# 1

## *Introduction*

### 1.1. Origins of thermodynamics

The increase of mechanization during the nineteenth century involved the construction of machines, such as the steam engine, for the conversion of heat energy into mechanical power. It was from the study of these heat engines that thermodynamics grew. The initial development was rapid. By 1900, the subject was firmly established, and although its application had at first been restricted to thermal engineering, its laws were soon recognized to be of such great generality as to be useful and important in many other branches of science also. Broadly speaking, thermodynamics is applicable to all processes in which temperature or heat play an important part. In physics, it provides a way of understanding phenomena as different as thermal radiation on the one hand and the low temperature properties of paramagnetic salts on the other. It supplies the basic theory of chemical reactions and underlies much of chemical engineering. It is applicable not only to steam engines but to refrigeration and rocketry.

With this very wide range of application, it is possible to adopt various terms of references within which to develop the subject. We shall choose examples which are primarily of interest to the physicist. The fundamental structure of the subject, however, is little affected by the applications one has in mind. This is because the basic theory can be developed in a precise and self-contained way with much of the rigour of a mathematical argument. To some this makes the subject seem too abstract and difficult, but we shall try to avoid this impression by developing the theory in the context of its applications to real physical systems.

## 1.2.   The macroscopic approach

Thermodynamics sets out to describe and correlate the directly observable properties of substances: the volume of a gas, the expansion of a wire, the polarization of a dielectric. These are all *macroscopic* quantities, properties of the materials in bulk. If we were interested in the pressure exerted by a gas on the walls of its container, we could, in principle, adopt the *microscopic* approach and start from the equations of motion of the individual molecules, examine the statistics of their motions and finally arrive at an expression for the macroscopic quantity, pressure, in terms of momentum exchange at the boundaries of the container. But for many purposes an analysis in microscopic terms is unnecessary. The laws of thermodynamics enable us to interrelate the macroscopic quantities *without making any microscopic assumptions* at all. The great generality of thermodynamics is a direct consequence of this. By avoiding commitment to any particular microscopic interpretation, thermodynamics is not limited to particular applications nor is it subject to the fashions of microscopic theory.

On the other hand, it is possible to associate particular kinds of macroscopic behaviour with certain general kinds of microscopic change: an anomaly in a specific heat, for instance, may result from a change of atomic ordering in a crystal. However, since no microscopic assumptions are built into the thermodynamics it is never possible to *identify* a microscopic process by thermodynamic reasoning alone.

It is perhaps because thermodynamics is not concerned with fundamentals in the microscopic sense that it sometimes does not appeal readily to the physicist; but he will disregard it at his peril. It is precisely because it avoids microscopic theories that it is so valuable. It often yields answers to problems where an understanding of the fundamental processes involved might be difficult or impossible. It also helps to prevent mistakes; for any result which does not satisfy the requirements of thermodynamics must be wrong. But, perhaps more important, a physicist's training is not only concerned with learning fundamental theories but also with developing a sensibility to the way in which physical systems behave, and here thermodynamics has a peculiar contribution to make by providing a very general framework of ideas from which the understanding of particular systems may more readily be achieved.

## 1.3.   The role of the laws

In seeking to derive relationships between directly observable quantities, thermodynamics is essentially formulating rules which these quantities must obey under given conditions. These may apply to a

substance undergoing a particular process, or they may be transformation rules which are useful in relating quantities which might be difficult to measure to ones which are more easily measured. For example, we will see that the ratio of the isothermal to the adiabatic compressibility is equal to the ratio of the principal heat capacities:

$$\frac{\kappa_T}{\kappa_S} = \gamma = \frac{C_p}{C_V}.$$

Such a relation follows logically from the laws of thermodynamics. If they are true, then this relationship must always be true. If an experiment gave some other result, then there would be something wrong with the experiment, for otherwise the whole structure of thermodynamics would collapse.

In order to derive these results in as simple a way as possible, it becomes necessary to define many new functions and concepts such as temperature, internal energy and entropy. With a given mass of gas, for example, we find that we usually only need specify its pressure and its volume to define its state precisely. These quantities are *direct observables*. But if we wish to describe how the pressure and volume will change if that gas flows down a tube of varying cross-section, it is convenient to introduce a quantity called *enthalpy* which is conserved in the process. These new and more abstract quantities enable us to characterize processes or conditions in a simple way. They might be constants in a given process or they might take some extremal value under given conditions. Having defined these new quantities, we must, of course, expect them to be related to each other and to the direct observables in a way which follows logically from their definitions.

