperature" would be unchanged, but the thermodynamic forces would be totally different, and so would be the climate. If, for example, we measure 20°C at one point on the surface and 30°C at another point 40 km away, we would be tempted to attribute an average temperature of 25°C to that area; but if we measured 25°C at both places, the average would be still 25°C . However, these two situations would give rise to two entirely different climate reactions, because in the former case one would have an air pressure difference and strong winds, while in the latter case there would be a calm and pleasant day.

Appendix A: Some Useful Mathematical Tools

The Gamma Function

The mathematical Gamma function $\Gamma(x)$ is an extension of the factorial function, valid for both real and complex numbers. Its analytical definition, due to the French mathematician Adrien-Marie Legendre (1752–1833), is:

$$\Gamma(t) = \int_0^{\infty} x^{t-1} e^{-x} dx \quad (2.76)$$

with t a real or complex number. This (only apparently) difficult integral can be integrated by parts:

$$\int_0^{\infty} x^{t-1} e^{-x} dx = \left[\frac{x^t}{t} e^{-x} \right]_0^{\infty} + \frac{1}{t} \int_0^{\infty} x^t e^{-x} dx \quad (2.77)$$

(the term within [...] being equal to 0), to obtain the recurrence formula of the Gamma function:

$$\Gamma(t+1) = t\Gamma(t) \quad (2.78)$$

From this result, it is immediately obtained $\Gamma(n) = (n-1)!$ when t is an integer n , which justifies the definition of *generalised factorial*. From the same integral, it is also easily seen that $\Gamma(1) = 1$.

In thermodynamics, it is often necessary to calculate the Gamma function for half-integer argument, $\Gamma(n/2)$. The recurrence formula (2.78) above can still be used, but ending up with the task of calculating the last term, $n = 1/2$. In the next Section, this will be shown to be:

$$\Gamma(1/2) = \sqrt{\pi} \quad (2.79)$$

Dirac's Delta Function

The Dirac delta function, indicated as $\delta(x)$, is a real function that is zero everywhere except at $x = 0$, and with an integral equal to 1:

$$\int_{-\infty}^{+\infty} \delta(x) dx = 1 \quad (2.80)$$

It was introduced by the English physicist Paul A. Dirac, and it may be physically interpreted as the density of an idealised point mass or point charge situated at the origin. Although it makes little sense mathematically, the delta “function” becomes

meaningful only when inside an integral, as the limit of a distribution becoming infinitely narrow about $x = 0$ while preserving its unitary integral.

The delta function admits any n -th order derivatives $\delta^{(n)}(x)$, a Fourier transform, and several other analytical manipulations typical of a true function. Some interesting properties of the delta are:

$$\int_{-\infty}^{+\infty} \delta(ax) dx = \int_{-\infty}^{+\infty} \delta(u) \frac{du}{|a|} = \frac{1}{|a|} \quad (2.81)$$

$$\int_{-\infty}^{+\infty} \delta'(x) f(x) dx = - \int_{-\infty}^{+\infty} \delta(x) f'(x) dx \quad (2.82)$$

$$\int_{-\infty}^{+\infty} f(x) \delta(x - x_0) dx = f(x_0) \quad (2.83)$$

$$\int_{-\infty}^{+\infty} f(x) \delta^{(n)}(x - x_0) dx = (-1)^n f^{(n)}(x_0) \quad (2.84)$$

The last two properties present the delta function as a kind of filter, by which a particular value of another function $f(x)$ can be extracted.

The Dirac delta may be thought as a continuous- x analog of the discrete Kronecker delta, $\delta_{ij} = 1$ for $i = j$ and 0 if $i \neq j$, which selects a discrete value out of a series $\{a_i\}$, $i = 1, \dots, n$:

$$\sum_{i=-\infty}^{\infty} a_i \delta_{ik} = a_k \quad (2.85)$$

Among the useful applications of the Dirac function, for a discrete distribution consisting of a set of points $x = \{x_1, \dots, x_n\}$, with corresponding probabilities $\{p_1, \dots, p_n\}$, a continuous probability density function $f(x)$ can be written as:

$$f(x) = \sum_{i=1}^n p_i \delta(x - x_i) \quad (2.86)$$

Gauss and Euler Integrals

Throughout this book, we will encounter several times integrals of the type:

$$I(n) = \int_u^v x^n e^{-\alpha x^2} dx \quad (2.87)$$

with $[u, v] = [-\infty, +\infty]$ or $[0, +\infty]$.

