1 Thermodynamics: a brief overview

Thermodynamics is a phenomenological theory based on empirical observations. The thermodynamic laws are a coherent account of the experimental analysis and provide a universally valid description of the behavior of macroscopic systems without referring to their detailed structure. In this chapter we introduce thermodynamics as a prerequisite scenario for statistical mechanics and we do not give an extended exposition of the theory, but only recall the basic concepts as an introduction to the statistical approach.¹

1.1 Equilibrium states and the empirical temperature

A thermodynamic system is any portion of matter, solid, liquid, mixtures, . . . , which can be described in terms of a small number of parameters, the macroscopic thermodynamic variables, as for example the pressure \( P \) and the volume \( V \) for a fluid or the magnetic moment and the magnetic field for a paramagnet. These variables are called extensive or intensive, depending whether their value does or does not depend on the amount of matter present in the system. Energy, volume and magnetization are extensive variables, whereas pressure, temperature and magnetic field are intensive.

Thermodynamics deals with macroscopic properties of macro-systems at equilibrium. A thermodynamic equilibrium state must then be characterized as a macroscopic phenomenon defined via macroscopic variables. A system is in equilibrium if the values of the observables we take into consideration and of the parameters describing its state do not change with time. The thermodynamic equilibrium is never fully static. The characterization of a thermodynamic state depends on the observation time, usually at least of the order of fractions of a second, which in any case must always be much longer than any characteristic time scale of the molecular motion (\( \sim 10^{-11} \text{ s} \)).

We use a simple example of common experience to illustrate this point. Soon after hot water is poured into a container, it appears to be at rest. If we ask whether the water is in a state of equilibrium, the answer depends obviously on the observation time and on the sensitivity of the measurement. Within the interval of time of a few seconds we do not feel changes in the temperature and in the volume of the water, which during this observation time can be considered at equilibrium by itself. If we prolong the observation, we start to appreciate that the temperature of the water is changing and the water cannot be considered

¹ For a more detailed exposition we suggest the book by A. B. Pippard (1957).
at equilibrium. After several minutes the temperature of the water will be equal to the room temperature and will not change anymore. Even if the molecules of the water continually evaporate, within a few hours also the volume of the water does not change appreciably. The water at the same temperature of the room in this interval of time can be considered in equilibrium. In a few days, the water will eventually evaporate and the equilibrium state of the water in the container is meaningless. Unlike mechanical equilibrium, thermodynamic equilibrium is never static. A thermodynamic state does not refer, as in mechanics, to instantaneous positions and velocities of all the microscopic constituents, the molecules of the water in the previous example, which define a micro-state or configuration. All the observation times considered above are macroscopic time scales much longer than the microscopic time characteristic of molecular motion. Provided the observation time considered remains within certain limits, the observables that we consider to characterize the thermodynamic equilibrium do not change their values appreciably. During this time the system goes through a large number of microscopic configurations and the proper macroscopic observables of thermodynamic equilibrium can be considered as averages in the complicated evolution of the motion of all molecules during this interval of time. A state of absolute thermodynamic equilibrium from which the system never departs cannot be considered, strictly speaking, as an experimental truth. Its occurrence, as we have seen, depends on the observation time and on the precision that we require or that we have in the experimental apparatus. In reality we can assume that for a given small set of external intensive parameters, such as temperature and pressure when the system is conditioned by the external world, or the volume and the energy for a system in complete isolation, a macroscopic system has definite equilibrium properties. With the caution discussed above, macroscopic violations of this working hypothesis, as we shall see later, would require observation times so long as to be considered infinite for any practical purpose.

If two systems are in equilibrium, their states and, therefore, their thermodynamic variables cannot be specified arbitrarily. Moreover, we can postulate the transitive property of equilibrium, zeroth law, that if two systems A and B are in equilibrium with a reference system C, then A is in equilibrium with B. For single component systems, as for instance fluids, whose states are specified by $P$ and $V$, their states must be conditioned in the sense that if two systems A and B are in equilibrium with another system C, specified by the two conditions

$$f_{AC}(P_1, V_1, P_3, V_3) = 0, \quad f_{BC}(P_2, V_2, P_3, V_3) = 0,$$

then A and B are in equilibrium with one another:

$$f_{AB}(P_1, V_1, P_2, V_2) = 0. \quad (1.2)$$

Here $(P_1, V_1), (P_2, V_2)$ and $(P_3, V_3)$ refer to systems A, B and C, respectively. The functions $f_{AC}, f_{BC}$ and $f_{AB}$ depend on the nature of the systems A, B and C.

