AN ELEMENTARY INTRODUCTION TO ENTROPY VIA IRREVERSIBILITY

R. GILES

Mathematics Department, Queen's University, Kingston, Ontario, Canada

ABSTRACT

An outline is given of an approach to thermodynamics in which entropy is derived directly, from a simple postulate of direct experimental significance, without reference to temperature or thermal equilibrium. The author first shows how the irreversibility of a natural process may be quantitatively measured; entropy is then defined so that its increase in any process equals the irreversibility. Lastly equilibrium states are defined and absolute temperature is derived from entropy by differentiation.

---

1. INTRODUCTION

The following is an outline, necessarily very condensed, of the initial stages of a one-term course in thermodynamics which was offered for several years to the honours class at Glasgow University. The characteristic feature of the treatment is that entropy is the first and not the last of the basic thermodynamic quantities to be formally introduced; and it is not introduced *ad hoc*, but derived from a very simple and plausible postulate (4.1 below) having a direct experimental meaning. In this way its fundamental significance is made apparent and it gains for the student an aura of reality which is often not realized by conventional treatments. Another didactic advantage is that instead of having to obtain entropy from temperature and energy by a process of integration we use *differentiation*—to the average student a much simpler procedure—to define temperature in terms of entropy.

The treatment can be regarded as a very much simplified version of that developed in *Mathematical Foundations of Thermodynamics*[1] (hereafter referred to as *MFT*). At the expense of some sacrifice in rigour, mathematical sophistication has been avoided and explanations of physical concepts have been largely replaced by illustrative examples. (For reasons of space, however, the number of examples in the present account has had to be severely limited.)

2. SYSTEMS AND STATES

Space limitations preclude any proper discussion of these concepts. We denote systems by capital letters $A, B, \ldots$ and states of A by $A_1, A_2, \ldots$. One example must suffice for illustration. Let L denote a (particular) solid metal cube. Let L_1, L_2 and L_3 denote the states of L in which its temperature is uniform and $0°C$, $50°C$ and $100°C$ respectively. Let L_4 denote the steady

503

state attained by L when two opposite faces are maintained at $0°C$ and $100°C$ respectively. Thus in L_4 there is a uniform temperature gradient across the block. Clearly L_4 is 'different' from the other states: if L is isolated the state L_4 will change, the temperature gradually becoming more uniform until eventually L_2 will result. (This unusual behaviour of L_4 is, of course, due to the fact that it is not an 'equilibrium state'. Since we make no restriction to equilibrium states, however, we do not need to give a formal explanation of the term at this stage.)

Two systems A and B may be *thought of* as together comprising a single system which we call their *union* and denote $A + B$. To form the union is a purely conceptual process: it is not necessary that the systems interact or even be in contact. However, in practice there is little point in considering $A + B$ unless some, possibly indirect, interaction between A and B is contemplated. Occasionally we will use such an expression as $2A$. This is shorthand for $A + A$ and denotes the union of A with *a replica of itself*. Finally, $A_1 + B_1$ will denote that state of $A + B$ in which the systems A and B are separated (i.e. not in interaction) and in states A_1 and B_1.

3. PROCESSES

If the state of a system changes a *process* is said to have occurred. We name the process by giving the initial and final states: for example, (L_4, L_2) denotes the process (mentioned above) of settling down which occurs in L if it is isolated and initially in the state L_4. That this notation for processes—involving only the naming of the initial and final states—is justifiable depends on the following circumstance: classical thermodynamics is concerned primarily with those properties of processes which depend only on the initial and final states, being independent of the particular manner by which the change of state took place. Thus if two processes have these features in common they need not be distinguished and will be called *equivalent*. As an important example, any process for which the initial and final states coincide is equivalent to the trivial process in which no change whatever takes place; this trivial process we call the *zero process*.

(L_4, L_2) is an example of a process which can occur *in isolation*: i.e. while the system concerned, namely L, is isolated. It is called a *natural process* and we write $L_4 \rightarrow L_2$. On the other hand it is not *reversible* for its *reverse* (L_2, L_4) cannot occur in isolation. (Of course we can by external action *compel* the process (L_2, L_4) to occur—for instance by enclosing L, initially in the state L_2, between two heat reservoirs.) We shall call (L_2, L_4) *anti-natural* (against nature) since its reverse is natural. We have thus $L_4 \rightarrow L_2$ but $L_2 \nrightarrow L_4$.

The process (L_1, L_2) is another example of a process which cannot occur in isolation, although it can of course be compelled to occur by bringing L into contact with a suitable heat reservoir. Exactly the same applies to its reverse (L_2, L_1). (Isolation implies, in particular, perfect thermal insulation so that cooling, just as much as heating, is impossible.)