The simplest of these integrals, $I(0)$, can be solved on the infinite real axis by an ingenious trick, due to Siméon-Denis Poisson (1781–1840), if we start from its square:

$$\begin{aligned}
 I^2(0) &= \left[\int_{-\infty}^{+\infty} e^{-\alpha x^2} dx \right]^2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\alpha(x^2+y^2)} dx dy = \\
 &= \int_0^{2\pi} d\theta \int_0^{+\infty} r e^{-\alpha r^2} dr
 \end{aligned} \tag{2.88}$$

The last identity follows from the change of Cartesian to polar coordinates, $x = r \cos \theta$, $y = r \sin \theta$, $dx dy = r dr d\theta$. The integral is now easily calculated:

$$I^2(0) = \int_0^{2\pi} d\theta \int_0^{+\infty} r e^{-\alpha r^2} dr = \int_0^{2\pi} d\theta \left[-\frac{1}{2\alpha} e^{-\alpha r^2} \right]_0^{+\infty} = 2\pi \cdot \frac{1}{2\alpha} = \frac{\pi}{\alpha} \tag{2.89}$$

from which we get the basic result $I(0) = \sqrt{\pi}$, for $\alpha = 1$.

If we take the integral on the interval $[0, +\infty]$ and make the substitution $x = \sqrt{t}$, we get the first Euler integral, $E(0)$:

$$\sqrt{\pi} = 2 \int_0^{+\infty} e^{-x^2} dx = 2 \cdot \frac{1}{2} \int_0^{+\infty} t^{-\frac{1}{2}} e^{-t} dt = \Gamma(1/2) \tag{2.90}$$

which proves the Eq. (2.79) of the previous Section.

The integrals (2.87) for odd n are equal to 0 on the interval $[-\infty, +\infty]$, since being the product between an even and an odd function.

The integrals for even n are obtained by differentiation with respect to the parameter α . By taking the first derivative of both sides of Eq. (2.87) we get:

$$\int_0^{+\infty} x^2 e^{-\alpha x^2} dx = \frac{\pi^{1/2}}{2\alpha^{3/2}} \tag{2.91}$$

By sequentially taking higher order derivatives, the following general result is obtained:

$$\int_0^{+\infty} x^n e^{-\alpha x^2} dx = \frac{1 \cdot 3 \cdot 5 \cdots (n+1) \pi^{1/2}}{2^{n/2} \alpha^{(n+1)/2}} = \frac{\Gamma[(n+1)/2]}{2\alpha^{(n+1)/2}} \tag{2.92}$$

The Stirling Approximation

When deriving Eq. (2.11) above, and several other times, in Chap. 3 and in other places in this book, we need to compute the factorial $n!$ of very large values of n , as well as its logarithm. Especially in the latter case, a very useful formula is the Stirling's approximation:

$$\ln n! = n \ln n - n + \frac{1}{2} \ln(2\pi n) + \mathcal{O}(1/n) \tag{2.93}$$

This formula is very accurate already for small n . For example, the relative error for $n = 20$ is less than 10^{-4} .

Stirling's formula can be applied also to the Gamma function, provided its argument t is real:

$$\ln \Gamma(t) \simeq (t - \frac{1}{2}) \ln t - t + \frac{1}{2} \ln 2\pi \quad (2.94)$$

We also note the important property, with α a real constant:

$$\ln(\alpha n)! = \ln(\alpha^n n!) = n \ln \alpha + \ln n! \quad (2.95)$$

Vector Calculus and Analysis

A scalar is a quantity characterised only by its magnitude (a number). Specification of a vector \mathbf{v} , instead, requires stating its direction as well as its magnitude $|\mathbf{v}| = v$. Unit vectors are vectors of unit length, while the zero vector has zero length and arbitrary direction. The unit vectors of the Cartesian coordinate system are written as \mathbf{i} , \mathbf{j} and \mathbf{k} , respectively along the axes $\{x, y, z\}$. Any other vector in the 3D space can then be expressed by giving its scalar components, $\{v_x, v_y, v_z\}$, as $\mathbf{v} = v_x \mathbf{i} + v_y \mathbf{j} + v_z \mathbf{k}$. The magnitude (or modulus) of the vector is: $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$.

Vectors \mathbf{u} and \mathbf{v} can be added, as $\mathbf{w} = \mathbf{u} + \mathbf{v}$, by adding their components: $\mathbf{w} = (u_x + v_x)\mathbf{i} + (u_y + v_y)\mathbf{j} + (u_z + v_z)\mathbf{k}$. A vector can also be multiplied by a scalar s , by multiplying its components: $s\mathbf{v} = sv_x\mathbf{i} + sv_y\mathbf{j} + sv_z\mathbf{k}$.

The *scalar product* (or dot product) of two vectors is defined as:

$$\mathbf{u} \cdot \mathbf{v} = u_x v_x + u_y v_y + u_z v_z = uv \cos \theta \quad (2.96)$$

with u, v the modulus of the vectors, and θ the angle comprised between the directions of two vectors, joined at a common origin. The result of the scalar product of two vectors is a scalar (a number).

The *vector product* (or cross product) of two vectors is defined as:

$$\mathbf{u} \times \mathbf{v} = (u_y v_z - u_z v_y)\mathbf{i} + (u_z v_x - u_x v_z)\mathbf{j} + (u_x v_y - u_y v_x)\mathbf{k} = \mathbf{w} \quad (2.97)$$

and its result is another vector \mathbf{w} , perpendicular to the plane containing \mathbf{u} and \mathbf{v} .