---

2 It can be experimentally verified that, independently from the process by which the pressure and the volume of the gas are fixed, its properties (viscosity, thermal conductivity, . . . ) are the same.
The above conditions allow us to introduce the notion of the empirical temperature. By solving the two equations in (1.1) with respect to $P_3$ there must be a function such that

$$\theta_A(P_1, V_1, V_3) = \theta_B(P_2, V_2, V_3).$$  \hspace{1cm} (1.3)

Since the relation (1.3) must have the same content of the thermal equilibrium condition (1.2) between the two systems A and B, the functional dependence on $V_3$ of $\theta_A$ and $\theta_B$ must be the same and can then be dropped by fixing the volume $V_3$ of the system C assumed as a reference system. The resulting function $\theta(P, V)$ defines then, the empirical temperature and its form depends on the chosen reference system. This argument can be generalized to any macrosystem and the empirical temperature $\theta$ will depend on all the parameters necessary to characterize the state of the system.

Once the above definition is assumed, the equilibrium implies a condition among the variables specifying a thermodynamic state and the expression of any one of them in terms of the other variables defines the equation of state. For a perfect gas the equation of state is given by the empirical observation for a very dilute gas, which can be summarized by Boyle’s law, according to which the product $PV = f(\theta)$. Using a mercury-in-glass thermometer, the function $f(\theta)$ is linear over a wide range of temperatures. The coefficient of proportionality, per mole of the gas, is identified with the gas constant $R$ once the unity of the scale coincides with the degree Celsius equal to the Kelvin (per one mole $R = 8.3145$ J/molK). The above temperature, for which from now on we use the standard notation $T$, also coincides with the absolute temperature $T$ to be defined later according to the second law.

The condition of mechanical and thermal equilibrium requires that the system pressure, $P$, and temperature, $T$, coincide with those of the environment. In the case of a multicomponent system, in addition to the standard thermodynamic variables, the specification of the relative concentration of each component is also required. Furthermore a system may or may not be homogeneous. If a system is homogeneous it is said to consist of a single phase. Hence the state of a system is given, in general, by specifying, besides $P$ and $T$, the number of phases $f$ and components $n$. However, not all of these are independent parameters. The actual number of degrees of freedom of a thermodynamic system, i.e., the number $v$ of independent parameters, is given by the Gibbs rule (which will be proved later on)

$$v = 2 + n - f.$$  \hspace{1cm} (1.4)

For instance, an ideal gas is a single-component single-phase system, and its state is entirely determined by two variables, pressure and temperature.

Depending on the value of the thermodynamic parameters, a single-component system may be in different phases. The information about the possible phases of a system is graphically represented in a phase diagram, which is a Cartesian diagram whose coordinates correspond to any two of the thermodynamic parameters $P$, $T$ and $V$. A generic thermodynamic state corresponds to a point in the phase diagram. According to Gibbs’ rule, the coexistence of two phases, for a single-component system, implies a number of degrees of freedom, $v = 1$, and hence different phases may coexist only along curves in the phase
Phase diagram of a one-component system (water) in the $T$–$P$ plane. The critical point (CP) is the end point of the liquid–vapor coexistence curve $T_c = 647.096 \text{ K}, P_c = 27.064 \text{ MPa}$. The triple point (TP) is the coexistence of three phases $T_{tr} = 273.16 \text{ K}, P_{tr} = 6.11 \times 10^2 \text{ Pa}$.