Now suppose we have two copies of the block L in the states L_1 and L_3. By bringing the two blocks together, waiting until no further change takes

place, and then separating them again we obtain the final state $L_2 + L_2$. Thus $(L_1 + L_3, L_2 + L_2)$ is a natural process†. Let us denote this process by α. The process α has involved two systems, each a replica of the block L, which have experienced the processes (L_1, L_2) and (L_3, L_2) respectively. Denoting these processes by β and γ we might say that α consists in the 'simultaneous occurrence' of β and γ. We express this by writing $\alpha = \beta + \gamma$: i.e. $(L_1 + L_3, L_2 + L_2) = (L_1, L_2) + (L_3, L_2)$. Notice that neither β nor γ is natural, although their sum α is.

The *sum* of any two processes is defined in the same way: the sum of the processes (A_1, A_2) and (B_1, B_2) of the systems A and B respectively, is the process $(A_1 + B_1, A_2 + B_2)$ of the system $A + B$. Observe that the sum of a process and its reverse is (equivalent to) the zero process:

$$(A_1, A_2) + (A_2, A_1) = (A_1 + A_2, A_2 + A_1) = 0$$

since $A_1 + A_2$ and $A_2 + A_1$ are really the same state. Hence we may call the reverse of a process α the *negative* of α and denote it $-\alpha$.

We have agreed to write $A_1 \to A_2$ and to describe the process (A_1, A_2) as natural not only when the process (A_1, A_2) can occur in isolation but also when it can be caused to occur by an arbitrarily small external interference. We now make a further relaxation of these conditions. Suppose that we can envisage some apparatus K which can be used to cause A to undergo the process (A_1, A_2) and suppose, moreover, that this can be done in such a way that the final state of the apparatus coincides with its initial state, K_0 say. In this case the apparatus has in no sense been 'used up' in the process (we shall say it is not *involved* in the process)—indeed it is at once ready to be employed in the same way again. We agree to allow this sort of use of auxiliary apparatus:

3.1. Definition

We write $A_1 \to A_2$ and call the process (A_1, A_2) *natural* whenever there exists some system K and some state K_0 of K such that the process $(A_1 + K_0, A_2 + K_1)$ can occur while the system $A + K$ is isolated, the state K_1 being equal to (or at least differing arbitrarily little from) K_0.

Sometimes the system K takes the form of an *engine* which works in cycles—if a whole number of cycles has been performed the initial and final states will coincide.

Using this definition we can establish two results that we shall need later:

3.2. Theorem

Let A and B be any systems, A_1 and A_2 states of A, B_0 a state of B. Then:
(a) If $A_1 \to A_2$ then $A_1 + B_0 \to A_2 + B_0$
(b) If $A_1 + B_0 \to A_2 + B_0$ then $A_1 \to A_2$.

Proof.
(a) By hypothesis there is an 'engine' K which, starting and finishing in

---

† It is true that some external agency has been used to move the blocks, so the system $2L$ has not been strictly isolated. However, we still regard the process as natural since—there being no limit on the time required—the external interference can be arbitrarily slight.

some state K_0, can take A_1 into A_2. We need now merely retain B in the state B_0 during this process, and observe what has happened to $A + B$†.

(b) If K, with initial and final state K_0, can implement $(A_1 + B_0, A_2 + B_0)$ —i.e. if $(A_1 + B_0 + K_0, A_2 + B_0 + K_0)$ can occur in isolation—then $B + K$, with initial and final states $B_0 + K_0$, implements the process (A_1, A_2).

The proof of the following theorem is similar:

3.3 Theorem
If $A_1 \to A_2$ and $A_2 \to A_3$ then $A_1 \to A_3$.

Using these results it is now easy to prove:

3.4 Theorem
If α and β are natural processes then so is $\alpha + \beta$.

4. IRREVERSIBILITY

We now introduce the basic postulate of our formulation:

4.1. Postulate
Let A_1, A_2, A_3 be any states of any system A. If $A_1 \to A_2$ and $A_1 \to A_3$ then either $A_2 \to A_3$ or $A_3 \to A_2$ (or possibly both).

On this postulate depends the construction of an entropy function and thus the whole structure of thermodynamics. From it we deduce:

4.2. Theorem
Given any two natural processes α and β, one of them is able to 'drive the other backwards': i.e. either $\alpha - \beta$ or $\beta - \alpha$ is natural (possibly both).