Of these new concepts, three are fundamental. Each follows from one of the laws of thermodynamics. From the zeroth law we are able to give a precise meaning to *temperature*. From the first law we are able to define *internal energy*, and from the second, *entropy*.

### 1.4. Systems, surroundings, and boundaries

A thermodynamic *system* is that portion of the universe which we select for investigation. A system may be simple or complex; it may be homogeneous or it may consist of many parts. A gas in a cylinder is a simple system. A mixture of phenol and water is a more complicated one: it contains two different substances or *components*, and for certain concentrations and temperatures separates into two phases. *Phase* is defined as a system or part of a system which is homogeneous and has definite boundaries. A phase may be a chemically pure substance, or it

4        *Equilibrium thermodynamics*

may contain more than one component as is the case with the phenol–water mixture. When the phenol–water mixture separates, both components are present in both phases, but in different concentrations.

Everything outside the system is called the *surroundings*, and the system is separated from the surroundings by its *boundary* (Fig. 1.1). In many cases, the boundary of a system may simply be its surface, as with a drop of liquid; but it is often convenient to contain the system within *walls* of some special kind that allow or prevent various sorts of interaction between the system and its surroundings. When we come to consider how a system may interact with its surroundings we shall find that interactions may be divided into two kinds. We may influence a system by doing *work* on it, or we may influence it *thermally*. Compression and magnetization are examples of work-like processes while heating in a flame is an example of a thermal process. Then a *rigid* wall prevents a system from changing its volume or shape so that no work of a mechanical nature may be done on it. Walls which prevent any thermal interaction are called *adiabatic*,[1] and a system enclosed in adiabatic walls is said to be *thermally isolated*. Such a system cannot exchange heat with its surroundings, but it may be possible to do work on it. Any changes which it undergoes will be *adiabatic changes*. A Dewar vessel provides a good approximation to adiabatic walls. A wall which prevents *any* interaction between the system and its surroundings is called an *isolating* wall, and the system is then said to be *(completely) isolated*. Walls which are not adiabatic (but through which a system may be influenced thermally) are called *diathermal* and two systems separated by a diathermal wall are said to be in *thermal contact*.

Fig. 1.1. **A thermodynamic system.**
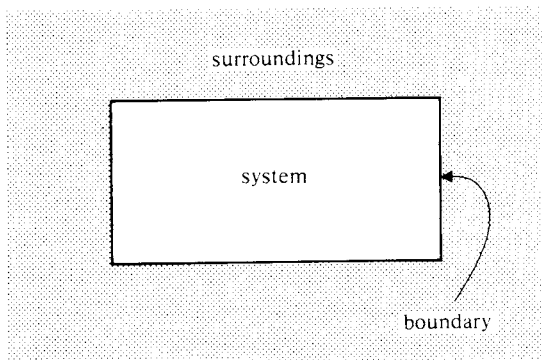


_____
[1] *Adiathermal* is also sometimes used.

In discussing chemical systems which contain different components, it is sometimes convenient to have a section of wall through which one or more of the components may pass while others are contained. Such a wall is called *semipermeable*. Hot quartz is permeable to helium but impermeable to other gases.

A system much used in developing the basic theory of thermodynamics is that of a gas contained in a smooth cylinder by a frictionless leakproof piston. This is a particularly helpful model system to refer to as it is easy to visualize how changes take place and how thermodynamic parameters may be varied.

### 1.5.    Thermodynamic variables

The thermodynamic variables comprise the direct observables and the 'new' quantities discussed in section 1.3. They may be divided into two classes. Those of the first class are essentially local in character and include such quantities as pressure, electric field, force, and density.[2] They are known as the *intensive* variables. Those of the second class correspond to some measure of the system as a whole and include such quantities as mass, volume, internal energy, and length. These are proportional to the mass of the system if the intensive variables are kept constant, and for this reason are known as *extensive* variables.