Consider a scalar function of the coordinates, $\psi(x, y, z)$, such as a temperature or a pressure distributed in the volume of a body. The **gradient** of ψ is a vector defined as:

$$\text{grad } \psi = \nabla \psi = \frac{\partial \psi}{\partial x} \mathbf{i} + \frac{\partial \psi}{\partial y} \mathbf{j} + \frac{\partial \psi}{\partial z} \mathbf{k} \quad (2.98)$$

It should be evident that the gradient of the scalar function (or *field*) is pointing to the direction where ψ changes more rapidly.

The operation producing the gradient can be thought of coming from the application of an operator ∇ on the scalar ψ :

$$\nabla = \mathbf{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\partial}{\partial y} + \mathbf{k} \frac{\partial}{\partial z} \quad (2.99)$$

If we now apply the same operator ∇ but to a vector \mathbf{v} , we obtain the **divergence** of that vector:

$$\text{div } \mathbf{v} = \nabla \cdot \mathbf{v} = \frac{\partial \mathbf{v}}{\partial x} \cdot \mathbf{i} + \frac{\partial \mathbf{v}}{\partial y} \cdot \mathbf{j} + \frac{\partial \mathbf{v}}{\partial z} \cdot \mathbf{k} \quad (2.100)$$

The physical meaning of the divergence is to calculate the net amount of the vector $\mathbf{v}(x, y, z)$ (actually a vector field) flowing in or out a closed surface. Imagine a vector field running parallel to the x -axis, such as water flowing at constant speed: the divergence calculates the integral of the flux across a surface perpendicular to x .

Finally, the vector product of ∇ times a vector \mathbf{v} , gives the **rotor**, or “curl” of the vector:

$$\text{curl } \mathbf{v} = \nabla \times \mathbf{v} = \left(\frac{\partial v_y}{\partial z} - \frac{\partial v_z}{\partial y} \right) \mathbf{i} + \left(\frac{\partial v_z}{\partial x} - \frac{\partial v_x}{\partial z} \right) \mathbf{j} + \left(\frac{\partial v_x}{\partial y} - \frac{\partial v_y}{\partial x} \right) \mathbf{k} \quad (2.101)$$

The rotor of a vector field is non-zero only if the field turns around some point, like in a vortex. For a vector field forming a vortex spinning circularly around a vertical line, the rotor calculates the value of $\mathbf{v}(x, y, z)$ along the perimeter of any circle drawn about the central line.

The square of the operator ∇^2 is called the **Laplacian**. Applied to a scalar, it gives another scalar:

$$\nabla^2 \psi = \frac{\partial^2 \psi}{\partial^2 x} + \frac{\partial^2 \psi}{\partial^2 y} + \frac{\partial^2 \psi}{\partial^2 z} \quad (2.102)$$

The Laplacian of a vector is a vector:

$$\begin{aligned} \nabla^2 \mathbf{v} &= \left(\frac{\partial^2}{\partial^2 x} + \frac{\partial^2}{\partial^2 y} + \frac{\partial^2}{\partial^2 z} \right) (v_x \mathbf{i} + v_y \mathbf{j} + v_z \mathbf{k}) = \\ &= \left(\frac{\partial^2 v_x}{\partial^2 x} + \frac{\partial^2 v_x}{\partial^2 y} + \frac{\partial^2 v_x}{\partial^2 z} \right) \mathbf{i} + \left(\frac{\partial^2 v_y}{\partial^2 x} + \frac{\partial^2 v_y}{\partial^2 y} + \frac{\partial^2 v_y}{\partial^2 z} \right) \mathbf{j} + \left(\frac{\partial^2 v_z}{\partial^2 x} + \frac{\partial^2 v_z}{\partial^2 y} + \frac{\partial^2 v_z}{\partial^2 z} \right) \mathbf{k} \end{aligned} \quad (2.103)$$

Some useful formulae of vector analysis are:

$$\nabla \cdot (\mathbf{u} \times \mathbf{v}) = \mathbf{v} \cdot \nabla \times \mathbf{u} - \mathbf{u} \cdot \nabla \times \mathbf{v} \quad (2.104)$$

$$\nabla (\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot \nabla \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{u} + \mathbf{u} \times (\nabla \times \mathbf{v}) + \mathbf{v} \times (\nabla \times \mathbf{u}) \quad (2.105)$$

$$\nabla \times (\nabla \psi) = 0 = \text{curl } (\text{grad } \psi) \quad (2.106)$$

$$\nabla \cdot (\nabla \times \mathbf{u}) = 0 = \text{div} (\text{curl } \mathbf{u}) \quad (2.107)$$

$$\nabla \cdot (\nabla \psi_1 \times \nabla \psi_2) = 0 \quad (2.108)$$

$$\nabla \times (\nabla \times \mathbf{u}) = \text{curl} (\text{curl } \mathbf{u}) = \text{grad} (\text{div } \mathbf{u}) - \nabla^2 \mathbf{u} \quad (2.109)$$

Note that a vector field with zero divergence is said to be *solenoidal* (a field with no point source, such as a magnetic field). A vector field with zero curl is said to be *irrotational* (such as a tube of water flowing in laminar flux without any turbulence). A scalar field with zero gradient is said to be *constant* (such as a temperature uniform everywhere in a body).

