The Origins of Time-Asymmetry in Thermodynamics: The Minus First Law

Harvey R Brown[*] *and Jos Uffink*[^]

This paper investigates what the source of time-asymmetry is in thermodynamics, and comments on the question whether a time-symmetric formulation of the Second Law is possible.

1. Introduction

Many authors have addressed the topic of time-reversal invariance in connection with thermodynamics and statistical mechanics. Most of them have assumed that the origin of the time reversal non-invariance (or time-asymmetry) of thermodynamics lies in its Second Law. However, in a recent paper in this journal one of us argued against this view. It was pointed out that modern and careful formulations of the Second Law do not presuppose or rely on asymmetric assumptions. This does *not* mean that thermodynamics is time symmetric. Rather, it was stated in the paper[1]:

> It is often said that [the approach to equilibrium] … is accompanied by an increase in entropy, and [is] a consequence of the second law. But this idea actually lacks a

[*] Sub-Faculty of Philosophy, University of Oxford, 10 Merton Street, Oxford OX1 4JJ, UK (*e-mail*: harvey.brown@philosophy.ox.ac.uk).

[^] Institute for History and Foundations of Science, Utrecht University, P.O. Box 80.000, 3508 TA Utrecht, The Netherlands (*e-mail*: j.b.m.uffink@phys.uu.nl).

[1] Uffink (2001).

theoretical foundation. … This aspect of time asymmetry is woven much deeper in the theory.

One of the objectives of the present note is to spell out in greater detail the claim that the approach to equilibrium does not follow from the Second Law. Section 2 is concerned with a simple, concrete application of this law, demonstrating that it does not *legislate* the approach to equilibrium so much as *characterise* it. In section 3, we state more precisely what the approach-to-equilibrium principle—the "Minus First Law"—is. We argue that its time-asymmetrical component lies in the postulated nature of the equilibrium states themselves. In section 4, we contrast this with the corresponding "problem of Boltzmann" in classical statistical mechanics, and make special reference to what we call the "second problem of Boltzmann", and its treatment by Boltzmann and Schrödinger.

Section 5 takes up the issue of whether the increase of entropy is a matter of definition, as Hawking has recently claimed. The question is particularly pertinent within the formal approach to thermodynamics initiated by Carathéodory. This leads us in section 6 to examine the claim that the counterpart of the Second Law in this approach and its recent elaboration by Lieb and Yngvason is time-symmetric. We finish with an aside in section 7 comparing the role of mathematical abstraction in thermodynamics and spacetime theory.

2. What the Second Law does not do

Consider claims of the following kind: that the "fundamental fact of irreversibility is summarised in the Second Law of Thermodynamics"[2], or that we might think of the Second Law "as nature's way of driving systems towards equilibrium"[3]. Such claims are common enough. But are they correct?

Imagine a cylinder, within which a gas is contained by a frictionless piston, and which can have occasional diathermal contact with a single heat reservoir. At a given

---

[2] Sklar (1993), p. 21.

[3] Davies (1999).

time t_0, the gas in found in the equilibrium state *A* and the piston is then made to undergo sudden motion, quickly increasing the volume of the cylinder by a definite amount. In the interval between t_0 and t_1, the gas undergoes adiabatic expansion until at it attains a new

state *B* of equilibrium. The states between *A* and *B* are not equilibrium states, so no path can be drawn between them in the accompanying entropy-temperature diagram below.
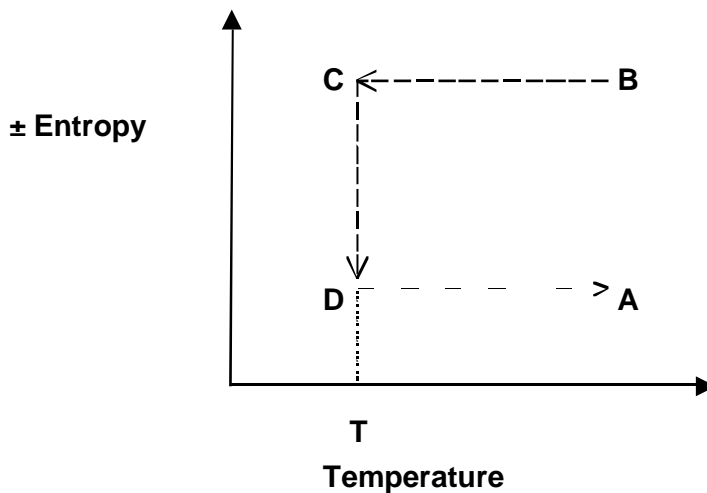
It is then arranged that the system undergoes a quasi-static, adiabatic process whereby the equilibrium state *C* is reached at time t_2, with the gas now at the temperature *T* of the heat reservoir. Between times t_2 and t_3, the cylinder is in contact with the reservoir, and the system undergoes a quasi-static isothermal change until the entropy associated with the new equilibrium state *D* at t_3 has retained the value it had for the initial state *A*. The cycle is completed when the system returns adiabatically and quasi-statically to state *A* at time t_4.