Within thermodynamics we consider also processes during which the state of a system changes. If a system is in equilibrium, its macroscopic state cannot change unless the external constraints are modified. If the external constraints are suddenly modified, the system will find itself, in general, in a non-equilibrium state. After some time, the system will reach a new equilibrium state corresponding to the modified constraints. A process from a non-equilibrium state to an equilibrium one is called natural or irreversible. To make our discussion more precise, let us suppose that $\tau_{\text{rel}}$ is the time over which the system relaxes to a new equilibrium state and $\tau_{\text{con}}$ the time over which the value of the external constraints is varied. If $\tau_{\text{con}} \ll \tau_{\text{rel}}$, the system will find itself, suddenly, in a non-equilibrium state. On the contrary, when $\tau_{\text{con}} \gg \tau_{\text{rel}}$ the system readjusts almost instantaneously with respect to the modification of the external constraints. In this sense the system passes from an equilibrium state to another one. Suppose that we change the external constraints in several steps and at each step we change them by a small amount such that the system is always arbitrarily close to an equilibrium state. The succession of these steps yields a quasi-static process. This process is called reversible, if by reversing back to the old values of the external constraints, the system returns to the original equilibrium state without any modification of the external environment. An isolated system (fully constrained by isolating walls which forbid any contact with any other system) eventually will reach equilibrium. A thermally isolated system is enclosed within adiabatic walls, which are defined by the following property. Any change of the system may only occur by a displacement of the walls or by forces acting at distance.
1.2 The principles of thermodynamics

First principle and the internal energy

When a system which is thermally isolated changes its thermodynamic state, we say that the transformation is adiabatic and we denote by $W$ the work exchanged between the system and the environment during the transformation. By convention, we define $W > 0$ when the system makes work on the environment. We can state in general that for each thermally isolated system there is a state function $U$ named internal energy such that exchanged work corresponds to the variation of the energy $U$ of the system in going from an initial to a final state and does not depend on the intermediate steps of the transformation. For an adiabatic transformation we have

$$\Delta U = -W,$$

(1.5)

where $\Delta U$ represents the variation of internal energy between the final and the initial state.

In general the energy $U$ of a system can be changed by external factors which can perform work on the system, e.g. by compressing a fluid. In analogy with the work done by a force $f$ to move an object a distance $dx$, the generalized displacements are introduced as $\eta_i$ modifications of the internal parameters of the system to adjust to the variation of the external factors. This work for an infinitesimal quasi-static change can then be expressed by the product of the generalized force $f_i = \partial U / \partial \eta_i$ times the generalized displacement $d\eta_i$. For instance, on a macroscopic fluid a compression by a pressure $P$ determines a volume contraction $dV$ with a work $dW = PdV$, where $P$ assumes the role of generalized force and $V$ is the displacement $\eta$. Similarly, for a magnetic body the magnetic field $B$ and the magnetic moment variation $dM$ along the direction of $B$ determine an increase of energy with a “work” $dW = BdM$. Another example consists of the energy variation $dW = \mu dN$ necessary to change the number $N$ of particles of a macroscopic system where the chemical potential $\mu$ is the “force” which produces the change in $N$. The variables representing the external conditions of the system, like the pressure, the magnetic field and the chemical potential are intensive. Their effect is on extensive variables, like the volume, the magnetic moment and the number of particles characterizing the macroscopic state of the system. The pairs of variables specified by the generalized forces and the generalized displacements of the system define the conjugate variables with respect to the energy $U$.

For a non-thermally isolated system undergoing a transformation, the equality (1.5) is not valid. The thermodynamic state and the energy of a system can also be changed without any mechanical work, for instance by radiation, by lighting a fire underneath or by running a current in a resistor immersed in it. The energy flow into and out of the system is called heat, $Q$. By convention, we define $Q > 0$ when the system receives heat from the environment.

