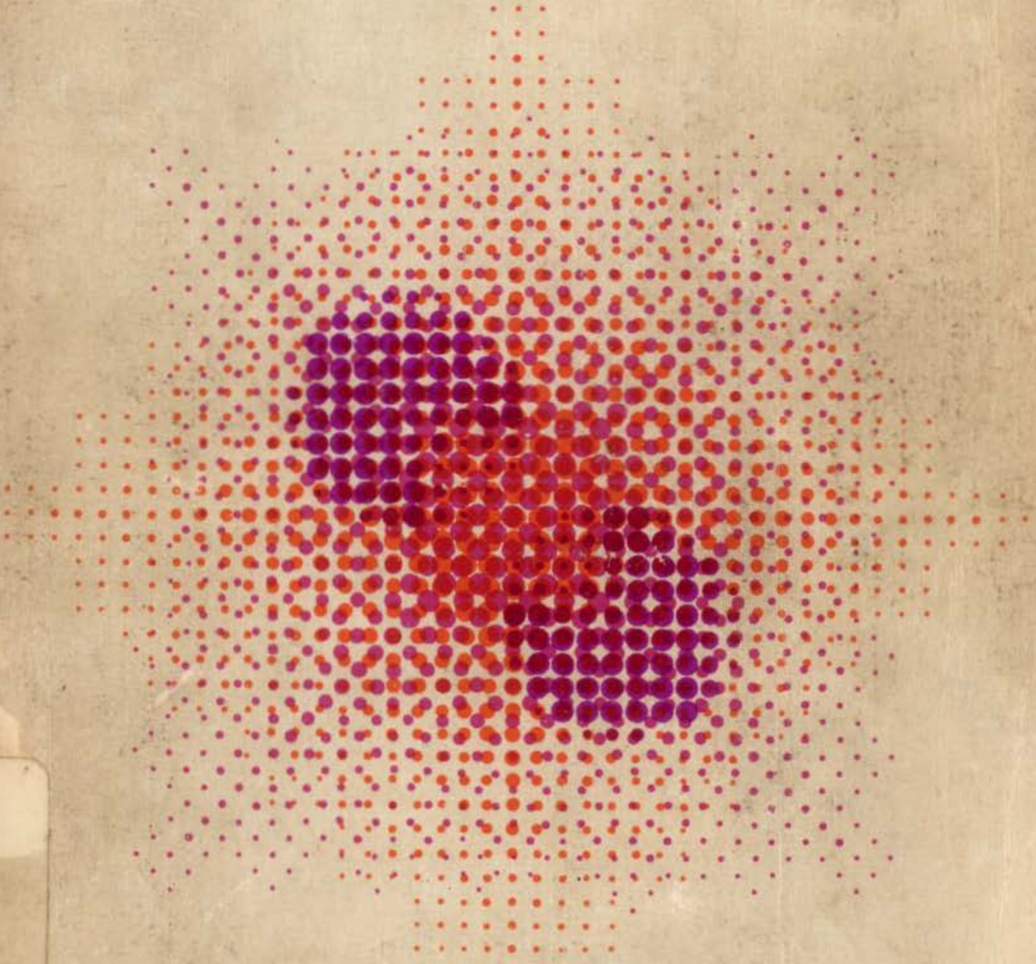
Penguin library of physical sciences: Physics

Classical thermodynamics

D. Elwell and A. J. Pointon



Classical thermodynamics

D. Elwell ar
A. J. Pointo

Chapter 1

The Description of Matter in Bulk

A sample of material – solid, liquid or gas – may be described in one of two basic ways. On the one hand, it may be described in terms of the component atoms and molecules of which it is formed, of the interactions between these components, and of the boundaries imposed upon them by the sample shape or the containing vessel. Alternatively, it may be described in terms of the bulk properties of the sample, which are determined in practical measurements, such as *mass, volume, density, specific heat, dielectric constant, elastic moduli, magnetic permeability, surface tension, thermal expansion coefficient, thermal conductivity* and so on.

These measurable quantities may be said to describe the various features of a body of material in its interaction with its surroundings. They also describe those phenomena observed during experiments on the body in which the state of the body is affected by external constraints. Their values may be varied by the application of pressure, or electric and magnetic fields, or by changing the temperature. In the study of thermodynamics it is these bulk properties or parameters which are of interest and, in particular, the dependence of these parameters on temperature. At the same time, since the temperature of a body is so closely related to the energy content of that body, this study must also involve consideration of the changes of energy which occur in bulk processes and in particular the effect of heat energy and work energy on the state of a body.

If any useful information is to be obtained about the behaviour of matter in bulk, in terms of these macroscopic properties, and if predictions are to be made of this behaviour under varying conditions, it is necessary that fundamental laws regarding these properties should be established in as simple a manner as possible. The fact that such laws can be established is the basis of the logical structure which is known as thermodynamics, and this leads to many relationships between the various properties. At first sight these relationships are surprising in their complexity, in view of the substantial omission of any detailed description of the component particles of the materials under consideration.

The four laws on which the structure of thermodynamics largely depends, and which will be described in the succeeding chapters, may be summarized as follows.

The zeroth law of thermodynamics defines the condition known as thermal equilibrium and hence allows a definition of temperature.

The first law of thermodynamics specifies the energy balance which exists in any closed thermodynamic system.

The second law of thermodynamics determines the direction in which thermodynamic processes will occur and the equilibrium state of a physical system.

The third law of thermodynamics sets a limit to both the thermodynamic temperature and entropy.

A *thermodynamic system,* as referred to in this summary (and also later in the text), may be defined conveniently as a bulk system of matter which has properties and measurable parameters which depend on temperature, pressure and other imposed conditions. (As a special case, an evacuated enclosure in which there is energy present in the form of electromagnetic radiation is also considered as a thermodynamic system.) The particular condition of such a system, which is specified by assigning values to these various properties, is termed the *state* of the system (which is *not* to be confused with the *phase* of the system as referred to in the term 'liquid state'). A *thermodynamic process* is any process which involves a change in this state. If the system does not interact with its surroundings, by for example exchanging heat, energy or material with them, it is said to be *isolated* from its surroundings. This state of isolation will not be the most common case, however, and normally the system will interact with its surroundings in some specified manner, so that any thermodynamic process involving the system will also involve its surroundings. In some cases, the system and its surroundings are referred to as the *universe*.

In any macroscopic thermodynamic system there will be components such as molecules, ions or photons. If the system is to obey the generalized macroscopic laws of thermodynamics it must necessarily consist of a sufficient quantity of matter, or a sufficient number of components, that fluctuations in the arrangement of these components, their position, concentration, momentum, etc. do not affect significantly any measured macroscopic properties of the system. This condition must apply even if, for mathematical convenience, an 'infinitesimally small' element of volume is considered.

Although the laws of thermodynamics are sometimes 'shown' to follow very closely from logic and common sense, it must be remembered that they are subject, at all times, to experimental test and may be regarded as being firmly based on experiment and observation. For this reason these laws are used to establish a methodology which may be applied in the treatment of practical problems, as will be seen from the many examples of such application which will be given by way of illustration.

Chapter 2
The Zeroth Law and Temperature

2.1 Thermal equilibrium

An important concept in the study of thermodynamics is that of *thermal equilibrium.* The concept of such equilibrium, involving as it does the establishment of a steady state from which there are no macroscopic fluctuations, is largely intuitive. For example, consider a thermodynamic system A, one property of which is monitored continuously. This property might be the electrical resistance of a piece of wire, or the pressure of a gas contained in some restraining enclosure. If the system is in an environment which is not subject to fluctuation or to varying external influences, the chosen property of the system will tend to a steady value which will cease to vary with time. If the system is then introduced into a new stable environment, the value of the chosen property will at first change with time and will then settle down to some new steady value. When this steady state is reached the system is said to be in equilibrium with its surroundings.

Consider now that the system A is, by some means, isolated from any contact with its surroundings. (Such isolation might be obtained approximately in practice by suspending the system by fine threads inside a highly evacuated container with highly reflecting walls.) The properties of the system A will now

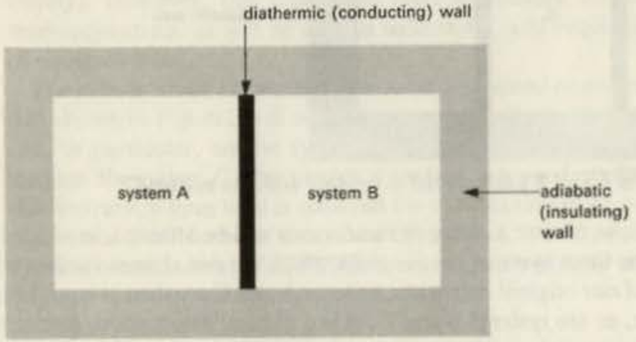


Figure 1 Systems A and B in thermal contact. (The adiabatic wall could be, for example, an evacuated region surrounding the system)

remain constant and no change will be observed, over any time period, in the value of the chosen property. If a second system B, which has a previous history very different from that of A, is introduced close to A but isolated from it, and from its surroundings, the properties of B will also remain unaltered with time. Suppose next that the systems A and B be brought into intimate contact within the enclosure in which they are isolated, as in Figure 1, such that energy can be exchanged between them through a diathermic or thermally conducting wall. As energy is exchanged through this wall, the properties of the two systems will change until, after a period of time, a steady state is reached in which there are no further observable changes. The two systems are then said to be in thermal equilibrium.

The establishment of thermal equilibrium does not, in fact, mean that the system will be in complete, or *thermodynamic*, equilibrium. The definition given above ignores, for example, any exchange of material, rather than energy, between the systems. The condition for *chemical* equilibrium is discussed in section 19.1. In addition, those cases have not been considered in which there is *mechanical* disequilibrium because one system exerts a force on the other.

2.2 The zeroth law of thermodynamics

Consider now that, within their isolating enclosure, the systems A and B in their initial disequilibrium states are each placed in thermal contact with a third system C, so that, as indicated in Figure 2, they are not in contact with each other.

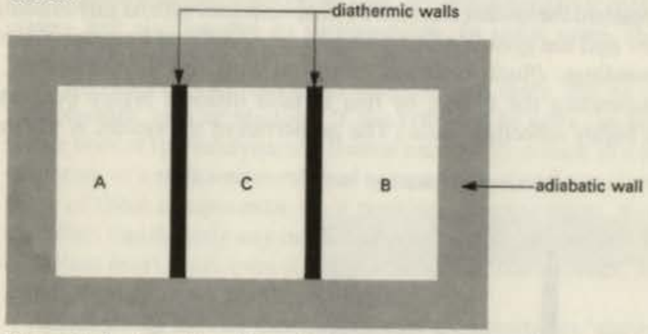


Figure 2 Systems A and B each placed in contact with the system C

Again after lapse of time a stable or steady state will be attained, in which the properties of the three systems reach values which do not change further with time. In terms of our original definition, systems A and C are then in equilibrium with each other, as are systems B and C. It is a natural assumption, intuitively based, that the systems A and B, though not in intimate contact with each other, are then also in thermal equilibrium. The postulate that this is so is referred to as the *zeroth* law of thermodynamics ('zeroth' because it was formulated after the 'first law' had been well accepted). There is no experimental evidence which runs contrary to this law. Indeed we may imagine its truth to be tested in the arrangement of Figure 2 by breaking the contact between the systems A and C and the systems B and C, once the steady state has been reached, and establishing such contact between A and B. All the evidence is that there would then be no further change with time of the values of any monitored properties of the systems A and B.

The zeroth law of thermodynamics is most usefully stated in the form:

When two thermodynamic systems are each in thermal equilibrium with a third system, they are also in thermal equilibrium with one another.

When a number of systems are in equilibrium according to this definition it is convenient to consider that there is some property which has the same value for each system, however different these systems may be in size and composition. This property is defined to be the *temperature* of the systems. Thus in the situation represented in Figure 1, when thermal equilibrium has been reached the temperature is the same for the two systems A and B; similarly in Figure 2, the three systems A, B and C will attain the same temperature.