---

[2] These quantities are not strictly local in the sense that it is possible to define them at a point. For example, we would define the local pressure in a gas by

$$p = \lim_{a \to 0} (F/a)$$

where $F$ is the normal force exerted across a small area $a$. However, when $a \lesssim l^2$, where $l$ is the mean free path of the gas molecules, the discrete nature of the molecular impacts becomes apparent and this quantity will fluctuate violently. As $a$ becomes smaller, it becomes necessary to average over longer and longer times in order to achieve any similarity between pressure as we have defined it and its macroscopic counterpart. Similar restrictions apply to other intensive variables. In the case of electric field, the limit is set by the uncertainty principle, for we would define the local electric field by

$$E = \lim_{V \to 0} (\dot{p}/e)$$

where $\dot{p}$ is the rate of change of the momentum of a particle with charge $e$ which we confine in the volume $V$. As we make $V$ smaller so as to define $E$ more nearly at a point, we eventually introduce a large uncertainty in the momentum of the particle by the restriction $\Delta p V^{1/3} \sim \hbar$, and it becomes impossible to observe $\dot{p}$. Fortunately, these restrictions do not concern us here, because by adopting the macroscopic approach in thermodynamics we can never hope to use it to describe systems on the atomic or quantum scale. Indeed, it is precisely because macroscopic quantities cease to have meaning that we cannot do so.

It is often convenient to refer to extensive quantities in terms of their values *per unit mass* of the system. They are then called *specific* variables. Often, extensive variables are represented by capital letters and the derived specific quantities by the corresponding small letter. Thus, the volume of unit mass is called the specific volume and is given the symbol $v$.

Another useful convention is to add a suffix $m$ to an extensive quantity when the amount of substance referred to is one mole. Thus, $C_p$ is the heat capacity at constant pressure (unit, $J\,K^{-1}$) and $C_{mp}$ is the molar heat capacity at constant pressure (unit, $J\,K^{-1}\,mol^{-1}$). The molar suffix is frequently dropped if there is no danger of confusion.

Many of the direct observables form conjugate pairs such that their product has the dimensions of energy. For these, the intensive member of each pair has the character of a force, and the extensive member that of a displacement. Some of these are listed in Table 1.1, together with the kind of system to which they particularly apply.

Any quantity which takes a unique value for each state of a system is called a *function of state*. The direct observables are obviously functions of state. In principle, it must be possible to express any function of state in terms of any set of variables which is sufficient to define the state of the system.

### 1.6.    Thermodynamic equilibrium

When a system suffers a change in its surroundings, it will usually be seen to undergo change. If the bulb of a thermometer is placed in a beaker of warm water, the mercury will begin to expand and will start

Table 1.1. *Some conjugate pairs of thermodynamic variables*

| System | Intensive variable | Extensive variable |
|---|---|---|
| fluid | pressure, $p$ | volume, $V$ |
| filament | tensional force, $f$ | length, $L$ |
| film | surface tension, $\gamma$ | area, $A$ |
| electrical | potential difference, $\mathscr{E}$ | charge, $Z$ |
| dielectric | electric field, $E$ | electric dipole moment,[a] $p$ |
| magnetic | flux density, $B$ | magnetic dipole moment, $m$ |
| all systems | temperature, $T$ | entropy, $S$ |
| generalized | force, $X$ | displacement, $x$ |

[a] We prefer to use the conventional symbol $p$ here despite the occasional possibility of confusion with pressure when problems involve both variables. In practice, the significance of the symbol is usually obvious from the dimensions of the quantities with which it appears.

to rise in the capillary. After a time, however, the system will be found to reach a state where no further change takes place and it is then said to have come to *thermodynamic equilibrium*. In general, the approach to thermodynamic equilibrium will involve both thermal and work-like interactions with the surroundings.

Similarly, if we place two systems in thermal contact, we generally find that changes will occur in both. When there is no longer any change (each has reached a state of thermodynamic equilibrium) the two systems are said to be in *thermal equilibrium*. In this case, we have prevented work-like interaction and allowed thermal interaction only. We shall eventually describe such a situation by saying that heat flows from one system to the other until they are at the same temperature.

Like mechanics, thermodynamics knows several kinds of equilibrium, and we define the stability of the equilibrium in a similar way. Thus a system is said to be in *stable* equilibrium if, after being slightly displaced, it returns to its original state. A system is in *metastable* equilibrium if it is stable for small displacements but unstable for larger ones. Certain systems also exhibit *neutral* equilibrium. Such systems may be displaced but will remain in the displaced condition when released. If a system is unstable to infinitesimal displacements it is said to be in *unstable* equilibrium.