$^3$ This can be taken as the definition of the chemical potential which will be clarified later.
In general the first principle of thermodynamic states that in any thermodynamic process the balance of exchanged work and heat equals the variation of the internal energy $U$, in the formula

$$\Delta U = Q - W, \quad (1.6)$$

where $\Delta U$ is fully determined by the parameters specifying the initial and final states. The first principle may be expressed in differential form for infinitesimal changes in the thermodynamic parameters by

$$dU = \delta Q - \delta W, \quad (1.7)$$

where the symbol $\delta$ implies that the corresponding quantities $\delta Q$ and $\delta W$ are not exact differentials, in contrast to $dU$. The first principle is then a formulation of the conservation of energy, generalized to include thermal exchange. For a one-component fluid system with $\delta W = PdV$ an adiabatic process is such that

$$dU + PdV = 0 \quad \text{or} \quad \left[ P + \left( \frac{\partial U}{\partial V} \right)_{P} \right] dV + \left( \frac{\partial U}{\partial P} \right)_{V} dP = 0, \quad (1.8)$$

while an isothermic process is determined by

$$dT = \left( \frac{\partial T}{\partial V} \right)_{P} dV + \left( \frac{\partial T}{\partial P} \right)_{V} dP = 0 \quad \text{or} \quad \left( \frac{\partial P}{\partial V} \right)_{T} = -\left( \frac{\partial T}{\partial V} \right)_{P}, \quad \left( \frac{\partial T}{\partial P} \right)_{V}. \quad (1.9)$$

As is easy to verify for a perfect gas, in the $V-P$ plane the curves describing the isothermic and adiabatic processes are decreasing functions of $V$ with the slope in the adiabatic case being larger. The proof is provided in the answer to Problem 1.1.

The first principle distinguishes between the two different forms of energy which can be exchanged by a system with the environment. The addition of heat to a body is not due to the application of a generalized force but by conduction, convection or radiation. In all these cases the body changes its thermodynamic state as a result of a change of motion at the microscopic level. In thermodynamics this change must be characterized by an appropriate pair of global conjugate variables, the intensive variable temperature and the corresponding extensive entropy, which will be defined next by means of the second principle.

**Second principle and the entropy**

Before introducing the second principle of thermodynamic we must discuss the concept of hotness and how it is related to the temperature. When two systems being out of thermal equilibrium (at different empirical temperatures) are put in contact with one another, we define as the hotter system the one which loses heat to reach equilibrium and the same final empirical temperature together with the other system. By the zeroth law, we can unequivocally define a correspondence between hotness and empirical temperature: if we say that the system at higher empirical temperature is hotter, then it will be hotter than all the other systems at the same lower temperature.

The second principle essentially establishes the empirical observation that there are two different forms of energy transmission and that, via the state function entropy, the heat
cannot be transformed into work without losses allowing for a thermodynamic characterization of irreversibility. It can be stated as follows:

no process is possible, the sole result of which is that heat is transferred from a body to a hotter one (Clausius’s formulation)

or

no process is possible, the sole result of which is that a body is cooled and work is done (Kelvin’s formulation)

or

in any neighborhood of any thermodynamic state there are states that cannot be reached from it by an adiabatic process (Carathéodory’s formulation).

The logical sequence that leads to the thermodynamic characterization of the irreversibility is first the introduction of the absolute temperature \( T \), through the universal efficiency of a Carnot reversible cycle; then the existence of a state function \( S(T, V) \) called entropy, via the Clausius theorem. This theorem states that for an arbitrary reversible cycle \( C_R \)

\[
\oint_{C_R} \frac{\delta Q}{T} = 0, \tag{1.10}
\]

from which the definition of the variation of the entropy function follows:

\[
S(V_2, T_2) - S(V_1, T_1) = \int_{\Gamma_{rev}} \frac{\delta Q}{T}, \tag{1.11}
\]

where \( \Gamma_{rev} \) represents a reversible process from the state \( V_1, T_1 \) to \( V_2, T_2 \). The difference (1.11) is independent of the path connecting the initial and final states; the entropy is therefore a state function completely determined for each state once its value is fixed for one state. Finally, the function \( S(V, T) \) quantifies the irreversibility by stating that for an infinitesimal generic transformation

\[
\frac{\delta Q}{T} \leq dS. \tag{1.12}
\]

In the above the (in)equality applies to (ir)reversible processes. Clearly the entropy cannot diminish for a thermally isolated system for which \( \delta Q = 0 \).