This very necessary, basic definition of the condition for thermal equilibrium does not define a *scale* of temperature any more than the definition of distance defines the metre. However, the corresponding establishment of scales of temperature, which is discussed in detail later, will be seen to require far greater elaboration than does the cutting of two notches in a piece of steel as a standard of length. On the other hand, it is clearly possible to recognize a *difference* in temperature between two systems from the fact that when they are placed in thermal contact these bodies approach a state of equilibrium from a state of disequilibrium. It is further possible to state that, during the establishment of this equilibrium, some quantity known as *heat* flows from one system to the other. If this heat flows unidirectionally from the system A to the system B, the system A is said to be 'hotter', or at a higher temperature, than B (and conversely). However, this discussion then encroaches on the second law of thermodynamics, as will be seen in section 4.1, and requires a close definition of the term 'heat'.

The rate at which the equilibrium state is attained in an arrangement such as that shown in Figure 2 will depend rather critically on the nature of the systems and, in particular, on the system C. If the equilibrium is established rapidly through the system C, this system is said to be a *good conductor* of heat. If, on the contrary, a long time is required for the establishment of equilibrium, then the system C will be a *bad* conductor of heat and, if C should be a high vacuum, the rate of attainment of equilibrium may be substantially zero.

Chapter 3
The First Law

3.1 Work and energy

If a body is subjected to a force F such that it is caused to move through a distance l, then the force is said to perform an amount of *work* W on the body. The quantity of work is defined in mechanics as the product of the force with the distance through which the point of application moves in the direction of the force. (In mathematical terms this is equal to $W = \mathbf{F}.\mathbf{l}$, the scalar product of the vectors $\mathbf{F}$ and $\mathbf{l}$.) In the case where the body on which the force acts is free to accelerate, it is shown in Newtonian mechanics that the work done results in the creation of an equivalent amount of kinetic energy. That energy is $W = \frac{1}{2}mu^2$, where m is the mass of the body and u is the resultant velocity (assuming that the body moves from rest). If the body is now decelerated and brought to rest so that it loses its kinetic energy, the force exerted by the body during deceleration will perform work exactly equal to the kinetic energy $\frac{1}{2}mu^2$.

It is thus possible to state that, in mechanics, work may be converted to an equivalent quantity of kinetic energy, and vice versa, so that, if work is considered as another form of energy, it would be possible to state that energy is conserved in this single system. However, it is also found that there are systems in mechanics which, although they are not in motion, are capable of performing work. Thus a coiled spring, a compressed gas, or a weight suspended above the earth, are each capable, on being released, of causing work to be performed and kinetic energy generated. To account for the appearance of this kinetic energy it has been found convenient to ascribe to the system which performs the work a latent or *potential* energy. Then, ignoring the intermediary role played by the work, it is possible to state that, if the initiating system, the spring, the gas or the suspended weight, produces in some object (including itself) a certain amount of kinetic energy, it must have lost an equivalent amount of potential energy. This statement appears in mechanics as the *law of conservation of energy* written in the form:

Gain in kinetic energy = loss in potential energy

or, *The sum of the potential energy and the kinetic energy is a constant.*

3.2 Forms of work

In thermodynamics the force which is exerted by one system on another, or by one body on another, need not be of the normal mechanical form. There can be, for example, work performed by an electric field on an electric charge, or by a magnetic field on a magnetic dipole. However, the main distinction which must be made is between *external work*, which may be performed *on* a system, and the *internal work*, which is performed *within* a system. When, for example, a heated gas expands and drives a piston, there will be work done by the gas on the piston, and this will be observed outside the gas as external work. At the same time, if there are attractive forces between the molecules of the gas, there will be work done by these molecules as they move apart during expansion. This work will not be observed in the production of kinetic energy or potential energy outside the gas, and is therefore referred to as internal work. We shall, in general, be concerned with the external work and it is useful here to consider a number of illustrative examples in which work is performed on a particular system by means of an externally impressed force.

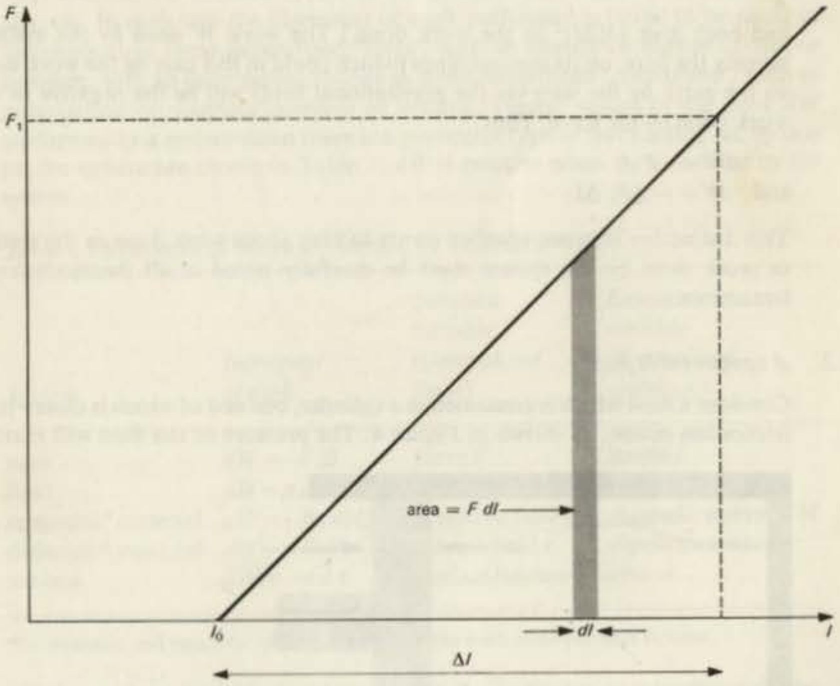


Figure 3 Wire obeying Hooke's law

3.2.1 *An extensible wire*

If a force F is applied to a wire, possibly by means of a weight suspended from the wire, and this force causes the wire to extend from a length l to a length $l+dl$, then the work done by the force on the wire will be $dW' = F\,dl$. For the case of a wire obeying Hooke's law, a force F will produce an extension from the original length l_0 to a length $l = l_0 + F/\lambda$, where λ, the modulus of the wire, is equal to the product of the area of the cross-section and Young's modulus divided by the original length l_0. This extension may be expressed in the form of the graph of Figure 3. The work $F\,dl$ may be represented by the area shown shaded and the total work done in extending the wire from l_0 to $l_0 + \Delta l$ as the force is increased from zero to F_1 is thus the total area up to Δl or

$$W' = \tfrac{1}{2}F_1\,\Delta l.$$

(The same result is obtained by evaluating

$$\int_0^{\Delta l} F\,dl,$$

and both give $\frac{1}{2}\lambda(\Delta l)^2$ as the work done.) The work W done *by* the system, namely the wire, on its surroundings (which could in this case be the work done *on* the earth *by* the wire via the gravitational field) will be the negative of the work done by the force. Thus

$$dW = -F\,dl$$

and $W = -\tfrac{1}{2}F_1\,\Delta l.$

This distinction between whether we are talking about work done *on* the system or work done *by* the system must be carefully noted in all thermodynamic treatments.

3.2.2 *A compressible fluid*

Consider a fluid which is contained in a cylinder, one end of which is closed by a frictionless piston, as shown in Figure 4. The pressure of the fluid will exert a force on the piston in a direction tending to increase the volume of the fluid and, at equilibrium, there will be an equal and opposite pressure acting on the outside surface of the piston. If the area of the piston is A and the pressure exerted is p, the force on the piston will be

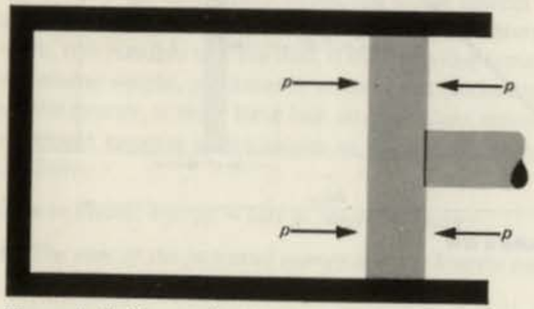


Figure 4 Fluid contained in a cylinder fitted with a frictionless piston

$$F = pA.$$

When the piston moves a distance dx in the direction of the force, the work done *by* the fluid will be

$$dW = F\,dx = pA\,dx.$$

The change in volume of the fluid during this movement dx is given by $dv = A\,dx$, so that $dW = p\,dv$. **3.1**

3.2.3 *General cases*

As has been stated, the work done by a system may be performed in a number of ways and may be the result of forces arising electrically, magnetically, mechanically, etc. In each case the increment of work performed is found to be equal to the product of some generalized 'force', such as electric or magnetic field or pressure, with an incremental change in some generalized 'coordinate', such as polarization or magnetic moment or volume. Typical values of the work dW performed *by* a system when there is a particular type of force acting on, or due to, the system are shown in Table 1; dW is positive when work is done by the system.

Table 1 Increments of Work Performed in Various Systems

System	*Increment of work*	*Intensive variable (Generalized force)*	*Extensive variable (Generalized coordinate)*
wire	$dW = -F\,dl$	force F	length l
fluid	$dW = p\,dv$	pressure p	volume v
magnetic* material	$dW = -B_0\,dM/\mu_0$	magnetic field B_0	magnetic moment M
dielectric* material	$dW = -E\,dP$	electric field E	dipole moment P
surface	$dW = -\sigma\,dA$	surface tension σ	area A

*In dielectric and magnetic materials we quote the work done per unit volume.

It is seen from the table that the generalized coordinates, or extensive variables, are proportional to the amount of material or to the extension of the system, while the generalized forces, or intensive variables, are independent of the size of the system.

3.3 Internal energy

When the compressible fluid referred to in 3.2.2 expands and performs work dW on its surroundings, this work will result, in an ideal mechanical context, in the generation of either a potential energy or a kinetic energy, equal to dW. If the fluid of Figure 4 is surrounded by adiabatic walls, the mechanical law of conservation of energy will cease to apply unless it is assumed that, in some way, the fluid has lost an amount of energy also equal to dW. Similarly, if a voltaic cell produces an electric current which results in the generation of kinetic energy of an electric motor, or if the action of surface tension causes the movement of a thread across the surface of a liquid, there is an imbalance in the energy equation. The problem is resolved if it is supposed that each system possesses a certain *internal* energy, which is a function of the state of that system. For example, consider that in its initial condition the fluid of Figure 4 has a definite internal energy U. During the expansion of the fluid, this internal energy changes to $U + dU$. It is then possible to maintain the principle of energy conservation by requiring that

$$dU = -dW \quad \text{or} \quad d(W + U) = 0, \tag{3.2}$$

so that the total energy of the system and its surroundings remains constant.

If the internal energy of a system is to be a definite quantity when the system is in a given state (that is, when the system is at a given temperature with a given set of constraints acting upon it), this internal energy can only be a function of that state of the system and not of the way in which that state was attained. Thus, if the expansion of the fluid is reversed, so that work dW is performed on the fluid while its temperature and other constraints return to their original values, the internal energy must return from the value $U + dU$ to the original value U.

3.4 Heat

It has been postulated that, during the establishment of thermal equilibrium between two bodies which were initially at different temperatures, there is a flow of *heat* from the hotter body to the colder body. That this quantity heat has the nature of energy (rather than, say, the nature of the postulated fluid *caloric*) has been long established by the work of Count Rumford and Humphrey Davy. These workers showed that the temperature of a system could be made to rise purely by the performance of work on the system and that the greater the amount of work which was performed the greater was this rise in temperature. The subsequent quantitative experiments performed to establish the exact equivalence of heat and work, and hence of heat and energy, were performed by such workers as Joule and Rowland and, more recently, by Laby and Hercus and others. The results of these experiments were largely discussed in terms of a quantity known as 'the mechanical equivalent of heat', which defined the amount of mechanical energy required to produce a given quantity of heat such as, for example, that required to raise the temperature of one gramme of water by one degree on the

centigrade scale. Defined in terms of the normal heat unit, which was derived to account for the changes in temperature when heat passed from one body to another, this mechanical equivalent of heat was found to be

$J = 4{\cdot}186$ joules per 15 °C calorie,

where one calorie is the amount of heat required to raise the temperature of one gramme of water by one degree centigrade at 15 °C. However, since it is established that heat and work are entirely equivalent, this statement really defines the specific heat of water at 15 °C to be 4·186 joules per gramme per degree centigrade.