Simple Tensor Algebra

If we multiply a vector by a scalar, $\mathbf{u}' = a\mathbf{u}$, the vector changes in magnitude (each of the components are multiplied by a) but not in direction. On the other hand, as shown in the previous Section, by multiplying two vectors we get either a scalar (dot product) or another vector perpendicular to the first two (cross product). But how do we get to change both the direction and magnitude of a vector into an arbitrary direction and different magnitude? This is obtained by introducing a more complex entity, the **tensor**, indicated by an underline:

$$\mathbf{u} = \underline{\mathbf{M}} \otimes \mathbf{v} \quad (2.110)$$

By writing the vector components explicitly, with $\mathbf{u} = u_x \mathbf{i} + u_y \mathbf{j} + u_z \mathbf{k}$ and $\mathbf{v} = v_x \mathbf{i} + v_y \mathbf{j} + v_z \mathbf{k}$, we see that the new mathematical object $\underline{\mathbf{M}}$ can be obtained by multiplying two original vectors component by component:

$$\begin{aligned} \underline{\mathbf{M}} &= (u_x \mathbf{i} + u_y \mathbf{j} + u_z \mathbf{k}) \otimes (v_x \mathbf{i} + v_y \mathbf{j} + v_z \mathbf{k}) = \\ &= u_x v_x \mathbf{ii} + u_x v_y \mathbf{ij} + u_x v_z \mathbf{ik} + u_y v_x \mathbf{ji} + \dots \end{aligned} \quad (2.111)$$

The new symbol \otimes represents this idea of distributing the product among all the vector components. By rewriting the 3×3 scalar quantities as M_{ij} , with $i, j = 1, 2, 3$:

$$\underline{\mathbf{M}} = M_{11} \mathbf{ii} + M_{12} \mathbf{ij} + M_{13} \mathbf{ik} + M_{21} \mathbf{ji} + \dots \quad (2.112)$$

this new mathematical entity appears as a square matrix, also called a tensor of rank 2, which can be multiplied by a vector or by another tensor, by using the usual rules of matrix algebra:

$$\underline{\mathbf{M}} = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix} \quad (2.113)$$

$$\mathbf{u} = \underline{\mathbf{M}} \otimes \mathbf{v} = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix} \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix} \quad (2.114)$$

$$\underline{\mathbf{M}} \otimes \underline{\mathbf{N}} = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix} \begin{pmatrix} N_{11} & N_{12} & N_{13} \\ N_{21} & N_{22} & N_{23} \\ N_{31} & N_{32} & N_{33} \end{pmatrix} = \sum_{k=1}^3 M_{ik} N_{kj} \quad (2.115)$$

In the following (see e.g. Chaps. 8–10) we will meet with tensors of rank 2, such as the stress and the strain of a deformed material element, and tensors of rank 4, such as the matrices of the elastic constants and elastic compliances of a material. However, in this book we will be exclusively concerned with quantities defined in orthogonal reference frames (i.e., the angles between the x , y , z axes are always at 90°), which simplifies much the tensor analysis, getting rid of (usually important) details such as covariant and contravariant components.

By extension, it might be tempting to deduce that a vector is nothing else but a tensor of rank 1, a mathematical entity with just one subscript index. However, care must be exercised. Although a seemingly mathematical trickery, tensor calculus is instead at the core of the concepts of invariance in physics. It turns out that all rank-1 tensors are also vectors, however *not* all vectors are rank-1 tensors: to be a tensor of any rank, a mathematical entity must be *invariant* with respect to the coordinate system. For example, if we consider the position vector \overrightarrow{OP} joining a point P with the origin O of a $\{x, y, z\}$ reference frame, and the position vector $\overrightarrow{O'P}$ joining the same point with the origin O' of a different reference frame $\{x', y', z'\}$, it is immediately seen that this vector depends on the reference frame, therefore it is not a rank-1 tensor. However, the *difference* between two position vectors $\overrightarrow{PQ} = \overrightarrow{OP} - \overrightarrow{OQ}$ does not depend on the reference frame, therefore the vector of the distance between two points is a rank-1 tensor.

Similarly, a tensor can always be written as a matrix, but a matrix is not necessarily a tensor. By multiplying a column vector (u_1, u_2, u_3) by a matrix \mathbf{M} , a new column of coefficients (q_1, q_2, q_3) is obtained: if these numbers are the components of another vector, then the matrix is a tensor, $\mathbf{M} = \underline{\mathbf{M}}$. That the resulting (q_1, q_2, q_3) is a vector can be easily checked: just change the basis (reference frame) to transform the first vector (u_1, u_2, u_3) into another vector (u'_1, u'_2, u'_3) ; then apply the same change of basis to (q'_1, q'_2, q'_3) ; if $\mathbf{q}' = \mathbf{M}\mathbf{u}'$, then \mathbf{q} is a vector, and \mathbf{M} is a tensor.