What does the Second Law of thermodynamics imply about the process of free expansion of the gas between t_0 and t_1? To some extent, the answer depends on the chosen formulation of the law, as we shall see later. Consider, for example, Kelvin's

version of the Second Law[4]: no cyclic process, such as $A \to B \to C \to D \to A$ or $A \to D \to C \to B \to A$, can have the sole effect of extracting heat from a reservoir and producing a corresponding amount of work. It is commonly argued that it follows from this principle that the transition $A \to B$, if it occurs, is 'irreversible', i.e. the converse transition $B \to A$ is impossible. However, the logic of this argument is always conditional on further assumptions, often left implicit. Clearly, Kelvin's principle only refers to *cycles* and does not assert the irreversibility of any non-cyclic process. It is only when one assumes the availability of both quasi-static processes $B \to C \to D \to A$ and $A \to D \to C \to B$, allowing the above transitions to be part of a cycle, that the principle becomes applicable to them. But these assumptions are themselves not part of Kelvin's principle.[5]

Hence, the implication of Kelvin's principle for the free expansion process can be summarized as a conditional statement: *if* the gas spontaneously expands to a new state of equilibrium, and *if* certain other processes are available, *then* the converse transition is impossible. But that this expansion occurs spontaneously is likewise not part of the content of the Law. What we resist is the supposition that the Second Law drives systems towards equilibrium, and that it is the most fundamental point of entry of time-asymmetry into thermodynamics.

3. The "Minus First Law"

---

[4] This is the version of the Second Law comonly attributed to Kelvin; his original formulation is discussed in Uffink (2001), section 5.

[5] It might also be worth noting that Kelvin's principle is not needed in this case to infer that the transition $A \to B$ involves an entropy increase. This fact can easily be seen to follow from the positivity of gas pressure, the invariance of internal energy in the transition and the First Law. However, if one considers a cycle of the kind given in the above entropy-temperature diagram, but without specifying the nature of the thermodynamic system, then Kelvin's principle can be used to infer entropy increase in the transition $A \to B$, subject to the availability of the quasi-static processes that close the cycle; see for example Dugdale (1996) pp. 60-62.

The fact that in thermodynamics the tendency of systems to approach equilibrium is logically prior to the Second Law may not be universally appreciated, but it has not escaped the notice of a number of commentators. The existence of this tendency has sometimes been referred to as the "zeroth law of thermodynamics"[6], in an unfortunate competition with R. H. Fowler's famous usage concerning the quite distinct principle of the transitivity of inter-body thermal equilibrium. The latter forms part of the basis of the notion of temperature[7], but the tendency towards equilibrium is a more basic principle. This point has been stressed by Joseph Kestin, for example[8]:

> The concept of temperature and our ability to perform reproducible temperature measurements rely on the fact that systems, however complex, which are made to interact across diathermal walls within a rigid adiabatic enclosure always reach a state of thermal equilibrium.

Let us articulate this phenomenological fact in the form of the following *Equilibrium Principle*:

> *When an isolated system finds itself in an arbitrary initial state within a finite fixed volume, it will spontaneously attain a unique state of equilibrium.*

The Equilibrium Principle can be broken into three distinct claims:

(A) *The existence of equilibrium states* for isolated systems. The defining property of such states is that once they are attained, the independent thermodynamic coordinates of the system are spatially homogeneous and remain thereafter constant in time, unless

---

[6] See Uhlenbeck and Ford (1963), p. 5, and Lebowitz (1994), p. 135.

[7] See, for example, Zemansky (1957), p. 6

[8] Kestin (1979), Vol I, p. 72; see also Uhlenbeck and Ford (1963), footnote 14.

the external conditions are changed. The claim that such states exist is not trivial—it rules out the possibility of spontaneous fluctuation phenomena[9].

(B) *The uniqueness of the equilibrium state*. The claim is that for any initial state of an isolated system bounded by a prescribed static surface, there is exactly one state of equilibrium.

(C) *The spontaneous approach to equilibrium from non-equilibrium*. A non-equilibrium state will typically come about as the result of a removal of internal constraints, such as the rapid displacement of adiabatic walls separating two bodies. (No indication of the speed of approach to the new equilibrium state is given of course: thermodynamics provides no equations of motion.)

The point we wish to stress here is that the time-asymmetry of thermodynamics arises, at the most basic level, through claim (A)[10]. The spontaneous motion towards equilibrium is time-asymmetric because of what equilibrium states are: once attained no spontaneous departure from them is possible without intervention from the environment.

---

[9] Readers who are squeamish about an existence claim being granted the status of a law (or part thereof) are invited to consider the case of Newton's first law of motion, which has as its content the claim that at least one ("inertial") reference frame exists, with respect to which the motion of *all* force-free bodies is uniform and rectilinear.