3.5 The first law of thermodynamics

The experiments by which the quantity J has been measured were carried out under substantially adiabatic conditions so that, by **3.2**, the work done *on* any system to raise its temperature simply goes to increase the internal energy of the system by an amount equal to the work done. Once the equivalence of heat and work have been established it becomes apparent that the addition of a given quantity of heat to a system will simply increase the internal energy by an equivalent amount, provided no other change takes place.

Consider again the expansion of the fluid of Figure 4. If the adiabatic walls are replaced by thermally conducting walls, it is possible that, during the expansion, heat dQ will flow through the walls into the fluid. The work dW done by the fluid, which decreases the internal energy of the fluid, will be counteracted by this heat dQ. With the internal energy content defined so that the energy balance is preserved, **3.2** will now be replaced by

$$d(W + U) = dQ.$$

This implies that the work done plus the change in internal energy must just equal the energy added to the fluid in the form of heat and the normal form of this energy-balance equation is

$$dQ = dU + dW, \tag{3.3}$$

which, in mathematical form, expresses the *first law of thermodynamics*. The extension of this expression to include any system, with the internal energy in the form of potential energy of a spring, the kinetic energy of the molecules of a gas, or the chemical energy of a voltaic cell, allows us to express this first law in the form:

In a conservative system energy is neither created nor destroyed but may be converted from one form to another.

In this case the work dW represents an interchange of energy between one system and a neighbouring system which together form a conservative system. The work dW may thus appear as kinetic or potential energy, internal energy or, where the work is done against frictional forces, it may be dissipated as heat. (The law could be generalized to include the conversion of mass to energy.)

3.6 The inexact differential

Although it may appear initially to introduce mathematical complications, it is of particular importance in thermodynamics to establish how various quantities, such as pressure and temperature, work and heat, depend on the *state* of a system as opposed to the method or manner by which that state was attained. This may be illustrated by considering the fluid discussed in relation to Figure 4 and imagining that this fluid is heated and expands while doing work upon its surroundings. Such an expansion could take the fluid from the state A (with pressure p_1 and volume v_1) of Figure 5(a) along the path I to finish at the state B (p_2, v_2). Since, for a small change dv in volume, **3.1** gives the work done as $p\,dv$, the work done by the fluid during the expansion would be

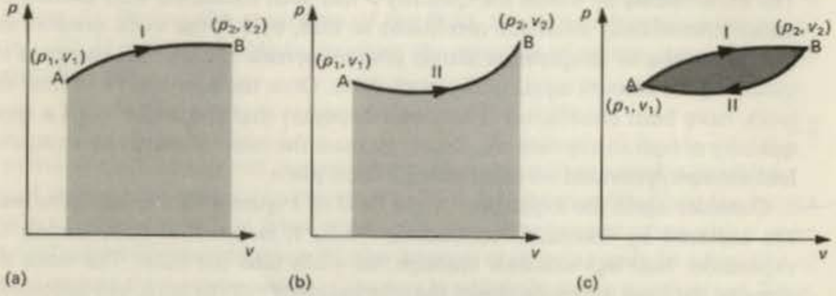


Figure 5 Work done by a fluid for different paths between two states

$$\int_A^B p\,dv = \int_{A\,(\text{path I})}^{B} dW,$$

which is equal to the shaded area under the line AB. If now the temperature of the fluid were varied during the expansion, so that its pressure at a given volume were different from the value along the path I, the expansion from A to B might be caused to occur, for example, along the path marked II in Figure 5(b), provided that the pressure and temperature and hence the volume returned to the required values at the end of the path. It is clear that the work done in this case, being the shaded area under this new path, is different from that in Figure 5(a). Thus

$$\int_{A\,(\text{path I})}^{B} dW \neq \int_{A\,(\text{path II})}^{B} dW, \qquad \textbf{3.4}$$

even though the initial and final states of the fluid are exactly the same for the two paths. In a similar manner it may be shown that this type of inequality may arise for the work done by any system when the system changes by two different routes between a given initial state and a given final state.

If a cyclic process is now constructed such that the fluid is expanded from A along the path I and compressed from B back to A along path II, the work done during such a process is clearly equal to the cross-hatched area between the paths, as shown in Figure 5(c). However, the system returns to its initial state at A, and we have postulated that there is a definite internal energy for a system in a given state. There can therefore be no change in the internal energy of the system over the complete cycle and hence, by the first law, any energy which was necessary to supply the work done by the fluid over the cycle must have been gained in the form of heat from some source or sources external to the fluid. It is possible to generalize therefore that, for a cyclic process, in which we require $\oint dU = 0$,

$$\oint dW \neq 0, \quad \oint dQ \neq 0$$

and, from **3.3**,

$$\oint dQ = \oint dW, \qquad \textbf{3.5}$$

where $\oint$ indicates integration round a complete cycle returning to the original state. It is usual, on the basis of these equations, to state that dU is a *perfect*, *exact* or *total* differential while dQ and dW are inexact differentials. An alternative, but entirely equivalent, statement is that U is a function of the state of the system alone, while Q and W are functions both of the state of the system and of the particular process by which that state has been attained.

Since it is obvious from the discussion of the processes involved in Figure 5 that the initial and final states at A and B can only be specified if the values of the temperature, pressure and volume are fixed for these states, it follows that we can write

$$\oint dT = \oint dp = \oint dv = 0.$$

Thus T, p and v are functions of the state of the system only, and dT, dp and dv are therefore exact differentials.

Chapter 4
The Second Law

4.1 Reversible and irreversible processes

The first law of thermodynamics, with its requirement that during any process which changes the state of a system the overall energy shall be conserved, provides a powerful tool for determining which thermodynamic processes are allowed and which are not. However, this law does not give any indication whether a particular process *will* occur, but only whether the process is allowed to occur by the condition that the energy is conserved. Thus, for example, it is just as possible for energy to be conserved when a person winds up a clock spring as when the spring is unwound during the driving of the clock mechanism. In the first case, manual energy, derived from stored chemical energy, is converted into potential energy of the clock spring. In the second, this internal potential energy is converted into kinetic energy of the clock mechanism and is substantially dissipated as heat. However, while the unwinding of the spring can occur spontaneously, the winding-up cannot. Similarly, in many systems in nature, it is possible for a process to occur spontaneously in one direction but for the reverse process to occur only under the conscious application of an external agency. It is the role of the second law of thermodynamics to predict the direction in which such natural processes occur once they are allowed by the condition imposed by the first law.

It will be seen to be difficult to express the second law of thermodynamics in a simple manner while at the same time covering the generality of processes and conditions to which it must apply. (The difficulty of making a simple statement of this law is frequently the cause of much confusion!) First of all, the direction in which a process can occur is intimately involved with the concept of time and with the 'arrow' of time. It is accepted as a natural phenomenon that time is increasing in one direction only. Therefore, if we state that a process occurs spontaneously from an initial state A to a final state B, it is understood that the state B occurs at a later time than the state A. What we should like as scientists is some logical general statement, which can preferably be written in a mathematical form, that expresses the fact that spontaneous processes occur in one direction only, so that the final state occurs at a later time than the initial state!

In order to be able to obtain and to state an understandable form for the second law of thermodynamics in its full sense, it is necessary first to distinguish between two types of processes, that is, between *reversible* and *irreversible* processes.

For a system to undergo a reversible process during which the state of the system is changed in some manner, it is necessary that the sense in which the change is occurring can be reversed exactly at any point in the process. This will require that the change is occurring under the influence of forces which are only infinitesimally in disequilibrium and which are of such a nature that, if the sense of the disequilibrium is reversed, then the whole sequence of events constituting the change will likewise be reversed in time. The whole process which has occurred up to a given point can then be reversed and the steps retraced to the original state without leaving any change in the rest of the universe. For these conditions to be satisfied it is clear that, during the whole of such a process, the system will at all times be in an equilibrium state with relation to its surroundings and to its own motion, so that no irreversible effects such as gas turbulence and electrical eddy currents can be produced, or work done against frictional forces. Thus, for example, if the pressure applied externally to the fluid of Figure 4 is instantaneously increased to be appreciably greater than the pressure within the fluid, the piston will accelerate during the compression of the fluid and the resulting kinetic energy of the piston will be dissipated either as a turbulence in the gas, which is subsequently converted to heat, or as sound waves as the piston oscillates about some equilibrium position. The process in which the fluid is involved in such a case will not therefore be reversible. Similarly, in a reversible process, there must be no expenditure of energy in overcoming friction or viscosity, since work which is so expended during motion in one direction cannot be regained when the motion is reversed. Some irreversible processes are illustrated in Figure 6.

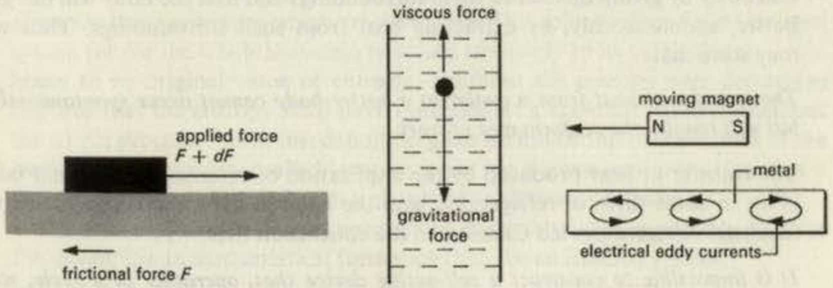


Figure 6 Examples of irreversible processes

The definition of a reversible process effectively ensures that no practical process can be exactly reversible. However, an ideal limiting system can be envisaged where the necessary equilibrium is maintained while some change in the state of the system occurs infinitely slowly, and there is no friction or viscosity, whether mechanical, electrical or magnetic. (Some practical systems may approximate closely to this ideal case.) Much consideration will be given to these ideally reversible processes.

Those processes which do not, even ideally, satisfy the conditions for reversibility are termed irreversible. Such a case, for example, is seen in the Joule paddle-wheel experiment, which contributed to the establishment of the first law of thermodynamics. Here a volume of water is heated as weights fall through a predetermined height, so causing a paddle to rotate against the viscosity of the water. The falling of the weights and the heating of the water occur spontaneously. However, no amount of heat subsequently applied to or extracted from the water will cause the weights to rise again to their original positions! Here the irreversibility of the turbulence of the water is the main contribution to the irreversible nature of the process.

In nature the processes which occur spontaneously are necessarily irreversible, although many do not obviously involve irreversible forces. Thus a plant may grow from a seed, wither and die but it cannot 'ungrow'. A small amount of coloured dye placed in a beaker of water will spread out and will become, in time, dispersed uniformly through the water. Since there is no violation of conservation of energy in either of these processes, it is possible to visualize another universe, still subject to the law of conservation of energy, in which plants reverse their growing sequence to return to seeds and a dye uniformly distributed in water gradually gathers itself into one small volume. However, this is not our universe, and our problem is to formulate some principle which will predict the observed behaviour in the universe which is familiar to us.

One well-known observation, which can be related in a convenient manner to the zeroth and first laws, has been taken as a basis for a statement of the required principle but describes the directionality in a rather restricted form. This observation is that a body placed in contact with cooler surroundings cools down spontaneously by giving up heat to these surroundings and that the body will not get hotter, spontaneously, by extracting heat from such surroundings. Thus we may state that:

The transfer of heat from a cooler to a hotter body cannot occur spontaneously, but will require the performance of work.