As it was already said, the tensor notation allows to expose underlying symmetries and invariances of the corresponding physical quantities. The *tensor invariants* are quantities derived from the tensor that do not change upon changing or rotating the

reference frame. They are defined as the coefficients of the characteristic polynomial of the tensor $\underline{\mathbf{M}}$:

$$p(\lambda) = \text{Det}[\underline{\mathbf{M}} - \lambda \underline{\mathbf{I}}] \quad (2.116)$$

where $\underline{\mathbf{I}}$ is the identity tensor (with 1 on the diagonal and 0 everywhere) and λ is an indeterminate quantity. For a 3×3 rank-2 tensor, the most commonly encountered in this book, there are only three invariants:

$$M_I = \text{Tr}\{\underline{\mathbf{M}}\} = M_{11} + M_{22} + M_{33} \quad (2.117)$$

$$\begin{aligned} M_{II} &= \frac{1}{2} [(\text{Tr}\{\underline{\mathbf{M}}\})^2 - \text{Tr}\{\underline{\mathbf{M}}^2\}] = \\ &= M_{11}M_{22} + M_{22}M_{33} + M_{33}M_{11} - M_{12}M_{21} - M_{23}M_{32} - M_{13}M_{31} \end{aligned} \quad (2.118)$$

$$M_{III} = \text{Det}\{\underline{\mathbf{M}}\} \quad (2.119)$$

The first one is called also the *trace* of the tensor (sum of the diagonal components of the corresponding matrix); the third one coincides with the determinant of the matrix; the second one has no obvious interpretation. Note that, since a tensor is written as a matrix, its eigenvalues can also be calculated; they would be M_1, M_2, M_3 for the 3×3 tensors. However, the number of invariants for a $n \times n$ tensor is just n , therefore the three invariants previously defined must be expressed in terms of the three eigenvectors, and vice versa:

$$M_I = M_1 + M_2 + M_3 \quad (2.120)$$

$$M_{II} = M_1M_2 + M_2M_3 + M_1M_3 \quad (2.121)$$

$$M_{III} = M_1M_2M_3 \quad (2.122)$$

The importance of invariants becomes evident when considering *objective functions* (i.e., not depending on the change of coordinates) of the tensor. Such objective functions depend only on the n invariants of the tensor, instead of its components. For example, when calculating the elastic potential energy as function of the strain tensor (see Appendix H to Chap. 9), this reduces to a function of three parameters rather than six (the strain tensor $\underline{\varepsilon}$ has $3 \times 3 = 9$ components, however by symmetry $\varepsilon_{ij} = \varepsilon_{ji}$, therefore its independent components are just six). Moreover, within the framework of linear elasticity, the energy must be a quadratic function of the strain, which eliminates an additional scalar. This is the reason why, for an isotropic material, only two independent parameters are needed to describe the elastic properties, known as *Lamé coefficients*.

Simple Fourier Analysis

The Fourier transform decomposes a function of a variable x into the normal components of a conjugate variable y . Physically useful pairs of conjugate variables are, e.g., time and frequency, or position and momentum. For example, a time signal can be decomposed into the frequencies that make it up; or, a movement in space can be decomposed into the wavevectors that correspond to elementary oscillation modes.

The basic rule to obtain the **Fourier transform** of a function $f(x)$ in the y -space is:

$$g(y) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{+\infty} f(x) e^{-ixy} dx \quad (2.123)$$

The function $g(y)$ can be anti-transformed, to obtain back the $f(y)$:

$$f(x) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{+\infty} g(y) e^{+iyx} dy \quad (2.124)$$

The Fourier transform is closely connected with the **Fourier series**, both named after the French mathematician J. J. Fourier (1768-1830). Any function $f(x)$ can be expressed by a series of Fourier components, as:

$$f(x) = \frac{a_0}{(2\pi)^{1/2}} + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx) \quad (2.125)$$

with the coefficients:

$$a_n = \frac{1}{L} \int_{-L}^{+L} f(x) \cos\left(\frac{n\pi x}{L}\right) dx \quad ; \quad b_n = \frac{1}{L} \int_{-L}^{+L} f(x) \sin\left(\frac{n\pi x}{L}\right) dx \quad (2.126)$$

With a little algebra, it can be shown that upon substituting these coefficients in the series development for $f(x)$, one obtains exactly the definition (2.123) of the Fourier transform.