The proof is simple (see MFT, p 34). This theorem makes it possible to measure quantitatively the irreversibility of a natural irreversible process. Indeed, we are going to assign to each possible‡ process α a scalar quantity $I(\alpha)$, the *irreversibility* of α, in such a way that:

(i) $\begin{cases} I(\alpha) > 0 \text{ if } \alpha \text{ is natural irreversible,} \\ I(\alpha) = 0 \text{ if } \alpha \text{ is reversible,} \\ I(\alpha) < 0 \text{ if } \alpha \text{ is antinatural irreversible;} \end{cases}$

(ii) I is *additive*: i.e. $I(\alpha + \beta) = I(\alpha) + I(\beta)$, for all possible processes α and β.

We measure the irreversibility of a natural process α by comparing it with that of a standard irreversible process γ, the irreversibility $I(\gamma)$ of γ being assigned arbitrarily§. We say α *is at least (most) r times as irreversible as γ* if

---

† It is necessary to assume that any state can be 'frozen', i.e. kept unchanged, when required. This may require some cunning. To freeze the state L_4, for instance, we may imagine that the block L is built out of a large number of thin square metal plates and that these are instantly separated from each other; on reassembly, the state L_4 is restored.

‡ 'Possible' means 'natural or antinatural or both'.

§ Any such process γ may be used [or, if there is none, we simply set $I(\alpha) = 0$ for all α]. However, to get the customary scales of entropy and temperature we may take for γ a natural process of the form $\gamma = (M_1 + R_1, M_2 + R_2)$ where M is a mechanical system (see § 5) with M_1 exceeding M_2 in energy by 1 erg, and R is a sealed container enclosing only a mixture of ice, water and water vapour (R is thus a 'heat reservoir' at the triple point of water); and set $I(\gamma) = (1/273 \cdot 16)$ erg/deg, *inventing* the new unit erg/deg to measure the new fundamental quantity, irreversibility.

$q\alpha - p\gamma$ is natural (antinatural) for positive integers p and q with $r = p/q$. A straightforward argument (*MFT*, p 43) based on theorems 3.4 and 4.2 shows that there is a unique real number r_0 such that, for any smaller (larger) rational number r, α is at least (most) r times as irreversible as γ. (Moreover, the principle mentioned in definition 3.1—that arbitrarily small changes in the environment are permissible— allows us to conclude that if r_0 is rational it belongs to both these classes.) We set $I(\alpha) = r_0 I(\gamma)$. It follows from the construction that the function I so defined has the property (i) above. That, for any natural processes α and β, $I(\alpha + \beta) = I(\alpha) + I(\beta)$ follows from the easily established fact (*MFT*, p 45) that if α and β are at least (most) r times and s times as irreversible as γ respectively, then $\alpha + \beta$ is at least (most) $r + s$ times as irreversible as γ.

With trivial changes the above definition of $I(\alpha)$ can be extended to every possible (i.e. natural or antinatural) process α.

5. ENTROPY

We now introduce the notion of a *mechanical system*: i.e. one of those idealized systems dealt with in elementary mechanics, from which dissipative forces (friction, viscosity, etc.) are absent. We assume as characteristic of mechanical systems that (a) the union of two mechanical systems is a mechanical system, and (b) any natural process involving only a mechanical system is reversible; thus, if α is such a process, $I(\alpha) = 0$. We define an *adiabatic process* of a system A to be a process of $A + M$ where M is any mechanical system: i.e. a process which 'involves', apart from A, only a mechanical system. Thus a *natural adiabatic process* of A means a natural process of $A + M$, and so on. Lastly, we assume (cf. Pippard[2] p 15) that any two states of a closed system can be connected, in at least one direction, by an adiabatic process. We can now prove:

5.1. Theorem
The irreversibility of a natural adiabatic process depends only on the system's initial and final states: i.e. if α and β are two natural adiabatic processes of A, both leading from A_1 to A_2, then $I(\alpha) = I(\beta)$.

Proof.

Let $\alpha = (A_1 + M_1, A_2 + M_2)$ and $\beta = (A_1 + N_1, A_2 + N_2)$, where M and N are mechanical systems. Since α and β are natural either $\alpha - \beta$ or $\beta - \alpha$ is natural, say $\alpha - \beta$. But $\alpha - \beta = (A_1 + M_1 + A_2 + N_2, A_2 + M_2 + A_1 + N_1) = (M_1 + N_2, M_2 + N_1)$ is a process involving only the mechanical system $M + N$. Being natural, it is also reversible. Thus $0 = I(\alpha - \beta) = I(\alpha) - I(\beta)$.