The transfer of heat produced by the application of external work would take place in some form of refrigerator, and the consideration of the operation of idealized refrigerators led Clausius to the conclusion that:

It is impossible to construct a self-acting device that, operating in a cycle, will produce no other effect than the transfer of heat from a cooler to a hotter body.

From the equivalence of heat energy and work energy, this statement may be shown to be equivalent – see Appendix A – to an alternative statement by Kelvin and Planck, namely:

It is impossible to construct an engine that, operating in a cycle, will produce no other effect than the extraction of heat from a reservoir and the performance of an equivalent amount of work.

These two statements are just different ways of expressing the *second law of thermodynamics* and, while they are exact in the context in which they are given, they are necessarily largely restricted to such a context.

4.2 Entropy

In order to obtain a more generally applicable formulation of this law which governs the direction of thermodynamic processes we may define a quantity, which is to be called the *entropy*, which necessarily increases during irreversible processes and which never decreases. For this quantity to be useful in describing the behaviour of a particular system it must necessarily be a function of the state of that system, so that whenever the system is in a particular state the entropy will have a particular value. If we can find some quantity which has this property, then the condition that this quantity increases will determine the direction in which any spontaneous process will occur, and it will be possible to state the second law of thermodynamics in the form:

In any thermodynamic process the entropy of the universe will increase or remain constant but will never diminish.

Three points regarding this 'law' may be noted. First, if the particular process occurs in an isolated system that system may be considered as its own universe. However, in most cases the systems will not be isolated from all their surroundings, and, in any process which occurs, the entropy of both the system itself and its surroundings will be altered. Strictly, therefore, in such cases the whole universe must be considered as a single system. Second, the condition that the entropy can remain constant during a thermodynamic change is necessary to allow for those processes which are reversible. This follows since for an isolated system (or for the whole universe) to return reversibly to its original state – and hence to its original value of entropy – without the entropy ever decreasing requires that the entropy shall have remained at a constant value throughout the whole process. Third, this definition gives no indication of the nature of the quantity which we have called 'entropy', nor has it given any indication that it will be possible even to find such a quantity for all types of systems.

If the entropy is, for convenience, given the symbol S it is possible to write the second law in mathematical form such that, for an isolated system,

$$\Delta S \geqslant 0 \qquad \textbf{4.1}$$

for any thermodynamic process. The inequality sign then applies to those processes which are irreversible and the equality sign to those which are reversible. The 'arrow' of time will then be specified by stating that if there are two states of the universe, or of an isolated system, the one with the higher entropy will occur at the later time. The entropy changes corresponding to the two cases of **4.1** are indicated in Figure 7 for an isolated system.

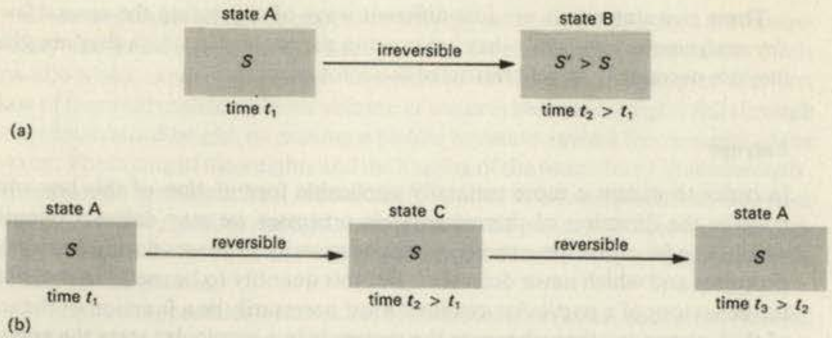


Figure 7 Changes in entropy with time for an isolated system

In order to advance usefully from this point it is necessary to obtain a specification for the entropy in terms of known thermodynamic quantities and, as we shall see, this may be partially achieved by the use of the classical statements of the second law which were given above.

4.3 Stages in the Carnot reversible cycle

The thermodynamic concept of entropy was first formulated from consideration of the efficiency of practical and ideal engines by a French engineer, Carnot, although the results were not expressed in terms of entropy until some time later. The essence of the discussion was an idealized reversible engine, postulated by Carnot, which may be described in the following manner which we have made as general as possible.

Consider that the working substance of an engine is connected by frictionless means to an external load on which it can perform work. The working substance may be, for example, a fluid, a rod of metal, a paramagnetic salt or a thermocouple, provided that, whatever substance is employed, the thermodynamic changes in which that substance is involved within the engine can occur reversibly. (The working substance is frequently taken to be an ideal gas but, although it will be seen in section 11.4 that the results obtained in this case are of particular interest, such a treatment is too restrictive for the present purpose.) There is considered to be available to the engine two heat reservoirs at different temperatures θ_1 and θ_2 with $\theta_1 > \theta_2$, the definition of the temperature difference being such that heat will flow from the higher temperature reservoir (the source) at θ_1 to the lower temperature reservoir (the sink) at θ_2.

The engine is now allowed to perform a cycle of completely reversible operations in which it does work on the external load and returns to its initial state. This cycle is performed in a series of four distinct stages.

(a) The engine is first placed in contact with the higher-temperature heat reservoir (or source), so that the working substance attains the temperature θ_1. The engine is then allowed to do work on the load as the state of the working substance changes reversibly at constant temperature θ_1, while heat Q_1 is taken in from the reservoir, as illustrated in Figure 8(a). The first stage of the cycle is thus *isothermal* and reversible.

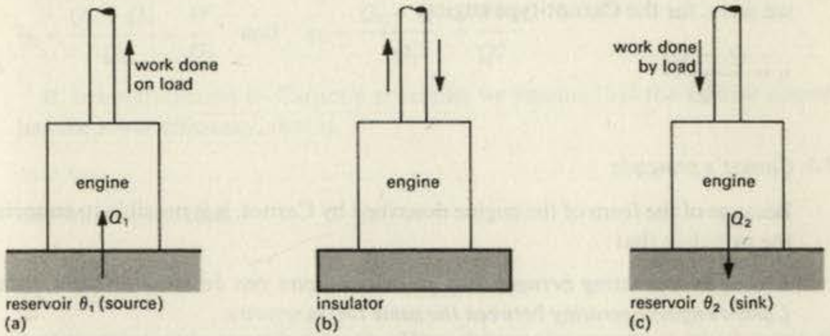


Figure 8 Stages in the Carnot reversible cycle

(b) The engine is now isolated from its surroundings by means of some insulator, as indicated in Figure 8(b), while it is allowed to do work on the load, so that the internal energy of the working substance decreases and its temperature decreases to the value of the second heat reservoir (or sink) at θ_2. This process is thus *adiabatic* and reversible.

(c) The engine at temperature θ_2 is placed in contact with the sink and, as indicated in Figure 8(c), the load is now allowed to perform work on the engine. The load thus returns towards its original position, while the working substance is held at the constant temperature θ_2. This stage of the cycle is carried out reversibly and isothermally while heat Q_2 is given to the sink and the process continues until a state is reached such that

(d) when the engine is insulated from its surroundings again and the load is allowed to perform further work, adiabatically and reversibly, on the engine, the load returns to its original position and the engine returns to its original state at temperature θ_1.

By the first law of thermodynamics, the work done by the engine during the above cycle of events must be equal to the difference between the heat Q_1 given to the engine by the source and the heat Q_2 absorbed from the engine by the sink, since the internal energy will return to its original value and there is no gain or loss of heat during the adiabatic processes. Hence

$$W = Q_1 - Q_2. \tag{4.2}$$

If the thermodynamic efficiency η of an engine (as opposed to any mechanical efficiency) is defined, in terms of the work done W and the heat Q_1 which must be taken in at the higher temperature in order to obtain this work, as

$$\eta = \frac{W}{Q_1}, \qquad 4.3$$

we have, for the Carnot-type engine,

$$\eta = \frac{Q_1 - Q_2}{Q_1}. \qquad 4.4$$

4.3.1 *Carnot's principle*

Because of the form of the engine described by Carnot, it is possible to enunciate the principle that:

No engine operating between two given reservoirs can be more efficient than a Carnot engine operating between the same two reservoirs.

This statement follows from, and can be shown to be equivalent to, the Kelvin–Planck and Clausius statements of the second law of thermodynamics. Consider that there is an irreversible engine E_I which drives a Carnot engine E_C in reverse, while both operate with the same source and sink. Suppose also that the engine E_I takes heat Q_1 from the source while the work W which it does during the cycle drives the engine E_C to give heat Q' back to this source, as indicated in Figure 9.

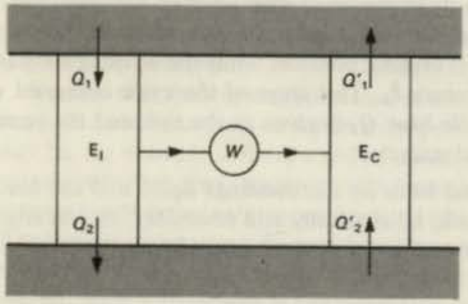


Figure 9 Carnot engine E_C driven in reverse by engine E_I. (It is not strictly important whether E_I is reversible or not, since only the Carnot engine is driven in reverse. It is important, however, that the heat exchange by E_I only occurs with the same source and the sink)

E_I will give heat Q_2 to the sink such that, by equation **4.2**,

$$Q_2 = Q_1 - W$$

and similarly E_C will take heat Q'_2 from the sink given by

$$Q'_2 = Q'_1 - W.$$

The efficiencies of the two engines E_I and E_C may be written as

$$\eta_I = \frac{Q_1 - Q_2}{Q_1} = \frac{W}{Q_1} \quad \text{and} \quad \eta_C = \frac{Q'_1 - Q'_2}{Q'_1} = \frac{W}{Q'_1}.$$

If, in contradiction to Carnot's principle, we assume that the Carnot engine has the lower efficiency, that is,

$$\eta_I > \eta_C, \qquad 4.5$$

this will require that

$$\frac{W}{Q_1} > \frac{W}{Q'_1},$$

and so Q'_1 must be greater than Q_1. Thus, during the process in which work W is done by E_I on E_C without involving any external agency, the source will become hotter by taking in heat $Q'_1 - Q_1$, which is exactly equal to the heat $Q'_2 - Q_2$ lost by the source. This result clearly contravenes the Clausius statement of the second law of thermodynamics, and hence the assumption **4.5** must be invalid. Thus Carnot's principle follows directly as a result of this statement, and we must write

$$\eta_I \ngtr \eta_C.$$

If the engine E_I were replaced by another Carnot reversible engine $E_{C'}$, it would be possible to reverse the direction of operation and thus to consider the inverse of the relation **4.5**. Then we could show by the same logic that

$$\eta_C \ngtr \eta_{C'}$$

as well as

$$\eta_{C'} \ngtr \eta_C.$$

The only way in which both these relations may be satisfied is that

$$\eta_C = \eta_{C'}$$

when the two Carnot engines are working between the same temperatures.

One consequence of Carnot's principle may be seen if it is realized that (a) all Carnot engines which are working between the same temperatures will necessarily have the same efficiency; (b) by virtue of the processes which make them irreversible, irreversible engines working between the same temperatures will be less efficient than the corresponding Carnot engines. Thus if we can determine these efficiencies it may be possible to obtain a measure of the irreversibility, and therefore to give a specification for the entropy. First of all, however, it is necessary to define, by means of the Carnot cycle, a universal scale of temperature on which the temperatures of the source and sink may be measured.