Some useful FT of elementary functions and operators (pairs (x, y) = generic; (t, ω) = time-frequency, or pulsation; (\mathbf{x}, \mathbf{k}) = position – momentum, or wavevector):

$$f(\omega) = \frac{1}{(2\pi)^{1/2}} \int_0^{+\infty} e^{-at} e^{-i\omega t} dt = \frac{2a}{(a^2 + 4\pi^2\omega^2)} \quad (a > 0) \quad (2.127)$$

$$f(\omega) = \frac{1}{(2\pi)^{1/2}} \int_0^{+\infty} e^{-at^2} e^{-i\omega t} dt = e^{-a\omega^2} \quad (2.128)$$

$$f(\omega) = \frac{1}{(2\pi)^{1/2}} \int_0^{+\infty} e^{-at} u(t) e^{-i\omega t} dt = \frac{1}{(a + i\omega)} \quad (a > 0) \quad (2.129)$$

$$f(\omega) = \frac{1}{(2\pi)^{1/2}} \int_0^{+\infty} e^{i\omega_0 t} e^{-i\omega t} dt = 2\pi \delta(\omega - \omega_0) \quad (2.130)$$

$$f(y) = \frac{1}{(2\pi)^{1/2}} \int_0^{+\infty} \sin^2(x) e^{-ixy} dx = \frac{1}{4} [2\delta(y) - \delta(y - \frac{1}{\pi}) - \delta(y + \frac{1}{\pi})] \quad (2.131)$$

$$f(y) = \frac{1}{(2\pi)^{1/2}} \int_0^{+\infty} \cos^2(x) e^{-ixy} dx = \frac{1}{4} [2\delta(y) + \delta(y - \frac{1}{\pi}) + \delta(y + \frac{1}{\pi})] \quad (2.132)$$

$$f(k) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{+\infty} [\nabla g(x)] e^{-ik \cdot x} d\mathbf{x} = ik g(k) \quad (2.133)$$

$$f(k) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{+\infty} [\nabla^2 g(x)] e^{-ik \cdot x} d\mathbf{x} = -k^2 g(k) \quad (2.134)$$

Problems

2.1 Basic nomenclature

Identify which of the following systems are either isolated, closed or open systems.

- (a) a car tyre; (b) the Milky Way; (c) a brain cell; (d) a refrigerator; (e) a hammer; (f) a frog.

2.2 Formal identities

(a) Show that, for a perfect gas, the pressure is $P = \frac{2}{3} \frac{E}{V}$.

(b) Show that the Helmholtz free energy, $F = U - TS$, is also equal to $F = -pV + \mu N$.

(c) If the Gibbs free energy is $G = F + PV$, state the thermodynamic independent variables on which G depends explicitly.

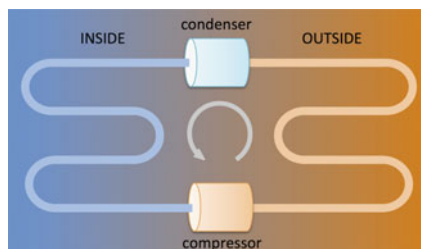
2.3 Thermal engine

An air conditioner is made of a sealed piping circuit, in which a coolant fluid flows. The piping goes through electrically-powered cooling stages: a condenser, which turns the hot liquid into a cold gas by expansion, and a compressor, which takes the room-temperature fluid and turns it to a high pressure, high temperature gas. A simplified scheme is represented in the figure below.

(a) Is this system closed or open? Can part of the system be identified as a closed or open subsystem?

(b) Describe the exchanges of energy, heat and entropy in the system, according to the thermal engine model.

(c) Could the system work without the condenser, by just disposing of the air from room temperature, to the exterior?



2.4 Exchanges of entropy

You add 50 L of hot water at 55 °C, with 25 L of cold water at 10 °C. What is the equilibrium temperature? How much entropy is produced by the time equilibrium occurs? Can you attribute part of the entropy to the hot water and part to the cold water?

2.5 Boiling, temperature and pressure

Boiling is the process by which a heated fluid turns into vapour. The *boiling point* is the temperature at which the pressure exerted by the evaporating liquid is equal to that of the surroundings. An open pot of water at the sea level, where the ambient pressure is $P_0 = 101.32$ kPa, boils by definition at the normal temperature of $T_0 = 100$ °C. The Clausius-Clapeyron equation:

$$T_B = \left(\frac{1}{T_0} - \frac{R \ln \left(\frac{P}{P_0} \right)}{\Delta H_{vap}} \right)^{-1}$$

relates the boiling point at a different pressure P , with ΔH_{vap} the heat of vaporisation of the liquid, equal to 40.65 kJ/mol for water (note that this is more than five times the energy required to heat the same quantity of water from 0 to 100 °C). Let us take a pressure cooker of volume $V = 6$ l, and fill it by half with water at $T = 23$ °C. At what temperature will the water boil?

2.6 Stefan-Boltzmann T^4 law

After Josef Stefan presented in 1879 his experimental T^4 law for the radiative heat transfer from a surface, Ludwig Boltzmann set out to give an explanation based on classical thermodynamics. He considered radiation “particles” to behave as a classical fluid, with energy density e . The idea that radiation could exert a pressure was quite new at that time, but it was a logical outcome of Maxwell’s equations. The pressure of this radiation fluid inside the familiar piston-cylinder ideal experiment (see Fig. 2.1a) would have been $p = e/3$. By writing the internal energy as $E = eV$, use the fundamental thermodynamic relation $dU = TdS - pdV$ to reobtain the T^4 law, as Boltzmann found out.