[10] One might debate the question whether an independent time-asymmetric element is involved in claim (B) and (C). The reason to think so is that the approach to equilibrium is typically a many-to-one transition: many different initial states will evolve to the same final state. The reversal of this transition would then be a one-to-many relation, which is excluded by claim (B). However, although this asymmetry is arguably there, it is so to say 'non-malicious'. Or in other words, it does not necessarily lead to an incompatibility with an underlying time symmetric theory. The reason is that in statistical mechanics many different microstates make up one thermodynamic macrostate. So at the microscopic level the transition from a non-equilibrium to an equilbrium state becomes one-to-one because the equilibrium state contains many more microstates than a non-equilibrium state.

The equilibrium state in thermodynamics is itself a time-asymmetric notion[11]— in contrast to a notion of equilibrium in statistical mechanics, as we shall see in the next section. Returning to the last section, it is seen that in the special case of the gas system the Equilibrium Principle not only pre-empts the Kelvin Principle in allowing just *one* of the processes $A \quad B$, $B \quad A$ to happen, it determines *which* one does.[12]

If lawlike status is to be conferred on the Equilibrium Principle, the existing appelation "Zeroth Law" clearly will not do; as we have mentioned it clashes with the now widely endorsed use of this nomenclature by Fowler for something that is logically distinct and less fundamental. Were the Equilibrium Principle the *most* fundamental tenet imaginable in thermodynamics, the term "Minus Infinite Law" might be appropriate. But is it? The Principle is, to repeat, a collection of claims, of which (A) is clearly the most basic. But (A) presupposes the ability to isolate the system of interest from the rest of the universe, and at least one author has argued that *this* is the most fundamental of all principles[13]. On the other hand, the term Minus First Law might (falsely) suggest that no further fundamental assumptions are needed between it and the Zeroth Law. Be that as it may, in this paper we adopt the term "Minus First Law"[14].

---

[11] This has been noted by Price (1996), p. 24.

[12] Some care must be taken in interpreting this last claim. The Equilibrium Principle discriminates between (i) a process in which a system, initially in equilibrium state A, is perturbed by some external intervention, and then evolves to a final equilibrium state B, and (ii) a process in which a system initially in equilibrium state B spontaneously evolves into a non-equilibrium state and is then, by external intervention brought to equilibrium state A. This latter process is ruled out. The principle remains, however, neutral in deciding between process (i) and (iii) a process in which a system initially in state B is perturbed and then spontaneously evolves to state A. So, the fact that a gas, after releasing the piston, expands to a greater volume, rather than contract to a smaller one, is not determined by the above Equilibrium Principle.

[13] Kestin (1979), vol II, p. 1.

[14] The two options in nomenclature—'law –1' and 'law – '—were suggested in Uffink (2001), footnote 93. An interesting question is whether the Minus First Law, or

4. Statistical mechanics and the problems of Boltzmann

In Boltzmann's approach to statistical mechanics, thermodynamical states correspond to so-called 'macrostates', i.e. certain subsets in a partition of the phase space Γ_M of the mechanical system. The equilibrium state corresponding to a fixed total energy E and particle number N is then identified with the subset with the largest volume under these constraints.

Uhlenbeck and Ford characterised as "the problem of Boltzmann" in statistical mechanics the following challenge[15]:

> … in its simplest form, one must 'explain' in which sense an isolated (that is conservative) mechanical system consisting of a very large number of molecules approaches thermal equilibrium, in which all 'macroscopic' variables have reached steady values.

The hope of many authors is that the problem of Boltzmann can be solved by "combining dynamics with phase space volume considerations".[16]

Such latter *combinatorial* considerations involve the recognition that the volume $|\Gamma_M|$ of the region of phase space Γ_M associated with the equilibrium macrostate M_{eq} ($|\Gamma_M| = \int_M \prod_{i=1}^{N} dr_i dv_i$ where N is the number of particles) is vastly larger than that associated with any other macrostate M when N is large

---

something close to it, itself implies any of the usual remaining laws of thermodynamics. Readers wishing to pursue this question may note that in Hatsopoulos and Keenan (1962), an attempt was made to ground virtually all of thermodynamics on a single axiom that is close in spirit to the conjuction of claims (A) and (B) above. This approach was elaborated in Kestin (1963) and (1979), Vol II, pp. 2 ff. (A correction to part of their 1962 argument was made by Hatsopoulos and Keenan in their (1965), p. 400.)

[15] Uhlenbeck and Ford (1963), p, 5.

[16] Lebowitz (1994), p. 138.

The *dynamical* considerations, on the other hand, are arguably necessary to justify the claim that almost any phase point associated with an arbitrary non-equilibrium macrostate is overwhelmingly likely to wander into the large equilibrium region of phase space. There is of course still debate as to what precise form these dynamical considerations must take (in particular, whether some type of ergodic behaviour is required[17]), and even as to whether *any* dynamical input over and above Hamiltonian flow on the phase space is necessary in solving the problem of Boltzmann. Moreover, there is debate over whether these two ingredients are sufficient. Indeed, as the words 'overwhelmingly likely' indicate, some independent probabilistic assumption seems to be needed as well. (Some authors argue that some notion of 'typicality' might take the place of 'probability') But for our purposes, these debates can be put aside.