4.4 A thermodynamic scale of temperature

Because the efficiency of all Carnot engines operating between the same two reservoirs is the same, independent of their form or of their working substances, it follows that this efficiency can only be a function of the temperatures of the source and the sink, since these temperatures are the only constant parameters in the description of the engines. We thus have a basis for defining a temperature scale independent of any property of any particular material, that is, a universal or absolute scale. For this purpose we define the temperatures of the source and sink of a given Carnot engine on an absolute temperature scale as T_1 and T_2 respectively, such that the efficiency of the Carnot engine operating between these temperatures is

$$\eta = \frac{T_1 - T_2}{T_1}. \qquad 4.6$$

By comparison with **4.4** we can see that this implies that

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

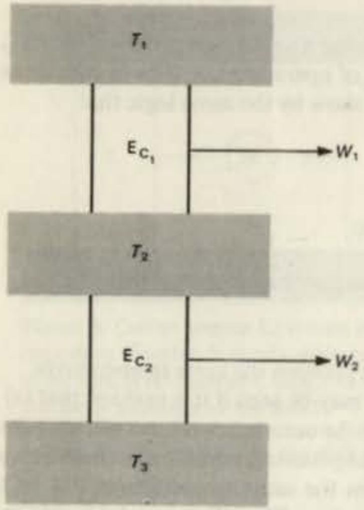


Figure 10 Two Carnot engines operating with the sink of one as the source of the other

or $$\frac{T_2}{T_1} = \frac{Q_2}{Q_1}, \qquad 4.7$$

so that the temperatures of the two reservoirs, in this definition, are in the ratio of the heats which the engine absorbed from one and rejected to the other.

That the definition of **4.6** gives a logical choice of a temperature scale may be seen by supposing that two Carnot engines operate as shown in Figure 10, where the sink of one engine acts as the source of the other. The heat Q_2 given out by the engine E_{C_1} at temperature T_2 is absorbed by E_{C_2} at the same temperature. Then, by the definition of **4.7**,

$$\frac{T_1}{T_2} = \frac{Q_1}{Q_2} \quad (\text{for } E_{C_1}) \quad \text{and} \quad \frac{T_2}{T_3} = \frac{Q_2}{Q_3} \quad (\text{for } E_{C_2}),$$

so that, whatever the value of the intermediate temperature T_2,

$$\frac{T_1}{T_3} = \frac{Q_1}{Q_3},$$

which is the relation which would apply for a single Carnot engine operating between temperatures T_1 and T_3. Thus, by taking a complete series of Carnot cycles, any range of temperatures may be specified in a self-consistent manner.

The temperatures defined by **4.6** and **4.7** will obviously contain an arbitrary constant. In practice this constant will be evaluated by taking some fixed temperature as a reference point for the whole temperature scale. This fixed point is normally taken at the triple point of water with the value 273·16 K (degrees Kelvin, or simply kelvins). Then it is possible to refer all other temperatures to this value since there will be an absolute zero of temperature, 0 K, such that, as the temperature of the sink of a Carnot engine approaches this zero of temperature, the value of Q_2, the heat rejected by the engine, approaches asymptotically to zero. With the fixed points chosen on this basis, the interval of temperature between the melting point of pure ice under standard atmospheric pressure and the boiling point of pure water under standard atmospheric pressure is a hundred degrees on the Kelvin scale, and with $T_3 = 273{\cdot}16$ K, $T_1 = 273{\cdot}16\, Q_1/Q_3$ K.

The fact that there are thus established a hundred degrees in the so-called 'fundamental interval' of temperature between the ice point and the steam point does not mean that the Kelvin scale, or thermodynamic scale, is equivalent to any other centigrade scale. This will be clear in the discussion of Chapter 19 where it will be seen that, when the practical measurement of a temperature depends on some particular property of some particular substance this practical temperature must be reduced to an absolute temperature by calibration against the Kelvin scale.

The establishment of the hundred steps or degrees necessary between the ice point and the steam point by use of the Carnot engine alone is theoretically achieved by establishing a sequence of a hundred such Carnot engines, as shown

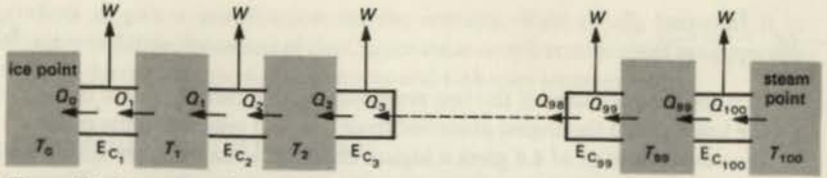


Figure 11 Sequence of a hundred Carnot engines, the sink of one acting as the source for its neighbour. Each engine performs the same amount of work and takes in the heat given out by the one above in the temperature sequence

in Figure 11. With each engine taking in the heat given out by its neighbour and each performing work W, it is clear that

$$W = Q_{100} - Q_{99} = Q_{99} - Q_{98} = \ldots = Q_1 - Q_0. \qquad \textbf{4.8}$$

It is clear also that, since the temperatures are defined in **4.6** to be in direct ratio to the heats taken in and given out, **4.8** implies that

$$T_{100} - T_{99} = T_{99} - T_{98} = \ldots = T_1 - T_0.$$

There are thus a hundred equal intervals between the ice point and the steam point, and each of these may be taken to be equal to one kelvin or, on the scale with the ice point at 0 °C and the steam point at 100 °C, to 1 °C.

4.5 The quantity Q/T

In the above discussion of the Carnot engine it was seen that the engine takes in heat Q_1 from the reservoir at temperature T_1 and gives out heat Q_2 to the sink at temperature T_2, the temperatures being effectively defined by the values of Q_1 and Q_2. We now distinguish between high-quality heat (i.e. heat at a high temperature) and low-quality heat (i.e. heat at a low temperature) by defining a quantity σ which is the ratio of the quantity of heat to the temperature at which that heat is available. Then, for the heat Q_1 taken from the source, this quantity will have the value

$$\sigma_1 = \frac{Q_1}{T_1},$$

while for the heat Q_2 delivered to the sink it will be

$$\sigma_2 = \frac{Q_2}{T_2}.$$

Over the whole Carnot cycle there will be a change in the quantity σ for the system consisting of the engine, the source and the sink by an amount

$$\Delta\sigma = \sigma_2 - \sigma_1, \qquad \textbf{4.9}$$

since σ_1 has been taken from the source and σ_2 delivered to the sink while the working substance returns to its original state and therefore suffers no change in σ. (There is, of course, no change in σ over the adiabatic processes where there is no exchange of heat.) Taking the values given for σ_1 and σ_2 we have

$$\Delta\sigma = \frac{Q_2}{T_2} - \frac{Q_1}{T_1}. \qquad \textbf{4.10}$$

However, since the definition of our absolute temperature is chosen to give $Q_1/Q_2 = T_1/T_2$ (see **4.7**), so that $Q_2/T_2 = Q_1/T_1$, the change in σ over the complete Carnot cycle is

$$\Delta\sigma = 0. \qquad \textbf{4.11}$$

Not only does this argument give $\Delta\sigma = 0$ for the complete cycle but, if any single part of the reversible cycle is considered, the same result is seen to be obtained. For example, when during the first stage of the cycle the engine takes in heat Q_1 from the source of temperature T_1, there is a decrease in the quantity σ of the source by an amount $-Q_1/T_1$ while the working substance of the engine gains an amount $+Q_1/T_1$ and so, over this stage, $\Delta\sigma = 0$ for the complete system of engine and source, although it is not zero for the source separately nor for the working substance separately. Similarly, as explained above, $\Delta\sigma$ is zero over the adiabatic stages.

The above argument may be extended to any reversible process, since it will always be found that any such process can be divided into sequences of infinitesimal reversible isothermal and adiabatic processes, each of which must separately satisfy the condition $d\sigma = 0$ for the complete system. This condition is normally expressed in terms of the sum of increments such that, if a part of the system gains a heat dQ while it is at temperature T, the change in σ over a reversible process is

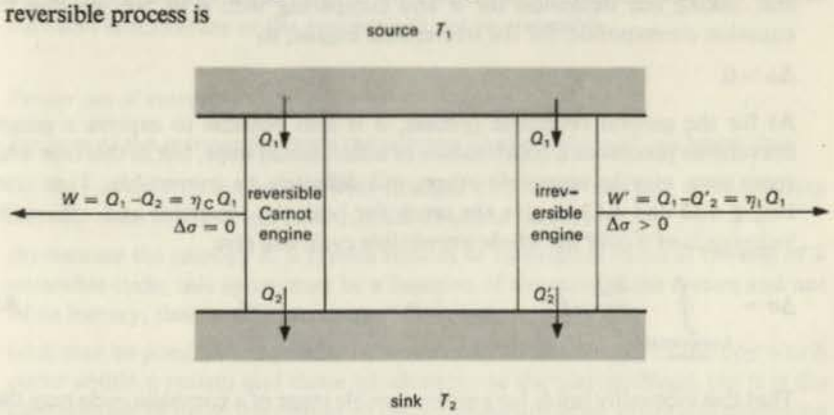


Figure 12 Reversible and irreversible engines between the same reservoirs

$$\Delta\sigma = \int \frac{dQ}{T} = 0, \tag{4.12}$$

where the changes in σ are considered for all parts of the system (and for the universe if the system is not isolated).

It is immediately clear from the above analysis that the quantity $\Delta\sigma$ satisfies the condition that we have required for the change in entropy ΔS for any reversible process, namely that $\Delta S = 0$, as in **4.1**. This result gives encouragement therefore to examine the behaviour of this quantity σ for the case of an irreversible process.

Let us suppose that, as expected from Carnot's principle, the irreversible engine shown in Figure 12 has an efficiency η_I which is, through the very processes which make it irreversible, less than η_C, the efficiency of the corresponding Carnot engine. Then, if heat Q_2' is given to the sink by the irreversible engine for heat Q_1 taken in from the source,

$$\eta_I = \frac{Q_1 - Q_2'}{Q_1}, \qquad \eta_C = \frac{Q_1 - Q_2}{Q_1} = \frac{T_1 - T_2}{T_1}$$

and hence, for $\eta_I < \eta_C$,

$$\frac{Q_1 - Q_2'}{Q_1} < \frac{T_1 - T_2}{T_1}.$$

Thus, on rearranging, we see that

$$\frac{Q_2'}{T_2} - \frac{Q_1}{T_1} > 0 \tag{4.13}$$

and, taking our definition for σ and comparing with **4.10**, we see that this equation corresponds, for the irreversible engine, to

$$\Delta\sigma > 0.$$

As for the general reversible process, it is also possible to express a general irreversible process as a combination of infinitesimal steps, but in this case while some steps may be reversible others will definitely be irreversible. Then combining **4.12** and **4.13** to give the result for both the reversible and irreversible increments of σ over the whole irreversible cycle will give

$$\Delta\sigma = \oint_{\substack{\text{irreversible}\\ \text{cycle}}} \frac{dQ}{T} > 0. \tag{4.14}$$

That this inequality holds for any irreversible stage of a complete cycle may then be demonstrated by considering that stage to be but a single part of an otherwise reversible cycle. Then, since $\Delta\sigma > 0$ for the whole cycle and $\Delta\sigma = 0$ for the part that is reversible, it is necessary that $\Delta\sigma > 0$ for the irreversible part. **4.14** may thus be generalized so that

$$\Delta\sigma = \int_{\substack{\text{irreversible}\\ \text{process}}} \frac{dQ}{T} > 0.$$

Also, although the value of σ of the substance which is undergoing the cyclic process may increase or decrease during different parts of a cycle, the value of σ will be unchanged at the completion of the cycle because the substance will return to its original state. Any total increase in $\Delta\sigma$ over the whole cycle must therefore occur in the system as a whole and is usually expressed as an increase in σ for the *universe*, since this will cover all interaction between the working substance and its surroundings.

It is thus seen that the quantity σ has all the properties required of the entropy by **4.1**. We therefore make the transfer by defining a change in entropy such that:

If a system or substance at temperature T increases its heat content by an amount dQ there is an increase in the entropy of the system or substance of

$$dS = \frac{dQ}{T} \tag{4.15}$$

and, over any complete process, the change in entropy of the system is the integral of dS over that process, so that

$$\Delta S = \int_{\text{process}} \frac{dQ}{T} \tag{4.16}$$

The condition, that in a reversible process the total change in entropy of the universe as a whole is zero, is satisfied because for every gain by a system of heat dQ at temperature T some other system must lose the same amount of heat at the same temperature or the process will not be reversible.