2.7 A negative temperature

Consider an ideal gas of N particles, each of which can exist in a “ground state” with energy $\varepsilon_i = 0$, or in an “excited state” with energy $\varepsilon_i = +e$. By taking that there

are no interaction in the ideal gas, can you show that in the thermodynamic limit ($N \rightarrow \infty$ and $V \rightarrow \infty$, with $N/V = \text{const}$) this system has a region of negative temperatures, as a function of the population of particles in the excited state? Does such a result make sense?

2.8 Greenhouse gases 1

Assuming the atmosphere is at equilibrium, the chemical potential of each gas is constant at any altitude h . This observation allows to predict the concentration of each greenhouse gas as a function of the altitude. The chemical potential of a molecular species is defined:

$$\mu = -T \left(\frac{\partial S}{\partial N} \right)_{E,V}$$

that is the derivative of the entropy with respect to N , at constant E and V . You know already the expression of entropy as a function of (N, V, E) , it is the Eq. (2.11), or Sackur-Tetrode formula. Use it to derive an expression for $\mu(h)$. Then by imposing that μ is constant at any h , obtain the number of molecules N as a function of h and of the mass of the molecule.

2.9 Greenhouse gases 2

The Earth's atmosphere is made up by 78 % of N_2 , 21 % of O_2 and 1 % of Ar_2 , not counting the fraction of water vapour which, according to the height and temperature, can vary from 1 to 4 %. Diatomic homonuclear molecules, N_2 , O_2 or Ar_2 , do not contribute to the greenhouse effect since their radiation absorbing power is zero. The “greenhouse effect” comes from the other gaseous species found in the atmosphere, in smaller concentration. To a first approximation, the efficacy of each gas in the greenhouse effect depends on: (1) its concentration, and (2) its capability of absorbing the electromagnetic radiation. In turn, the *global-warming potential* (GWP) of each gas depends on its efficacy and of its lifetime in the atmosphere. For the five main greenhouse gases (CO_2 ; methane; N_2O ; ozone; water vapour), the table below gives typical concentration and lifetime; their radiation absorption capability can be deduced from the following figure.

- Which of the five gases is mostly effective in absorption? Explain why.
- Rank in order of importance the five greenhouse gases, by taking into account the combination of the three parameters (concentration, lifetime, absorption).
- Why the gas apparently the most important in this ranking seems to be, conversely, the least important in global warming? (Fig. 2.7)

Gas	Concentration (ppm)	Lifetime (years)
CO_2 (carbon dioxide)	400	30–90
CH_4 (methane)	2	12
N_2O (nitrogen protoxyde)	0.3	115
O_3 (ozone)	0.03	0.05
H_2O (water vapour)	(°)	0.02–0.06

(°) water vapour makes about 1–4 % of the atmosphere

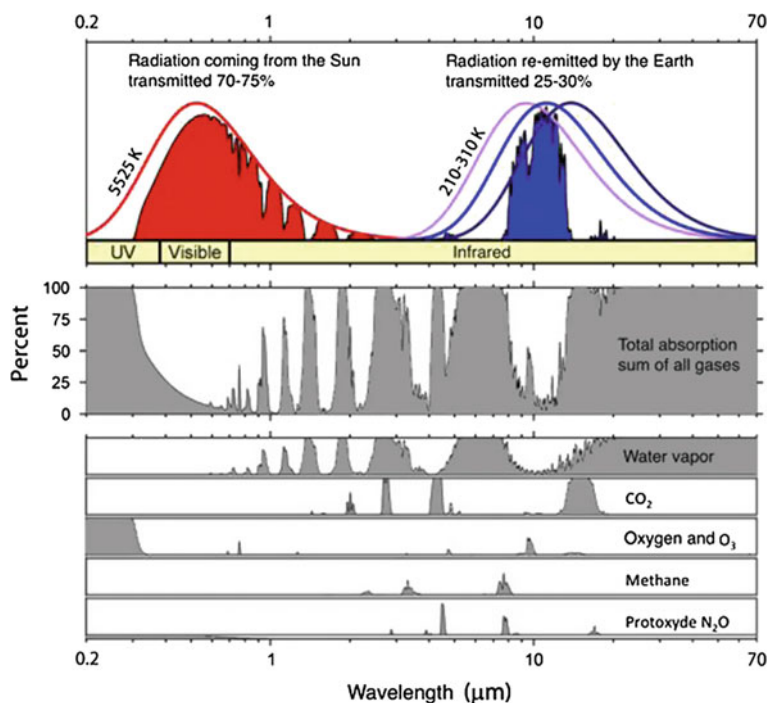


Fig. 2.7 Top row spectra of the electromagnetic radiation emitted by the Sun (red) and by the Earth (blue). Middle row: total absorption fraction by wavelength, for all the greenhouse gases. Bottom row absorption fraction by wavelength for the five main greenhouse gases. [Image © by R. A. Rohde, Global Warming Art project, repr. under CC-BY-SA 3.0 licence, see (*) for terms.]