The main point we want to emphasize is that whatever the solution to the problem is, it need not involve any time-asymmetrical elements. Indeed, it is clear that the notion of equilibrium (the equilibrium macrostate as defined above) is time-symmetric[18]; unlike its counterpart in thermodynamics. The equilibrium state cannot, in general be permanent: fluctuations out of equilibrium will occur spontaneously sooner or later for almost every initial mechanical state. Thus, Claim (A) above is not valid here. (Part of the problem of Boltzmann is specifying 'in what sense' claim A can hold.)

The highly significant fact that Boltzmann saw $\log\big|_M\big|$ as a generalisation of the thermodynamic notion of entropy—applicable now to non-equilibrium as well as equilibrium states— is not strictly relevant to the present considerations, as we are interested here in the mechanical counterpart of the Minus First Law, not the Second. But is worth noting that owing in Boltzmann's approach, the logical independence of the two laws in thermodynamics does not hold for its mechanical counterparts. The spontaneous

---

[17] Sometimes even individual commentators seem to waver on the ergodic issue; compare Lebowitz (1994), p. 135 and his (1993), p. 3. A careful, critical analysis of Lebowitz's treatment of the (first) problem of Boltzmann is found in Ridderbos (2000).

[18] However, note that not all approaches to statistical mechanics have to share this feature. In a Gibbsian approach, one can define notions of equilibrium which, like the thermodynamical notion, are time asymmetric. See van Lith (1999).

approach to equilibrium automatically involves an increase in Boltzmann entropy in statistical mechanics—a point we shall return to below.

The above problem of Boltzmann refers to the behaviour of an isolated macro-system subsequent to the preparation of a non-equilibrium state; i.e. it refers to *prediction*. The reason that an awkward time-asymmetrical element enters the picture is that if the above (time-symmetric) considerations are applied to the problem of *retrodiction*, an unacceptable conclusion seems inevitable. It is that at any time *prior* to instant at which the system—which may in principle be the observable universe—is found displaced from equilibrium, the system was, with high probability, closer to equilibrium. This conclusion is, to say the least, hard to reconcile with much of the information stored in our memories, and no one seems to believe it. This difficulty—how to prevent unacceptable retrodictions once one has acceptable predictions—we shall call *the second problem of Boltzmann*. It was nicely expressed by Schrödinger in 1950[19]:

> The following is known and is universally agreed upon: the overwhelming majority of all those micro-states that would impress our crude senses as the same observable (=macro-) state do lead to identical, moreover to the actually observed *consequences*. What ails us only, that we can equally well scan the *antecedents*. And they are—again for an overwhelming majority—entirely wrong, inasmuch as the antecedents are the mirror image in time of the aforesaid consequences; it would thus appear that the system has reached its momentary state by an "anticipation" of its actual future history in reversed order.

It is fairly well known that Boltzmann himself addressed this second problem in a number of brief passages.[20] In his (1897) he offered a choice between two options. The first is to assume as a contingent fact that at present the entire universe is in a very low entropy state. The second option is that the observable world is the result of a localised,

---

[19] Schrödinger (1950), p. 189-190. For useful recent discussions of the second problem, see Price (1996), Ch. 2, and Schulman (1997), section 2.6.

[20] Boltzmann (1895, 1897a, 1897b) and (1898). Excerpts from the first two papers (1895, 1897a) are found in Barbour (1999), pp. 341-342.

spontaneous fluctuation from a permanent equilibrium state of a vastly larger and older universe.

Critics have noticed that even granting this picture of a "great enough" universe, by far the most likely fluctuation consistent with our observations would still put the lowest entropy state at the present moment, and not into the ancient past, thus making our present memories strictly illusory.[21] Boltzmann's second suggestion does not really provide a clean solution to the retrodiction problem. Moreover, it has of course been overtaken by developments in modern cosmology. In all fairness, however, it should be remarked that Boltzmann never took his own suggestions very seriously.[22]

It is worth making two final remarks. The first is to emphasise again that the time-asymmetrical element, if any, that enters Boltzmann's picture only does so in addressing the second, and not the first problem of Boltzmann. (It is essentially this consideration that has led Price to complain that "the problem of explaining why entropy increases has been vastly overrated."[23]) This assertion is the mechanical counterpart of our claim in the previous section that no time-asymmetry was necessarily being introduced in claim (C) associated with the Equilibrium Principle. Of course, what prevents anything like the second (retrodiction) problem even arising in thermodynamics is the time-asymmetric nature of equilibrium states in this theory: no retrodictive approach to equilibrium is possible because it would imply the existence of fluctuations away from equilibrium, in conflict with claim (A).

Secondly, the common response to the second problem of Boltzmann is to attribute a very special, low-entropy state to the universe immediately after the Big Bang. If this is the correct response, it would seem that the origins of time asymmetry are

---

[21] See Price (1996), p. 35, and Lebowitz (1994), p. 142.