4.6 Properties of entropy

Because of the manner in which the entropy has been defined, we know that:

(a) we can only be concerned with changes in entropy and not with absolute entropy since no basis has been laid for absolute values;

(b) because the entropy of a system returns to its original value at the end of a reversible cycle, this value must be a function of the state of the system and not of its history, that is, dS is an exact differential.

(c) it may be possible to consider independently those changes of entropy which occur within a system and those which occur in the surroundings, but it is the combination of these, the change in the complete universe, which will determine whether the process is reversible or not;

(d) in a *reversible*, *adiabatic* process there is no change of heat content of any system at any point and hence

$$\Delta S = \int \frac{dQ}{T}$$

is identically zero and the process is therefore referred to as *isentropic*. (In some discussions adiabatic processes are frequently stated to take place *rapidly* in order to ensure that $dQ = 0$. It would not be expected that such a process would be reversible since we require that, to be reversible, the changes should follow equilibrium states. Hence we expect there to be an increase in entropy in such an adiabatic process performed rapidly and the distinction must be made between isentropic and non-isentropic adiabatic processes.)

On the basis of these criteria and definitions, and equations **4.15** and **4.16**, it is possible to derive a number of simple, though important, relationships involving the entropy.

4.6.1 *The temperature–entropy diagram*

Any process in which a system can exchange heat with its surroundings may be represented on a temperature–entropy diagram by representing any point in the process by the value of temperature and entropy at that point. If the process is reversible, any change in the heat content may be represented by writing $dQ = T\,dS$. (This relation cannot be assumed for an irreversible process since, in such cases, there may be entropy changes associated with processes other than the addition of heat to a system.) Since only changes in entropy are considered, the actual value of entropy will be taken relative to some arbitrary zero.

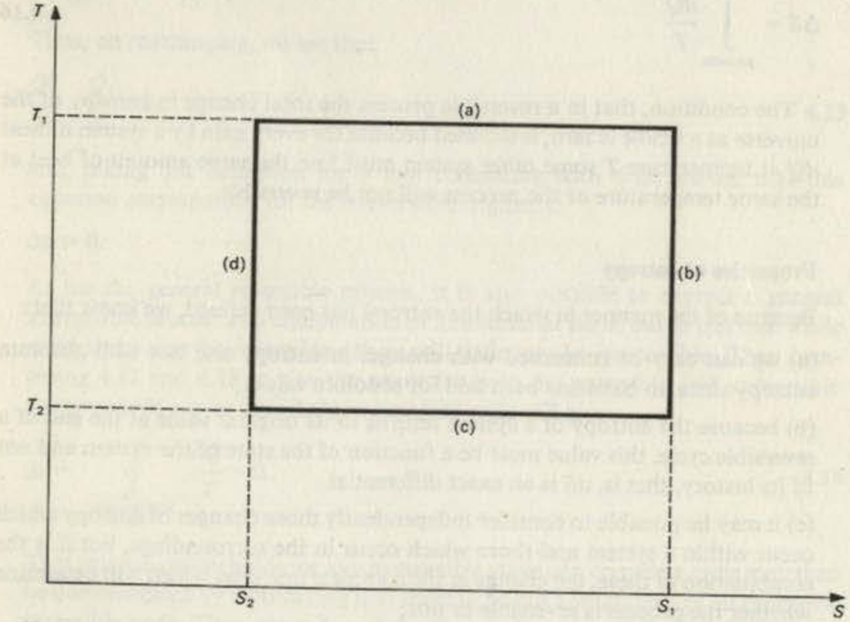


Figure 13 *T–S* diagram of the Carnot cycle

Thus, for example, the Carnot cycle may be represented by two processes (the isothermal processes) at constant temperature (T_1 and T_2) and two (the adiabatic processes) at constant entropy (S_1 and S_2). The form that this representation of the cycle will take is shown in Figure 13 where (a) and (c) represent the isothermal processes and (b) and (d) represent the adiabatic processes. The heat taken from the source along the stage (a) is clearly

$$T_1(S_1 - S_2) = Q_1,$$

while the heat given to the sink along (c) is

$$T_2(S_1 - S_2) = Q_2.$$

The work done during the cycle is, as before,

$$W = Q_1 - Q_2 = (T_1 - T_2)(S_1 - S_2),$$

which is the area of the cycle on the *T–S* diagram. (Not surprisingly, because of our original definition, we see that the thermodynamic efficiency η is

$$\frac{Q_1 - Q_2}{Q_1} = \frac{T_1 - T_2}{T_1}.)$$

4.6.2 *Available work*

When a system undergoes an irreversible change, which necessarily involves an increase in the entropy of the universe, there will be a decrease in the total work which is available from any energy sources involved. This may be illustrated very simply by considering a system consisting solely of two vessels each containing an identical mass of liquid but at two different temperatures. Useful work could be obtained from the masses of liquid in these vessels if they were used respectively as the source and sink of a Carnot engine. However, if the two volumes of liquid are first mixed together so that they become a single volume with double the mass and with a single temperature, they can no longer be so used. The useful work which was originally available and which has been lost by the mixing process is directly proportional to the gain in the entropy, ΔS, which occurs during the irreversible mixing. In fact, as shown in Appendix B, if the lowest temperature sink which is accessible to the system is at temperature T_0, the loss of available work is $T_0\,\Delta S$.

On the other hand, if the two volumes of liquid had been brought to the same temperature reversibly there would be no gain in entropy and the maximum amount of work would be extracted from them.

The fact that an increase in entropy is associated with a decrease in available work indicates that as the entropy of the universe increases with time so there is a decrease in the useful activity available from the original energy sources. We thus see the increase in entropy as involving at the same time the 'running down' of the universe.

4.6.3 *Irreversible changes in systems*

We have seen from **4.1**, which we used as a definition for a change in entropy, that the condition $\Delta S > 0$ must be satisfied for an irreversible process which involves the exchange of heat. However, there are irreversible processes in which no exchange of energy occurs, whether in the form of heat or otherwise. Such processes usually involve a change in the form or arrangement of the system, as will happen for example when one pure gas diffuses into another pure gas to form a mixture of gases. Similarly, but in a different system, there will be a disordering, or change in the arrangement, of the atomic magnetic moments when a magnetized paramagnetic material (thermally isolated from its surroundings) becomes demagnetized on the sudden removal of the magnetic field. The processes are certainly irreversible, there being no question that they proceed under equilibrium conditions, while there is equally no change of heat content of the system. In each of these processes there will be a gain in entropy although

$$\int \frac{dQ}{T} = 0$$

since $dQ = 0$ identically. It is clear that for processes of this type we must write

$$\Delta S > \int \frac{dQ}{T} \tag{4.17}$$

if we wish the entropy to increase for all irreversible processes.

It is fairly easy to account for this additional entropy (which obviously does not arise in these cases from such effects as friction, eddy currents and so forth, which would involve exchange between different forms of energy) by the method outlined in the following discussion. However, we must note at once that, because of this result, the calculation of changes in entropy according to equation **4.16** is only strictly permissible for processes which are reversible. For this reason it will be observed that in subsequent calculations the change in the entropy of a system will be calculated, where possible, as if the system had followed a reversible path between the initial and final state, even if it has not. Then, since the entropy of the initial and final states will be a function of those states alone, the difference between the two entropies will be the same when calculated for a reversible path as for any other sequence of changes between the two states. Inevitably, though, there will be seen to be cases where no reversible path is in fact possible.

Consider an example. Two beakers each contain a mass m of a liquid which has temperature-independent specific heat c. The one beaker is at temperature T_1, the other at temperature T_2. The two masses of liquid are added together in a third beaker (which is insulated from its surroundings and has zero thermal capacity) to give a total mass $2m$ at temperature $\frac{1}{2}(T_1 + T_2)$. Because the process is irreversible there will be a gain in the total entropy of the universe. However, in order to calculate the change in entropy we must use the fact that the entropy of the final state is independent of the path by which this final state is attained. We thus compute the change in entropy of the two masses of liquid as if they had followed reversible paths to the temperature $\frac{1}{2}(T_1 + T_2)$, so that it is possible to write the change in entropy as in **4.16**. Such reversible paths would necessarily consist of a series of equilibrium stages. The reversible heating (or cooling) of the two beakers would therefore require a series of heat reservoirs, each at an infinitesimally lower temperature than its neighbour. However, since these are only required for a 'thought' experiment, they present us with no problem.

Putting $dQ = mc\,dT$ (by the definition of the specific heat), the change in entropy of the liquid which starts at temperature T_1 is

$$\Delta S(\text{liquid at } T_1) = \int \frac{dQ}{T} = \int_{T_1}^{(T_1+T_2)/2} \frac{mc\,dT}{T} = mc \ln \frac{T_1 + T_2}{2T_1},$$

while the change in entropy of the liquid which starts at T_2 is

$$\Delta S(\text{liquid at } T_2) = \int_{T_2}^{(T_1+T_2)/2} \frac{mc\,dT}{T} = mc \ln \frac{T_1 + T_2}{2T_2}.$$

If the two volumes of liquid are now added together in the original isolated beaker to form a single volume, we shall have reached the same state as in the case of direct mixing. The process of the mixing will now be entirely reversible and involve no change in entropy, and so the total change in entropy of the liquid will be

$$\Delta S(\text{total}) = mc\left[\ln \frac{T_1 + T_2}{2T_1} + \ln \frac{T_1 + T_2}{2T_2}\right]$$
$$= mc \ln \frac{(T_1 + T_2)^2}{4T_1 T_2},$$

which is necessarily positive since

$$(T_1 + T_2)^2 > 4T_1 T_2 \quad \text{for } T_1 \neq T_2.$$

In the reversible processes which we have envisaged there will be no change in the universal entropy, since our imaginary reservoirs will have lost an amount of entropy equivalent to ΔS(total). In the case of the direct mixing no external systems are involved and the entropy gain is an absolute gain for the whole universe.

If the two liquids contained by the beakers had been of different chemical composition and entirely immiscible, so that they could exchange heat but could not mix, this type of calculation would still be valid, since the 'mixing' process would still be reversible. However, if the liquids had been different but had been miscible, it would not have been possible to separate the liquids by any simple process after they had been mixed together and the entropy change would have been greater. (For a discussion of such mixing see section 11.5.)