It is worth pointing out that in the strictest sense neither mechanics nor thermodynamics knows truly unstable equilibrium, and the reason in both cases is the same. Equilibrium is defined in terms of macroscopic variables which are large scale averages of quantities which, on the microscopic scale, are subject to fluctuations. Pressure exerted by a gas is the macroscopic average of the impulses from discrete molecular impacts. The atoms of a solid are always in thermal motion. Although in large systems these fluctuations may be relatively unimportant, any fluctuation, however small, is, by definition, sufficient to destroy unstable equilibrium. Thus no truly unstable equilibrium exists although in some systems the size of the displacement for which the system remains in metastable equilibrium may be so small that the system is loosely spoken of as being unstable. Being defined in terms of macroscopic quantities, equilibrium is itself a macroscopic concept. We may only apply the idea of equilibrium to large bodies, to systems of many particles. The Brownian movement of a colloid particle shows that it is certainly not in equilibrium. On the other hand, the mean density of colloid particles at different heights in a suspension does obey rules which may be derived from our ideas of equilibrium. Some examples of the different kinds of equilibrium are illustrated in Fig. 1.2.

8          *Equilibrium thermodynamics*

### 1.7.     Thermodynamic reversibility

When a system undergoes a series of changes a thermodynamic *process* is said to take place. A process is said to be reversible if, and only if, its direction can be reversed by an *infinitesimal* change in the conditions; it is not enough if it may only be reversed by a finite change. Thermodynamic reversibility requires two conditions to be satisfied: the process must be *quasistatic* and there must be *no hysteresis*.

*Quasistatic processes.* To be quasistatic, a process must be carried out so slowly that every state through which the system passes may be considered an equilibrium state. Strictly speaking, this means that the process should be carried out infinitely slowly. Fast changes cause disequilibrium between different parts of a system. For example, suppose a gas is to be compressed from the state $(p_1, V_1)$ to the state $(p_2, V_2)$ (Fig. 1.3). If the compression is performed sufficiently slowly, there will

**Fig. 1.2. Different kinds of equilibrium.**

*Stable equilibrium.* A ball in a depression and a cylinder of gas at constant temperature will both eventually return to their initial states if displaced and released.

*Neutral equilibrium.* A ball on a horizontal plane may be displaced to any position on the plane and will remain there when released. Similarly, a system consisting of a liquid and its vapour at constant temperature also shows neutral equilibrium, for the vapour pressure depends only on temperature so that as long as both phases are present, change of volume simply causes condensation or vaporization without change of pressure and the system remains in equilibrium with its surroundings.

*Metastable equilibrium.* A ball in a small hollow on an otherwise convex surface is only stable to small displacements. A mixture of hydrogen and oxygen in a thermally isolated vessel is also stable to small displacements but a large compression could raise the temperature sufficiently for the mixture to explode.
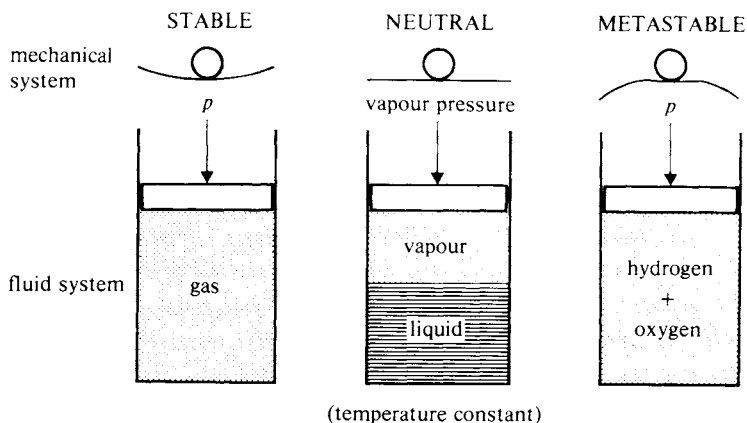
always be time for the gas to adjust to the changing environment and it will never depart significantly from equilibrium with it. Every state which the gas passes through will be an equilibrium state, and, clearly, the process may be reversed at any time by reversing the operations on the system. If, instead, we compress the gas rapidly by a sudden movement of the piston, sound waves or shock waves will be set up in the gas creating regions of different pressure and temperature. Clearly, such a change is not reversible. (We cannot extract the sound waves by moving the piston out again.)