[22] For example, in his (1897b) Boltzmann offers the idea for the benefit of those "who cannot resist the temptation to phantasize about the universe", and his (1898) ends with the remark: "Certainly nobody will regard speculations of this kind as important discoveries or … the highest goal of science". The only virtue he claims for his speculations is that they might stimulate "the mobility of thoughts".

[23] Price (1996), p. 39.

different in kind in thermodynamics and statistical mechanics. In the former, the asymmetry is built into the theory by way of the fundamental notion of equilibrium. In the latter, it would appear that the asymmetry must come from a source distinct from the structure of the laws of the fundamental interactions, which can only mean global initial or final conditions. The price to be paid, as is widely appreciated, is that these conditions are staggeringly ungeneric, or "improbable". It is conceivable, however, that a situation closer to that of thermodynamics may re-emerge in canonical quantum gravity[24].

5. Is entropy increase true by definition?

Conventional treatments of the Second Law in thermodynamics all presuppose a distinction between past and future. Consider, for example, the basic notion of a Carnot cycle. Carnot himself clearly took it for granted that given any two distinct states within a single cycle, one of them can be picked out as the earlier one. All the traditional formulations of the Second Law likewise presuppose the distinction between 'earlier' and 'later' (or 'initial' and 'final'). To which pre-thermodynamic arrow, or arrows, of time were the founding fathers of thermodynamics implicitly referring? It is not clear whether this was a question they asked themselves, or whether, if pushed, they would not have fallen back on psychological time on some other arrow. Note, however, that it is not necessary that this pre-thermodynamic arrow need not necessarily be provided by physical laws. Any appropriate contingent facts will suffice for this purpose, and indeed, as Earman teasingly noted, one might simply choose an arrow defined by the order in which most people eat main course and desert[25].

Let us now move from thermodynamics to statistical mechanics. Boltzmann's speculations contain a startling suggestion, which has attracted considerable attention.

---

[24] This would be the case if the speculation of Barbour (1994; 2000, chapter 22), concerning the role of time-capsules in the solution of the Wheeler-DeWitt equation, is borne out.

[25] Earman (1974). For discussions of non-thermodynamic arrows of time, see Zeh (1992), Price (1996), and Schulman (1997).

Consider a localized fluctuation, on the scale of a galaxy or so. From the time symmetry of statistical mechanics it follows that it is equally likely that the local state will develop by evolving towards equilibrium as by moving away from it. However, a living being who happens to find itself within an environment that shows a local entropy gradient might simply define a time direction as going from lower to higher entropy[26].

Perhaps the most elaborate development of this position was proposed by Schrödinger (1950). Schrödinger noted that it might happen that a large system in a non-equilibrium state separates into several subsystems, which, at least for some period of time, remain isolated from each other. These 'branch' systems can then be assigned their own individual phase spaces, and hence their own individual Boltzmann entropies. Now although for each branch system an evolution towards equilibrium is still as likely as an evolution away from equilibrium, Schrödinger claimed that in all probability the collective behaviour is such that these entropies change (if at all) both monotonically and *in the same way,* i.e. decreasing or increasing together[27]. It follows from the time-reversal symmetry of statistical mechanics that both behaviours are equally probable.

Schrödinger proceeded to propose what he called *the entropy law* for any pair of such branch systems:

$$\left(S_{1B} - S_{1A} \right)\left(S_{2B} - S_{2A} \right) \quad 0$$

where S_{1B} is the entropy of system 1 at time t_B, etc., and where $t_B > t_A$. This time-symmetric law allows us to define "phenomenological time" differently from the time t

---

[26] See the translated excerpt from the 1897 Boltzmann paper in Barbour (1999), p. 342; see also the discussion in Price (1996), pp. 32-37.

[27] This is the so-called "branching hypothesis" discussed in the philosophical literature; for some acute comments about the latter, see Price (1996), pp. 44-46. It is interesting to note that although this idea is usually attributed to Reichenbach (1956), Schrödinger (1950) predates this source. Indeed, the idea, together with the term 'abgezweigte Teilsysteme' appears already in an abstract (Schrödinger, 1933) of a talk that was never published.

that appears in statistical mechanics: it is, as Schrödinger stressed, neither t nor $-t$ but rather "either t or $-t$"! His point was that from the perspective of phenomenological time, entropy always increases, whether or not such increase occurs with increasing t. In other words, phenomenologically the entropic behaviour of the world is the same whether one is dealing with the ascent phase of Boltzmann's cosmic fluctuation or the descent—as Boltzmann himself appreciated.

The idea that the direction of time might be reduced to an entropy gradient has recently been advocated by Stephen Hawking, who bases it on the assertion that our brains act like computers, which supposedly incur an entropic cost in the process of using memory[28]. It follows from this assertion that the states of the world we remember are those with lower entropy than present and future states, since to repeat Schrödinger's point, all subsystems of the universe partake in the same entropic flow. For Hawking, this "reasonable" notion that "the psychological arrow of time" coincides with the thermodynamic arrow of entropy increase leads to a striking conclusion[29]:

> So the second law of thermodynamics is really a tautology. Entropy increases with time, because we define the direction of time to be that in which entropy increases.