4.7 Entropy, probability and disorder

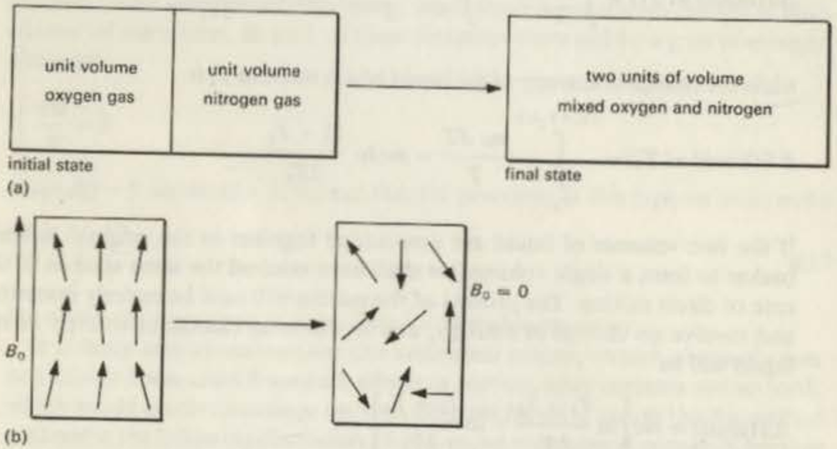


Figure 14 Increase in disorder when (a) two different gases at the same temperature and pressure are mixed together, and (b) when a magnetic field is removed from a paramagnetic specimen

We have already mentioned that such processes as the mixing together of two different gases or the demagnetizing of a paramagnetic material will result in a gain in entropy, not because of an exchange of heat but because of the very irreversibility of the changes which occur. The property which these processes have in common is that they involve a spontaneous change from a more ordered arrangement to a less ordered arrangement of the components of the particular system. Thus, in Figure 14(a) the initial state of the separated gases, indicated for convenience as oxygen and nitrogen, is more ordered than the final mixture. That the process is irreversible is obvious if it is considered that the mixing process follows spontaneously on the removal of the partition between the gases but that the replacement of the partition will not cause the gases to separate out from the mixture. Indeed, quite a complicated procedure would be required to resegregate the oxygen and nitrogen molecules. Similarly, in Figure 14(b), if the field B_0 which causes the magnetization is instantaneously removed, the atomic magnetic moments will instantaneously be partially aligned with no field applied. The magnetization which results from this alignment will disappear with time as the system tends spontaneously to a state where these moments are arranged in a random manner. Again the direction of the spontaneous change from the magnetized to the demagnetized state will result in a more disordered state replacing the ordered state of the magnetized specimen. Since the transfer from the initial state to the final state in both of the cases illustrated in Figure 14 is irreversible, there is clearly a gain of entropy in each case. We may therefore conclude that a disordered state has a higher entropy than an ordered state. (Although in both the cases illustrated here the final state is that of higher entropy, in most cases both the energy and the entropy of the system have to be considered in order to determine the final state – see section 6.2.)

If there is to be any exact relation established between the disorder and the entropy of a system, it is clearly necessary that the disorder shall be expressible in some way in terms of either the macroscopic or microscopic parameters of the system. The way in which such an expression can be obtained is seen most simply if it is noted that a disordered state of a system can be achieved in more ways (or by more arrangements of the components) than can an ordered state. Thus, in the final (or mixed) state of Figure 14(a), where both gases have spread through twice their original volumes, each molecule can be found in twice as many positions as were possible in the initial state where the gases were confined by the partition, each to one half of the complete enclosure.

Consider that the unit volume of oxygen contains N molecules and that these can be arranged in the limited volume in $\Omega_i(O_2)$ distinct ways. When the partition is removed, each molecule can take up two positions for every one in the initial state and so the number of arrangements will increase by a factor 2^N to

$$\Omega_f(O_2) = \Omega_i(O_2) \times 2^N.$$

Similarly, if $\Omega_i(N_2)$ is the initial number of ways of arranging the nitrogen molecules in one half of the enclosure, there will be

$$\Omega_f(N_2) = \Omega_i(N_2) \times 2^N$$

ways of arranging the N nitrogen molecules once the partition has been removed. Consequently, if the total number of arrangements possible in the initial and final states of the whole gas system are $\Omega(\text{initial})$ and $\Omega(\text{final})$, then

$$\Omega(\text{initial}) = \Omega_i(O_2) \times \Omega_i(N_2)$$

$$\text{and} \quad \Omega(\text{final}) = \Omega_f(O_2) \times \Omega_f(N_2) = \Omega_i(O_2) \times 2^N \times \Omega_i(N_2) \times 2^N,$$

since each arrangement of the oxygen molecules may be taken together with each arrangement of the nitrogen molecules. Thus

$$\frac{\Omega\,(\text{final})}{\Omega\,(\text{initial})} = 2^{2N}. \qquad 4.18$$

Since N will be of the order of 10^{22} for a litre of gas at standard temperature and pressure, this ratio is considerable and the number of ways of arranging the molecules with the partition removed is considerably greater than the number of arrangements available before this removal. Consequently, once the partition has been removed, the molecules moving in a random manner will have a far greater probability of being found in the completely mixed state than in the segregated state. The irreversible mixing of the gases associated with the increase in entropy is therefore also associated with an increase in the likelihood of finding the mixed state relative to that of finding the unmixed state once the partition has been removed.

A similar argument applied to the magnetized specimen of Figure 14(b) would show that there is only one way of finding the magnetic moments completely aligned, but many ways of finding them orientated at random. The final disordered state is, in both cases, the most probable state and it appears therefore that there is a relation between high probability and high entropy.

If we now consider that heat energy, involving as it does a random motion of molecules in a gas or of ionic vibrations in a solid or of electromagnetic radiation in space, is a disordered form of energy, it is possible to associate the increase in entropy which occurs with the addition of heat to a body with an increase in disorder. However, the manner in which the disorder due to the addition of heat energy can be expressed in terms of arrangements of the molecules or ions is beyond the scope of our present text.

Suppose that for some system the number of distinct arrangements of its components, which we shall call the statistical weight of the system, is Ω. The relation between the entropy and the disorder can then be envisaged in the form

$$S = f(\Omega). \tag{4.19}$$

The actual form of the function $f(\Omega)$ may be simply established by considering a composite system composed of two separate systems A and B which do not interact in any way, as indicated in Figure 15. Then the entropy of the composite system will be

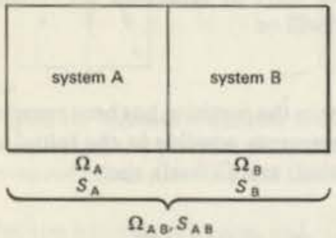


Figure 15 Two non-interacting systems A and B

$$S_{AB} = S_A + S_B, \tag{4.20}$$

where S_A and S_B are the individual entropies of A and B. At the same time the total weight of the combination will be given by the product of the weights Ω_A and Ω_B of A and B respectively, i.e.

$$\Omega_{AB} = \Omega_A \times \Omega_B, \tag{4.21}$$

since every one of the Ω_A possible arrangements of the system A may be taken together with every one of the arrangements Ω_B. For equations **4.20** and **4.21** to be consistent with each other the relation between S and Ω must be of the form

$$S = k \ln \Omega, \tag{4.22}$$

where, after the originator of this expression, the constant k is called Boltzmann's constant. Then

$$S_{AB} = k \ln \Omega_{AB} = k \ln \Omega_A + k \ln \Omega_B = S_A + S_B.$$

This problem of the relation between entropy and the weight of a system will be discussed further in subsequent sections of the text. However, it is interesting at this point to note that, by virtue of **4.22**, the result of mixing the two equal volumes of the two gases shown in Figure 14(a) (the gases both being at the same temperature and pressure) is to increase the entropy by the amount

$$\Delta S = S_{\text{final}} - S_{\text{initial}} = k \ln \Omega(\text{final}) - k \ln \Omega(\text{initial}),$$

which, on substituting from **4.18**, gives the entropy of mixing as

$$\Delta S = 2Nk \ln 2. \tag{4.23}$$

(The value of Boltzmann's constant which is required to give the same change in entropy for a reversible process from both the equations

$$dS = \frac{dQ}{T} \quad \text{and} \quad dS = k\, d(\ln \Omega)$$

$$k = 1{\cdot}38 \times 10^{-23}\ \text{J K}^{-1}.$$

This constant will occur later in our treatment of the kinetic theory of gases.)

4.8 The combined first and second laws

Because of the introduction of the concept of entropy it is possible to rewrite the mathematical form of the first law of thermodynamics for a *reversible* process with dQ replaced by $T\,dS$. Then **3.3** becomes, for a system at temperature T,

$$T\,dS = dU + dW. \tag{4.24}$$

The advantage of this expression compared with the original is that, while dQ is an inexact differential, dS is an exact differential. Then, because dU is an exact differential and dW may also be expressed in terms of an exact differential (which depends on the particular system and its constraints, as is seen from Table 1), it is possible for mathematical relationships to be established between the various parameters of the system in question, as will be discussed in Chapter 6.

For irreversible processes, the combined law gives

$$T\,dS > dU + dW,$$

but, as is very clear, such an expression has more physical significance than mathematical application and, as already explained, changes in entropy will be calculated, using **4.24**, for reversible paths between the initial and final states.

The substitution $T\,dS = dQ$ is often used to obtain an expression for the specific heat of a material. The thermal capacity of a body is defined as dQ/dT and the specific heat (strictly, the specific thermal capacity) is the thermal capacity per unit mass (denoted c) or per mole (denoted C). The value of the specific heat will depend on the conditions under which the heat dQ is supplied and the so-called principal specific heats C_p and C_v are the values with pressure and volume respectively maintained constant. These specific heats may be expressed in terms of changes in entropy, since $dQ/dT = T\,dS/dT$, as

$$C_p = T\left(\frac{\partial S}{\partial T}\right)_p \quad \text{and} \quad C_v = T\left(\frac{\partial S}{\partial T}\right)_v. \qquad \mathbf{4.25}$$

If the volume v of a fluid is maintained constant, then $dW = p\,dv = 0$, and so it follows from **4.24** that

$$C_v = \left(\frac{\partial U}{\partial T}\right)_v. \qquad \mathbf{4.26}$$

The usefulness of these expressions will be seen in Chapter 6.

4.9 Carathéodory's principle

The method by which the entropy has been introduced so far has required an appreciation of physical method and procedure. It is, however, possible to introduce the entropy in a more formal mathematical manner, of which we shall give here only the briefest indication, and that for a particular case.

Consider that, by the first law of thermodynamics, the heat supplied to a dielectric system is given by

$$dQ = C_v\,dT + p\,dv - E\,dP, \qquad \mathbf{4.27}$$

where work is performed *by* the system as it expands against the pressure and *on* the system by the electric field as the polarization P increases. The substitution $dU = C_v\,dT$ has been made from **4.26**. Consider further that the system is in the state X defined by a given set of values of T, v and P as indicated in Figure 16.

It is known that the states which are accessible from X by an adiabatic process, i.e. by a process in which $dQ = 0$, are limited to a surface in (P, T, v) space, while

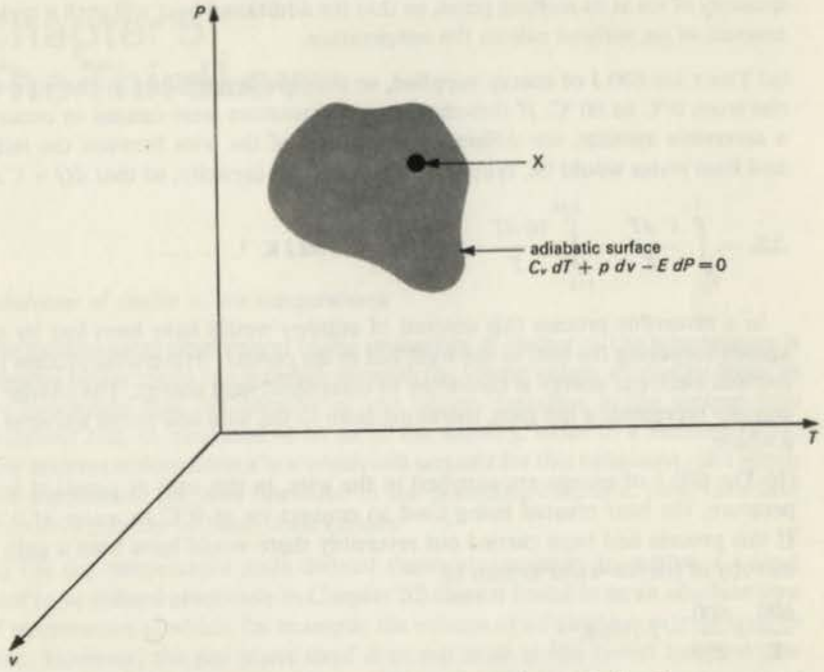


Figure 16 State X of system represented in the coordinates P, T and v

any state lying off this surface is not accessible from X by such a process. It is Carathéodory's principle that, for such inaccessible states to exist, it must be possible to write **4.27** in the form

$$dQ = C_v\,dT + p\,dv - E\,dP = Y\,dy,$$

where Y and y are both functions of T, v and P, while dy must be a perfect differential. It is then possible to identify Y with T and y with a function of state S, the entropy, so that

$$dQ = T\,dS$$

and the heat change is expressed in terms of a perfect differential.