Since an equilibrium state corresponds to definite values of the system parameters, we may represent a quasistatic process by plotting how the parameters vary as the system passes from the initial to the final state (Fig. 1.3). Such a representation is known as an *indicator diagram*. In a non-quasistatic change, system parameters do not define the states through which the system passes nor can they describe the processes it undergoes. Non-quasistatic processes should therefore not be represented by a line on an indicator diagram.

*Hysteresis.* When a process is reversed in a system with hysteresis, the system does not retrace its previous path, but proceeds by a different one. A common example is found in the magnetization of iron (Fig. 1.4). If carried out sufficiently slowly, each state through which the system passes may be considered as an equilibrium state. The variables are, at all times, well defined and the process may be represented on an indicator diagram. Nevertheless, here also, it is clear that the system parameters do not uniquely define the state of the system since their

**Fig. 1.3. The indicator diagram for a reversible process in a simple fluid.**
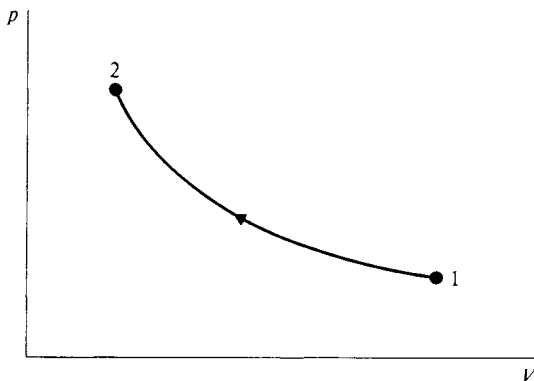
relationship depends on the previous history of the system. Friction is a common cause of hysteresis.[3]

### 1.8.    Degrees of freedom

For any particular system we may list many thermodynamic variables which describe different aspects of its properties. If they are functions of state, the values of these variables will be determined by the state of the system: for a particular state, they will have a particular set of values. Many of them, however, will be related to one another in some way. For example, density is the ratio of mass to volume. We may therefore ask: what is the minimum number of variables whose values

**Fig. 1.4.  An example of hysteresis: the magnetization of iron.**
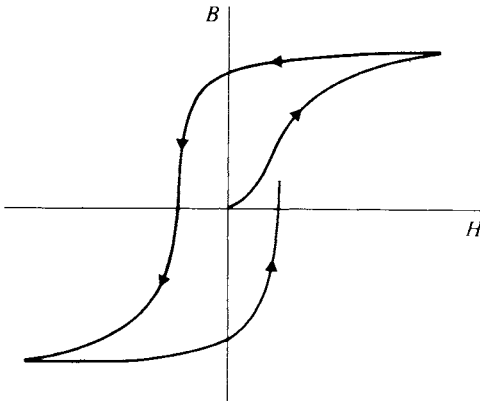


[3] Strictly speaking, the distinction between a non-quasistatic process and a hysteretic process is only one of time scale. If we place a piece of iron in a magnetic field there is a unique state of the iron corresponding to the lowest energy configuration of the system. As soon as the field is applied, we are really placing the iron in a *metastable* state from which it would have to proceed to the truly stable state by a series of minute changes on a molecular scale. The potential barriers between these steps, however, are so large that the approach to equilibrium only proceeds at an extremely slow rate, one which is quite negligible on any normal time scale. Since a hysteretic process is, in this sense, a non-quasistatic one on an enormous time scale, we might expect there to be inhomogeneities within the system as with normal non-quasistatic processes. From our knowledge of ferromagnetism we know this to be the case. The irregularities in the motion of domain walls, which may be observed under a microscope, indicate the presence of inhomogeneities which hinder magnetic rearrangement. It is precisely these inhomogeneities which prevent the attainment of the true equilibrium state within any normal period of time. Since, by and large, the times taken by thermodynamic systems to reach true equilibrium tend to one or other of the extremes, it is convenient to retain the distinction between the non-quasistatic and hysteretic processes.