(*) The terms of the Creative Commons Attribution-ShareAlike 3.0 and 4.0 International License (<http://creativecommons.org/licenses/by-sa/3.0/>, <http://creativecommons.org/licenses/by-sa/4.0/>) permit use, duplication, adaptation, distribution, and reproduction in any medium or format, as long as appropriate credit is given to the original author(s) and the source, providing a link to the Creative Commons license and indicating if changes were made. If remixing, transforming, or building upon this chapter or a part thereof, such contributions must be distributed under the same license as the original.

References

1. S. Weinberg, *Dreams of a final theory: the scientist's search for the ultimate laws of nature* (Vintage, Random House, New York, 1992)
2. R. Newburgh, H.S. Leff, The Mayer-Joule principle: the foundation of the first law of thermodynamics. *Phys. Teach.* **49**, 484–487 (2011)
3. H.J. Steffens, *James Prescott Joule and The Concept of Energy* (Science History Publications, New York, 1979)
4. R.H. Swendsen, Statistical mechanics of classical systems with distinguishable particles. *J. Stat. Phys.* **107**, 1143–1166 (2002)
5. J.F. Nagle, Regarding the entropy of distinguishable particles. *J. Stat. Phys.* **117**, 1047–1062 (2004)
6. O. Sackur, Die Anwendung der kinetischen Theorie der Gase auf chemische Probleme. *Ann. Physik* **36**, 958–980 (1911)
7. H. Tetrode, Die chemische Konstante der Gase und das elementare Wirkungsquantum. *Ann. Physik* **38**, 434–442 (1912)
8. S. Carnot, *Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance* (Bachelier, Paris, 1824)
9. R. Clausius, *The Mechanical Theory of Heat* (transl. by W.R. Browne of nine Clausius' papers from German) (MacMillan and Co., London, 1879)
10. C. Bustamante, Unfolding single RNA molecules: bridging the gap between equilibrium and non-equilibrium statistical thermodynamics. *Quart. Rev. Biophys.* **38**, 291–301 (2006)
11. L. Boltzmann, Über das Wärmegleichgewicht zwischen mehratomigen Gasmolekülen. *Wiener Berichte* **63**, 397–418 (1871)
12. J. Perrin, *Le mouvement Brownien et la réalité moléculaire*. *Ann. Chimie Phys.* **18** (8me ser.), 5–114 (1909)
13. J.W. Gibbs, On the equilibrium of heterogeneous substances. *Trans. Connecticut Acad. Arts Sci.* III, 198–248 and 343–520 (1874–1878) [The voluminous (300 pages) Gibbs' paper was published in several parts in the Connecticut Academy bulletin, and remained largely unknown until it was translated in German and French, between 1891 and 1899.]
14. A. Kleidon, A basic introduction to the thermodynamics of the Earth system far from equilibrium and maximum entropy production. *Philos. Trans. Roy. Soc. (London) B: Biol. Sci.* **365**, 1303–1315 (2010)
15. S.G. Brittin, G.A. Gamow, Negative entropy and photosynthesis. *Proc. Nat. Acad. Sci. USA* **47**, 724–728 (1961)

Further Reading

16. S.B. Braun, J.P. Ronzheimer, M. Schreiber, S.S. Hodgman, T. Rom, I. Bloch, U. Schneider, Negative absolute temperature for motional degrees of freedom. *Science* **339**, 52–55 (2013)
17. E. Schrodinger, *What is life?* (Cambridge University Press, Cambridge, 1944)
18. W. Yourgrau, A. Van der Merwe, Entropy balance in photosynthesis. *Proc. Nat. Acad. Sci. USA* **59**, 734–737 (1958)
19. E. Fermi, *Thermodynamics* (Dover Books on Physics, Reprint of the 1937 edn.) (Dover, New York, 1965)
20. D. Kondepudi, I. Prigogine, *Modern Thermodynamics: From Heat Engines to Dissipative Structures* (John Wiley, New York, 1998)
21. D. Schroeder, *An Introduction to Thermal Physics* (Pearson, Boston, 1999)
22. F. Frenkel, B. Smit, *Understanding Molecular Simulation. From Algorithms to Applications*, 2nd edn. (Academic Press, New York, 2002)

23. S.E. Jorgensen, YuM Svirezhev, *Towards a Thermodynamic Theory for Ecological Systems* (Pergamon Press, New York, 2004)
24. J.D. Neelin, *Climate Change and Climate Modeling* (Cambridge University Press, 2011)

<http://www.springer.com/978-3-319-30645-2>

The Physics of Living Systems

Cleri, F.

2016, XXIV, 620 p. 204 illus., 173 illus. in color.,

Hardcover

ISBN: 978-3-319-30645-2