In our opinion, some care must be exercised in interpreting this claim.

It is probably right to interpret Hawking's point as similar to Boltzmann's speculation regarding what Schrödinger called "phenomenological time": as long as entropy is changing, in a world sufficiently rich to contain observers it will appear to such

---

[28] Hawking (1994, p. 348) states that the process of recording information in memory in computers must involve an increase of entropy, but the claim need not take this precise form. Others argue that it is memory *erasure* that *necessarily* involves a finite entropic cost, even if in practice entropy normally increases with memory formation. Be that as it may, the relevant claim is that in practice, for all known computers and brains, the use of memory in the widest sense of the term is entropically costly. Doubts as to the validity of this claim are expressed in Earman (1974), p. 34.

[29] Hawking (1994), p. 348.

observers that entropy is increasing[30]. But a word of warning is perhaps not out of place. The Boltzmann claim is made in the context of statistical mechanics. Even in that context, the conclusion is dependent on the validity of contingent facts (e.g. the branching hypothesis): the Second Law is not strictly tautological (or perhaps better: analytic). If Hawking's reference to *thermodynamics* above is taken literally, his (Hawking's) claim is even more questionable.

This can best be seen by considering the formal, axiomatic approach to thermodynamics found in the work of Carathéodory[31], and based on the existence of an "adiabatic accessibility" relation defined on the non-metric, differentiable space   of equilibrium states. The statement corresponding to the Second Law in this approach is that known as Carathéodory's Theorem, which states that for so-called 'simple systems', state functions exist on   which play the role of entropy and absolute temperature. In his original 1909 work, Carathéodory discusses an arbitrary adiabatic process of a simple system (say: the free adiabatic expansion of a gas), and argues that it both is irreversible and involves a change of entropy—but that his principles are incapable of determining whether the entropy increases or decreases.

What is determinate in Carathéodory's approach is that entropy can be defined for independent systems consistently, i.e. it will either increase or decrease for all such simple systems. For Carathéodory, it was explicitly a matter of *appealing to experience* to determine whether entropy increases or decreases. For certain commentators (such as Ehrenfest-Afanassjewa), this feature of the approach was a virtue, for others (such as Planck) it was a defect.[32] More important for our purposes is that in this approach one obtains Schrödinger's entropy law, namely that all isolated systems undergoing adiabatic

---

[30] This is distinct from the Leibnizian claim defended by, for example, Reichenbach (1956) and Gold (1966), that the model of the universe created by taking the total history of our universe, say, and inverting it with respect to time, is just our universe under a different description. This latter position comes under sustained attack in Earman (1974), section 4.

[31] Carathéodory (1909), (1925).

[32] See Uffink (2001), sections 9, 10.

processes suffer a change of entropy of the same kind, i.e. either all positive or all negative.

Again, it is a subtle matter to interpret this claim, and the brief presentation of Carathéodory's 1909 discussion on irreversible processes in simple systems in Uffink (2001) may be slightly misleading. For the purpose of concreteness, this discussion was framed in terms of the example of the free adiabatic expansion of a gas. This is unfortunate in the sense that it may suggest that the argument *presupposes* that the gas spontaneously expands, not contracts, as time progresses, and that all that Carathéodory's axioms leave undetermined is whether entropy increases or decreases in the process of *expansion*. But these two possibilities, surely, cannot be regarded as each other's time reversal. (Perhaps an analogy might help. One can imagine a theory of the gravitational interaction, which leaves to experiment the question whether the gravitational interaction is attractive or repulsive. Obviously, these two possibilities are not each other's time reversal.) The possibility of entropy decrease in this context would correspond to a world vastly different from ours in which it not even clear that sentience is possible.

However, this is not the correct construal of Carathéodory's argument. The statement that a gas expands during a free adiabatic process does not follow from his axioms. In fact, it is exactly the empirical input which he needs to gauge the sign of entropy changes. The difference between the view of Carathéodory and that of Boltzmann and Schrödinger is that Carathéodory relies on a presupposed distinction between earlier and later, and uses the behaviour of a gas as a standard to fix the sign of changes in thermodynamic entropy; whereas the latter authors start from an explicitly defined sign for the Boltzmann entropy, and then use the behaviour of a gas (or any other irreversible process) to gauge the direction of phenomenological time.

So in the context of Carathéodory's thermodynamical discussion, it seems far from analytic that entropy increases.

6. Time symmetric versions of the Second Law

Now, even though Carathéodory's approach to the second law leaves it to experience whether entropy increases or decreases in adiabatic processes of simple systems, this is not the same as saying that his second law is time-symmetric. In order to see whether a law is time-symmetric or not, it is best to state a clear criterion. In an influential paper, Earman[33] formulated a condition for a theory to be time-reversal invariant. The idea is here that one specifies a transformation * on the states of the theory, implementing their reversal, and a theory is time-reversal invariant if the class of dynamically allowed models is invariant under the mapping: $s(t) \quad s * (-t)$. Note that not all commentators are in sympathy with such a criterion of symmetry based on inter-world, rather than *intra*-world considerations[34]. But let us put any such doubts aside.