4.10 Calculation of change in entropy

As an example of change in entropy, consider that heat is supplied to a wire of resistance ten ohms by a current of one ampere flowing for one minute. Consider two cases. (a) That the wire is isolated from its surroundings and, having a temperature-independent thermal capacity of 10 J K^{-1}, increases in temperature from its initial state at 0 °C. (b) That the wire is in thermal contact with a large quantity of ice at its melting point, so that the additional heat will melt a certain amount of ice without raising the temperature.

(a) There are 600 J of energy supplied, so that the temperature of the wire will rise from 0 °C to 60 °C. If this change in temperature were caused to occur in a reversible manner, the difference in entropy of the wire between the initial and final states would be, taking C as the thermal capacity, so that $dQ = C\,dT$,

$$\Delta S_a = \int_{T_i}^{T_f} \frac{C\,dT}{T} = \int_{273}^{333} \frac{10\,dT}{T} = 10 \ln \frac{333}{273} = 1{\cdot}98\,\mathrm{J\,K^{-1}}.$$

In a reversible process this amount of entropy would have been lost by the agency supplying the heat to the wire, but in the present, irreversible process the ordered electrical energy is converted to disordered heat energy. The change in entropy represents a net gain, therefore, both to the wire and to the universe as a whole.

(b) The 600 J of energy are supplied to the wire, in this case at constant temperature, the heat created being used to convert ice at 0 °C to water at 0 °C. If this process had been carried out reversibly there would have been a gain in entropy of the ice–water system by

$$\frac{600}{T} = \frac{600}{273} = 2{\cdot}19\,\mathrm{J\,K^{-1}},$$

which, since the process is irreversible, represents a net gain of entropy for the universe.

Chapter 5
The Third Law

5.1 Behaviour of matter at low temperatures

The experimental observation of the properties of matter as the temperature is lowered below room temperature towards the lowest values attainable leads to the conclusion that the disorder of any system decreases as the temperature decreases and, as measured in terms of the entropy, tends to a limiting value. Any attempt to formulate a law which will account for this behaviour, and which will supplement the laws discussed in the preceding chapters, must take into account a number of basic observations.

(a) On the temperature scale defined thermodynamically in section 4.4 (and later to be defined practically in Chapter 20) there is found to be an absolute zero of temperature at which, for example, the volume of an ideal gas extrapolates to zero. However, the gas phase itself does not exist at the lowest temperatures since, as this zero of temperature is approached, the gas undergoes changes of phase to more ordered condensed states, first to a liquid and then, generally, to a solid. Once the solid phase is formed its internal energy will extrapolate to some zero-point value at the absolute zero of temperature.

(b) If we maintain our definition for a reversible process that

$$\Delta S = \int \frac{dQ}{T},$$

then, if C is the thermal capacity of a given condensed system so that $dQ = C\,dT$, we have

$$\Delta S = \int \frac{C\,dT}{T}.$$

It is clear that there will be a singularity in the entropy as T approaches zero unless C also approaches zero at least as fast as T. As will be discussed in Chapter 10, this required temperature dependence of the specific heat is confirmed by experimental observations on solids and is, furthermore, predicted by the method of statistical physics.

(c) The experimental evidence from chemical reactions which were allowed to take place at low temperatures suggested to Nernst that:

The entropy change which occurs in a chemical reaction between two ordered solids tends to zero as the temperature tends to zero.

This statement is known as the *Nernst heat theorem.*

(d) It is observed that in any magnetic system the spin magnetic moments from which the magnetism arises will tend to a completely ordered arrangement as the zero of temperature is approached, so that the weight of the state of the magnetic moments at absolute zero tends to unity (there being only one way of attaining complete order). The corresponding entropy, given by the logarithm of the weight, is zero.

The experimental results summarized here led Simon to produce a general statement governing the thermodynamic behaviour of such properties as specific heat, paramagnetism of solids and many others as absolute zero is approached. The result was the 'Nernst–Simon' statement, viz.:

The entropy change associated with any reversible isothermal process of a condensed system approaches zero as the temperature approaches zero.

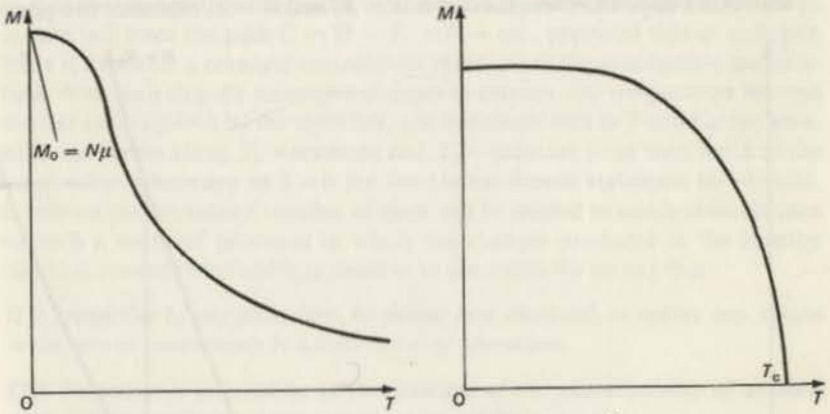
This more general statement is the *third law of thermodynamics*. (In this statement 'condensed system' normally refers to a solid but will, in the case of liquid-helium-II, refer to the liquid phase.)

Table 2 Some Results of the Third Law

System	*Quantity considered*	*Result*
voltaic cell	electromotive force V	$\lim_{T\to 0}\frac{dV}{dT}=0$
elastic solid	isothermal bulk modulus β	$\lim_{T\to 0}\frac{d\beta}{dT}=0$
surface of liquid	surface tension σ	$\lim_{T\to 0}\frac{d\sigma}{dT}=0$
magnetic material	magnetic moment M	$\lim_{T\to 0}\frac{dM}{dT}=0$
superconductor	critical field B_c	$\lim_{T\to 0}\frac{dB_c}{dT}=0$
solid	specific heat C_v	$\lim_{T\to 0}C_v=0$
solid–liquid phase change	pressure p for change of phase at temperature T_t	$\lim_{T_t\to 0}\frac{dp}{dT_t}=0$

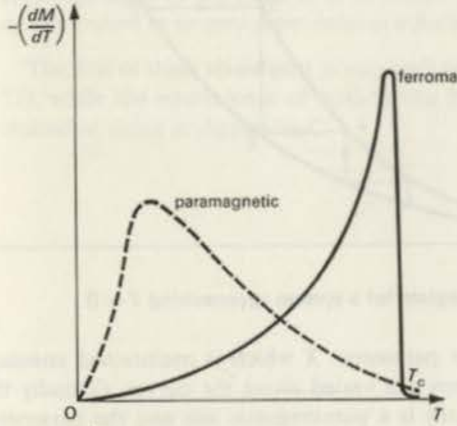
Although we have not dealt with detailed thermodynamic properties so far, it is instructive to list in Table 2 a few of the significant results of this law which will illustrate its importance. (It must be pointed out, however, that the absolute zero of temperature has not actually been reached and that the verification of these predictions is based on extrapolation to $T=0$.)

The results of the third law of thermodynamics for magnetic materials are illustrated in Figure 17. This shows the variation with temperature of the magnetization for a paramagnetic material in the presence of a magnetic field,



paramagnetic material
(a)

ferromagnetic material
(b)



slope of magnetization curve
(c)

Figure 17 Behaviour of magnetic materials as absolute zero is approached and for a single domain ferromagnetic material of Curie temperature T_c, and also the variation of dM/dT for each case. It is clear that the experimental results indicate agreement with the third law.

5.2 The unattainability of absolute zero

An alternative form of the third law to that given by the Nernst–Simon statement is obtained by considering the representation of a system on a temperature–entropy diagram as the temperature approaches zero. Such a representation is shown in Figure 18, where the curves $X = X_1$ and $X = X_2$ represent two possible values of a particular constraint parameter X which is maintained constant while the temperature and entropy are varied along the curves. (Usually the system chosen for these discussions is a paramagnetic salt and the parameter held constant is the magnetic field. Then we normally have $X_1 = B_{01}$, where B_{01} is zero, and $X_2 = B_{02}$, where B_{02} is a large magnetic field capable of producing a high degree of magnetic ordering. However, X could equally be the pressure applied to a gas or fluid, the force on a wire or the electric field on a dielectric.)

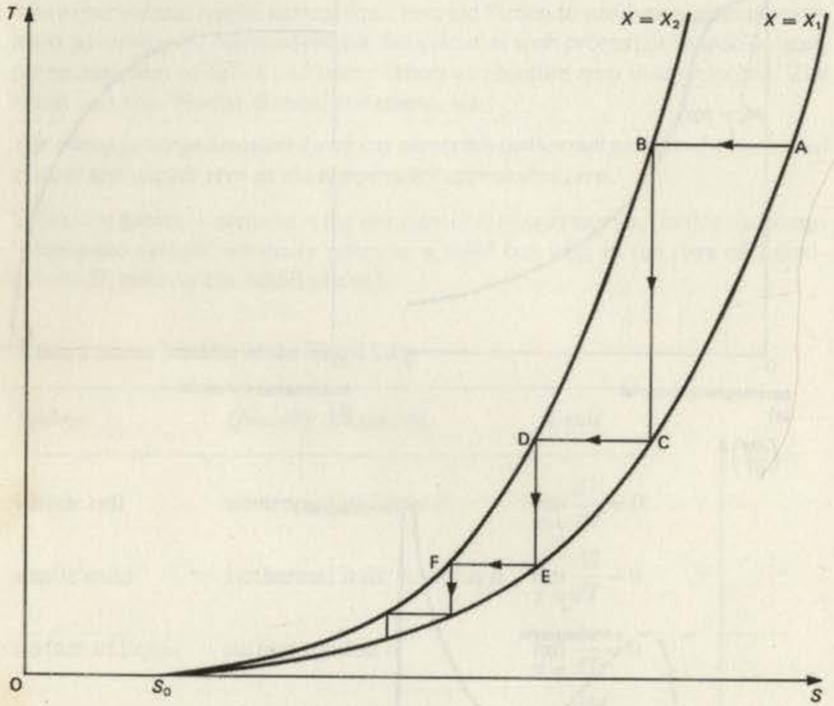


Figure 18 Temperature–entropy diagram for a system approaching $T = 0$

In the figure S_0, the entropy remaining at $T = 0$, is the *zero-point entropy*, which may be considered to be due to the disorder which remains 'frozen' into the solid even at absolute zero.

Now suppose that, starting with the system in the state represented by the point A, the value of X is changed isothermally from X_1 to X_2, so that the entropy changes to the value at B. (This process would require some heat reservoir at constant temperature to supply or remove heat as required to maintain the isothermal condition.) Once the system is at the state B it may be isolated from its surroundings while the value of X is changed isentropically from X_1 to X_2 so that the temperature falls to the value at C. A repetition of each of these steps in turn will trace the path C → D → E → F → etc., provided that at each step there is available a constant-temperature reservoir at the appropriate temperature. With each step the successive changes in entropy and temperature become smaller as is required by the third law, and approach zero as T approaches zero, since the curves along X_1 = constant and X_2 = constant must both finish at the same value of entropy at $T = 0$ for the Nernst–Simon statement to be valid. It follows that an infinite number of steps will be needed to attain absolute zero by such a series of processes in which the changes produced in the entropy diminish towards zero and it is possible to generalize by stating that:

It is impossible by any procedure, no matter how idealized, to reduce any system to the zero of temperature in a finite series of operations.

This statement is referred to as the *principle of the unattainability of absolute zero*, and by exactly the same reasoning we see that:

It is impossible by any procedure, no matter how idealized, to reduce the entropy of any system to its zero-point value in a finite number of steps.

The first of these statements is very well borne out by experience (see Chapter 15), while the equivalence of both to the Nernst–Simon statement is demonstrated in detail in Appendix C.