5.2. Theorem
If α and β are natural adiabatic processes of systems A and B, leading from A_1 to A_2 and from B_1 to B_2 respectively, then $\alpha + \beta$ is a natural adiabatic process of $A + B$ leading from $A_1 + B_1$ to $A_2 + B_2$, and $I(\alpha + \beta) = I(\alpha) + I(\beta)$.

Proof.

The first statement follows immediately from the definitions and the second is just property (ii) of §4.

An inspection of these theorems shows that 'natural' may be replaced by 'possible' without affecting the proofs.

We can now define an entropy function S. For each system A first choose arbitrarily† a reference state A_0 and assign it zero entropy: $S(A_0) = 0$. Let A_1 be any other state. By assumption there exists a natural adiabatic process connecting A_0 and A_1. If it leads from A_0 to A_1 call it α; if it leads from A_1 to A_0 call its reverse α. In either case define $S(A_1) = I(\alpha)$.

With this definition the entropy of every state of every mechanical system is automatically zero. The following theorem can now be easily proved.

5.3. Theorem

(a) For any states A_1 and B_1 of systems A and B, $S(A_1 + B_1) = S(A_1) + S(B_1)$.

(b) Let α be a natural adiabatic process of a system A leading from A_1 to A_2. Then $I(\alpha) = S(A_2) - S(A_1)$.

Now, any natural process (A_1, A_2) involving only a system A can be regarded as a special case of a natural adiabatic process of A [by writing it in the form $(A_1 + M_1, A_2 + M_1)$]. Applying this to the case when A is the union of several other systems, we have, in view of the additivity of entropy:

5.4 Corollary

In any natural process the total entropy of all the systems involved never decreases, and it remains constant only if the process is reversible.

6. EQUILIBRIUM STATES AND TEMPERATURE

The introduction of entropy in §5 involved no reference to temperature. This is not surprising since most states—e.g. L_4 or $L_1 + L_3$—do not 'have' a temperature at all. However, we now define an *equilibrium* state in such a way that every equilibrium state has a temperature. The usual meaning of 'equilibrium' is somewhat vague and involves reference to the internal structure of the state; ours is quite specific, involving only the concepts that we have already introduced. Roughly speaking, an equilibrium state is a state of 'maximum settled-down-ness':

6.1. Definition

A_1 is an *equilibrium state* of a system A if there is no state A_2 such that (A_1, A_2) is natural irreversible process.

It is easy to deduce from this definition that $A_1 + B_1$ can be an equilibrium state only if A_1 and B_1 are equilibrium states. However, this condition is not sufficient: for instance $L_1 + L_3$ is not an equilibrium state (see §3).

To introduce temperature we must first construct an internal energy function E. Our route is the usual one[2], differing only in certain details. We assume that every state of a mechanical system has a definite energy and

---

† Except that if A_0 and B_0 are the reference states for A and B then the reference state chosen for $A + B$ must be $A_0 + B_0$.

that for mechanical systems energy is additive and always conserved (i.e. in any natural process 'involving' only a mechanical system the initial and final energies are equal). We define the *work* W done on A in an adiabatic process to be the decrease in energy of the mechanical system involved. We can then *prove*† the first law of thermodynamics in its usual form[2,3].

The introduction of temperature is most simply described in the case of a simple fluid, or *chemical system* in the sense of Zemansky[3]. In the present context such a system is best defined as one with the following two properties:

(a) every state of the system has a definite *volume* V,

(b) if two states A_1 and A_2 have the same energy and the same volume then either $A_1 \to A_2$ or $A_2 \to A_1$ (or both).

It follows that two equilibrium states of the same energy and volume must have the same entropy. The equilibrium states of a simple fluid thus lie on an *equilibrium surface* $S = S(E, V)$ in a 'space' with coordinates E, V, S. If we assume, as is customary in physics, that this surface is sufficiently smooth we can now define the temperature T of any equilibrium state by the equation $1/T = \partial S/\partial E$. At the same time the pressure P may be defined by the equation $P/T = -\partial S/\partial V$.

It is a simple matter to show that T and P have the qualitative and quantitative properties associated with the terms absolute temperature and pressure.

REFERENCES

[1] R. Giles. *Mathematical Foundations of Thermodynamics.* Pergamon: Oxford (1964).
[2] A. B. Pippard. *The Elements of Classical Thermodynamics.* Cambridge University Press: London (1957).
[3] M. W. Zemansky. *Heat and Thermodynamics,* 4th ed. McGraw-Hill: New York (1957).

† That this is possible is due to the strength of the assumption stated in the previous parenthesis. The proof is practically that of theorem 5.1, with W replacing I.