Applying this criterion to thermodynamics is not completely straightforward. Thermodynamics has no equations of motion and it is hard to specify its dynamically allowed models. The only clear instance where reference to time is made explicitly in Carathéodory's approach is in the distinction between 'initial' and 'final' that appears in the adiabatic accessibility relation. Hence, a natural option is to call a law or set of laws time-symmetric iff *they continue to hold under reversal of the adiabatic accessibility relation on the space of equilibrium states*.

Now, Carathéodory's version of the second law states, roughly speaking, that in the neighbourhood of every equilibrium state there exists another, which is adiabatically inaccessible to it. That is: there exists no state from which it is possible to reach all neighbouring states by some adiabatic process. Under the reversal of adiabatic accessibility, this would transform into the statement that there exist no state which can be reached by all its neighbouring states. Carathéodory's formalism, however, allows the latter situation. Thus the principle is not time-reversal invariant.

However, both cases mentioned here are somewhat pathological. For a gas, the first situation would obtain if the entropy function has an absolute or local minimum (say a state of zero absolute temperature). The second case would obtain if its entropy function had a local maximum (i.e. a state in which the system would be unable to absorb

---

[33] Earman (1974), p. 23.

[34] See for instance Brown and Sypel (1995) and Budden (1997).

more heat). But neither possibility occurs in ordinary systems. Hence one can, at little risk, strengthen Carathéodory's principle by excluding *both* cases to obtain a time-symmetric formulation.

A similar consideration applies to the recent axiomatisation of thermodynamics due to Lieb and Yngvason[35]. This latter approach is related to that of Carathéodory, again being based on a suitably defined relation of adiabatic accessibility on the state space (though with a somewhat different interpretation of 'adiabatic'). The main result Lieb and Yngvason establish is the existence, for all 'simple' systems (again, for a somewhat different interpretation of 'simple' systems) of an entropy function which increases under adiabatic processes. The principal advantages of their approach are that it does not presuppose the differentiability of — and so is capable in principle of handling phase transitions and critical states— and more importantly that it guarantees that the entropy and absolute temperature functions are defined *globally* on [36].

Remarkably, the axioms needed to obtain their version of the second law (which they call "the Entropy Principle") are time-reversal invariant in the sense indicated above. In the case of Lieb and Yngvason, what non-invariant axioms there are, are not directly related to the derivation of their Entropy Principle[37]. (Of course, the invariance of the axioms under this reversal implies trivially that, under the same conditions needed to obtain their Entropy Principle, there exists also a function that decreases under adiabatic processes.)

Again, as in the case of Carathéodory, the Lieb-Yngvason approach actually obtains more than just the result that an entropy function exists for simple systems, with the property that entropy increases (or decreases) for adiabatic processes. An important additional feature is that these entropy functions can be chosen *consistently* for all simple systems, i.e. as either increasing or decreasing for all such systems. What is even more, these entropy functions are additive, which implies the highly non-trivial fact that the possibility of adiabatic processes in combined systems is determined just by the sum of

---

[35] See Lieb and Yngvason (1998, 1999, 2000).

[36] For a review of the Lieb-Yngvason approach, see Uffink (2001), section 10.

[37] Uffink (2001), section 11.

their individual entropies. This in turn implies Schrödinger's entropy law: for any two non-interacting simple systems, entropy changes will always occur in the same sense.

Some final words of clarification may be in order.

In the above analysis, time reversal is implemented, to repeat, by the reversal of the adiabatic accessibility relation. Such a reversal should then correspond to a world in which processes occur which look like those occurring in a film of our world but played backwards. In particular, the reversal of the spontaneous adiabatic expansion of a gas discussed in section 2 above would correspond to a spontaneous adiabatic contraction. But this behaviour is inconsistent with claim (A) of the Minus First Law, which as we have seen rules out spontaneous deviations from equilibrium. Indeed, if one starts from the notion of equilibrium of section 3, which has time asymmetry built into it, one may well argue that the time-reversal of an equilibrium state is not an equilibrium state. From this point of view, we would not be entitled to put $s = s^*$, and this analysis would fail.

This leaves us with several options. We could ignore claim (A) and merely focus on the formal structure of the axioms, in which the elements of the state space   are given little to no physical interpretation. The disadvantage of this option is that it is less clear how physically significant this symmetry is.

Alternatively, we could turn to some of the additional axioms in the Lieb-Yngvason approach (in particular, those called A6 and T4). These axioms, roughly speaking, postulate the existence of special kinds of adiabatic processes (namely, mixing and thermal equilibration) whose properties are non-invariant under the reversal of the adiabatic accessibility. In this option, the Entropy Principle is embedded in a more general time asymmetric theory.

Or, third, we could rely on the intended meaning of equilibrium and, in particular, on the validity of claim (A). In that case, the time-reversal of an equilibrium state is no longer an equilibrium state, and the earlier analysis is no longer applicable. A problem is then, that one would still like to be able to represent the distinction between those states which will remain steady in the future, and those that have remained steady in the past, in the structure of the theory. In as far as the intention of the formalism is merely to reach a description by which the equilibrium state is completely characterized by a small number

of variables as, volume, temperature, entropy and so on, it is unclear how one could achieve this.