We now follow more closely the logical sequence sketched above. The Carnot reversible cycle starts from a state \( (P_1, V_1) \) and returns to it via an isothermal expansion from \( (P_1, V_1) \) to \( (P_2, V_2) \) (the empirical temperature \( \theta_1 \), with the old notation for the time being, is maintained fixed by absorbing the heat \( Q_1 \) from a heat bath or reservoir), followed by an adiabatic expansion from \( (P_2, V_2) \) to \( (P_3, V_3) \), an isothermal compression from \( (P_3, V_3) \) to \( (P_4, V_4) \) (the empirical temperature \( \theta_2 \) is maintained fixed by giving the heat \( Q_2 \) to a heat bath), and finally an adiabatic compression back to \( (P_1, V_1) \). The balance of work on the basis of the first principle and the efficiency \( \eta \) defined as \( \eta = W/Q_1 \) are

\[
W = \oint PdV = Q_1 - Q_2, \quad \eta = 1 - \frac{Q_2}{Q_1}. \tag{1.13}
\]

4 For a discussion of this less common formulation, one may consult the book by Zemansky (1968).

5 A heat bath is a sufficiently large system capable of absorbing and yielding a finite amount of heat without changing its natural temperature.
Consider now two systems $R$ and $R'$ undergoing two Carnot cycles between the temperatures $\theta_1$ and $\theta_2$. If for simplicity we choose $Q'_1 = Q_1$, then $Q'_2 = Q_2$ or $W' = W$. If it were not so, by combining the Carnot cycle for $R'$ with the reverse one for $R$, one would violate the Kelvin statement of the second principle. See the answer to Problem 1.2 for the proof. The heat ratio must be an universal function solely depending on $\theta_1$ and $\theta_2$, $Q_2 / Q_1 = f(\theta_1, \theta_2)$.

Moreover, by considering two cycles working between $\theta_1$ and $\theta_2$ and between $\theta_2$ and $\theta_3$ and the combined cycle working between $\theta_1$ and $\theta_3$, one has

$$Q_3 / Q_1 = f(\theta_1, \theta_3) = f(\theta_1, \theta_2) f(\theta_2, \theta_3).$$

From the last equation one sees that the function $f$ factorizes:

$$f(\theta_1, \theta_2) = \frac{\phi(\theta_2)}{\phi(\theta_1)}.$$

The absolute temperature $T$ is then defined as the universal function of the empirical temperature $T = \alpha \phi(\theta)$. The efficiency $\eta$ is then

$$\eta = 1 - \frac{Q_2}{Q_1} = 1 - \frac{T_2}{T_1},$$

and therefore for a Carnot cycle

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} = 0,$$

for any reversible closed cycle the Clausius theorem (1.10) holds, as can be proved by covering the cycle by a net of Carnot cycles and going to the limit of finer and finer networks. The entropy function (1.11) is then defined. The scale $\alpha$ can be determined by the equation of state of a perfect gas.
For any cycle acting between the temperatures $T_1$ and $T_2$, to avoid contradiction with the second principle the Clausius theorem is modified into

$$\frac{Q_1}{T_1} + \frac{-Q_2}{T_2} \leq 0,$$

where the equality holds for reversible cycles. In general, the corresponding algebraic sum for all the transformations occurring in a cyclical process can only be less than zero, leading to the inequality for the Clausius theorem:

$$\oint_C \frac{\delta Q}{T} \leq 0$$

and to (1.12) in differential form for an infinitesimal change.

We remark that the existence of the entropy as state function is related to the fact that there exists an integrating factor for the infinitesimal heat exchange $\delta Q$ and that such an integrating factor is the inverse absolute temperature. By considering $T$ and $V$ as independent variables, from the first principle one has

$$\delta Q = \left(\frac{\partial U}{\partial T}\right)_V dT + \left[\left(\frac{\partial U}{\partial V}\right)_T + P\right] dV,$$

and the condition for $\delta Q$ being an exact differential is

$$\frac{\partial^2 U}{\partial V \partial T} = \frac{\partial}{\partial T} \left[\left(\frac{\partial U}{\partial V}\right)_T + P\right].$$

By considering, for instance, the perfect gas ($U = C_V T$ with the heat capacity at constant volume $C_V$ constant, and $PV = nRT$, $n$ being the number of moles and $R$ the gas constant), one soon sees that the above equation cannot be satisfied. On the other hand, by considering $\delta Q/T$, the condition for being an exact differential becomes

$$\frac{\partial}{\partial V} \frac{C_V}{T} = \frac{\partial}{\partial T} \left[\frac{1}{T} \left(\frac{\partial U}{\partial V}\right)_T + \frac{P}{T}\right],$$

and is obviously satisfied.

The definition of the state function entropy is the missing quantity to complete the set of thermodynamic variables arranged in pairs of conjugated variables. For a reversible transformation the entropy appears as the generalized displacement variable whose conjugate generalized force is the temperature.

The first principle establishes how a process can take place without violating energy conservation. The second principle fixes the direction of these processes. For example if two bodies at different temperatures are put in thermal contact the energy flows from the hotter to the colder and eventually the final temperature is the same. Generally speaking, however, we could imagine that the energy flows in the opposite direction as long as the total energy is conserved in agreement with the first principle. We know that such a process which would correspond to a diminishing of the entropy of the total isolated system never occurs. The second principle may be considered to be empirical and up to now there have been no known violations of it.
Third principle and the unattainability of the absolute zero

The third principle of thermodynamics, also known as Nernst's principle, is about the behavior of the entropy when the temperature decreases towards absolute zero. As for the second principle, there exist different formulations, starting from the one originally put forward by Nernst (1906), all logically equivalent.

The simplest formulation, which has a direct interpretation within the statistical mechanics, as we shall see later (see the discussion in Section 4.2), is the following.

The entropy of a system at the absolute zero of temperature is a universal constant, which can be set to zero.

In mathematical terms one has

$$\lim_{T \to 0} S(T) = 0.$$  \tag{1.17}

Alternatively, the third principle may be stated as the impossibility of attaining the absolute zero of temperature in a finite number of reversible thermodynamic transformations.

Finally, the original formulation of Nernst is usually enunciated in the Nernst–Simon form (see the book by Zemansky (1968)): the entropy variation of any isothermic reversible transformation of a condensed system (liquid or solid) tends to zero when the temperature decreases indefinitely. This can be mathematically expressed by

$$\lim_{T \to 0} (S(T, X_f) - S(T, X_i)) = 0.$$  \tag{1.18}

In the above $X$ indicates (besides the temperature) any other thermodynamic parameter (such as pressure or magnetic field), which can be varied reversibly during an isothermic transformation. Clearly, if one assumes the first formulation of Eq. (1.17), there can be no dependence on the value of $X$, and Eq. (1.18) follows. The logical connection of the first and third formulations with the impossibility of attaining absolute zero in a finite number of transformations requires the mathematical development of the consequences of the first two principles. This development will be carried out in the next two sections and the equivalence between the unattainability formulation and Eq. (1.18) is shown in Problem 1.4.

We will not then embark here in a detailed discussion of the equivalence of the various formulations and, for the time being, we limit ourselves to a graphical illustration following Zemansky (1968). To this end we refer to Fig. 1.3, where the behavior of the entropy as a function of temperature is shown for two different values of a parameter $X$. Both curves go to zero when $T \to 0$ according to Eq. (1.17). We can now imagine performing an isothermic reversible transformation (see thin arrow (1) in Fig. 1.3) starting on a point of the upper curve corresponding to $X = X_i$ and letting $X$ vary from $X_i$ to $X_f$. Then, we perform an isentropic transformation during which the parameter $X$ is reversibly changed back to the original value $X_i$ (see thin arrow (2) in Fig. 1.3). The net result of the two transformations is a decrease of the temperature of the system. In experimental practice, when the system is a magnetic salt and the parameter $X$ is the external magnetic field, such a sequence of transformations is called adiabatic demagnetization. Clearly, one can iterate the two-step transformation (see thin arrows (3) and (4)) and reach a further lower temperature. However, the amount of decrease of temperature in this second iteration is lower than in