Regardless of which option one prefers, the point we wish to emphasize is that strictly speaking, in accordance with Earman's criterion, time asymmetry is not needed to obtain a rigorous statement of the second law; it is required only in the interpretation of the notion of equilibrium.

7. An aside on the role of abstraction in thermodynamics

Nearly all attempts to axiomatise thermodynamics have been inspired by the work of Carathéodory, which has been compared with the development of Minkowski geometry in special relativity[38]. It may be useful to pursue this point further.

It is widely known that Einstein was guided to his 1905 formulation of special relativity—more specifically, his method of deriving the Lorentz transformations—by thinking about the structure of thermodynamics. Einstein realised that "constructive" attempts to determine the effects of motion on rigid rods and clocks, based on theories governing the fundamental forces holding together the constituent parts of these bodies, were hopelessly premature. He thus retreated to the secure but less ambitious position of using phenomenological principles, akin to the laws of thermodynamics, as constraints on the form of the inertial coordinate transformations[39]. As he said[40], the relativity principle, for instance, in ruling out the possibility of ever detecting the effect of inertial motion on experiments, is similar to the "restricting principle" associated with the impossibility of a certain kind of perpetual motion machine. Special relativity, like thermodynamics, is a "principle theory" in Einstein's terminology. It is clear that for Einstein a constructive account of relativistic kinematics in terms of the dynamical behaviour of rods and clocks

---

[38] Uffink (2001), section 9.

[39] Einstein (1919).

[40] Einstein (1969).

would have had greater explanatory power. But such an account could not be provided in 1905.[41]

Minkowski showed that the content of special relativity could be formulated in terms of the metrical structure of an abstract four-dimensional spacetime, with a class of privileged global coordinate charts. At almost the same time, Carathéodory showed that the principles of thermodynamics could be reformulated in the language of an abstract state space coordinatized by the state variables of the system in question, in which a local topological, rather than metric, property is connected with the statement of the Second Law. In the recent work of Lieb and Yngvason the analogy between the set of states adiabatically accessible to a given state and the future light cone associated with (set of all points causally accessible to) a given point in Minkowski spacetime is even more transparent.

The simple point we wish to make is this. The success of attempts to reformulate thermodynamics in terms of more abstract structures does not change the fact that the theory is ultimately based on phenomenological principles—in other words that it is a "principle theory". The microscopic constitution of thermodynamic systems is still of no concern. Nor is the existence of a temporal metric: dynamical equations of motion involving time derivatives are still absent. In similar fashion, Minkowski's geometric reformulation of special relativity does not alter the principle-theory status of Einstein's 1905 theory. The absolute Minkowski spacetime offers no more of a "constructive" account of relativistic kinematical effects than the topological properties of the state space do of the entropic behaviour of systems in thermodynamics. Provision of such a constructive picture in the latter case is the prerogative of statistical mechanics. In the case of special relativity, the corresponding prerogative lies with the theory of the fundamental non-gravitational interactions, which provides in principle an account of the behaviour of rods and clocks—understood now, in Einstein's words, as "moving atomic configurations".

---

[41] For further discussion of this issue, and its reappearance in the work of J.S. Bell, see Brown and Pooley (2001).

References

Barbour J. (1994) 'The Emergence of Time and Its Arrow from Timelessness', in Halliwell *et al*. (1994), pp. 405-414.

Barbour, J. (1999) *The End of Time* (London: Weidenfeld & Nicolson).

Boltzmann, L. (1895) 'On Certain Questions of the Theory of Gases', *Nature* **52**, 413-415.

Boltzmann, L. (1897a) 'Zu Hrn Zermelos Abhandlung "Über die mechanische Erklärung irreversibler Vorgänge"' *Wied. Ann*. **60**, 392-398.

Boltzmann, L. (1897b) 'Über einige meiner weniger bekannten Abhandlungen uber Gastheorie und deren Verhältnis zu derselben' *Verh. der 69. Vers. D. Naturf. und Ärzte,* Braunschweig, p.19-26.

Boltzmann, L. (1898) *Vorlesungen über Gastheorie,* Vol. II, (Leipzig: Barth)

Brown, H.R. and Pooley, O. (2001) 'The origins of the spacetime metric: Bell's Lorentzian pedagogy and its significance in general relativity', in C. Callender and N. Huggett (eds.) *Physics Meets Philosophy at the Planck Scale* (Cambridge: Cambridge University Press) pp. 256-272. e-Print: gr-qc/9908048.

Brown, H.R. and Sypel, R. (1995) 'On the meaning of the relativity principle and other symmetries', *International Studies in the Philosophy of Science* **9**, 235-253.

Budden, T. (1997) 'Galileo's Ship and Spacetime Symmetry', *British Journal for the Philosophy of Science* **48**(4), 483-516.

Carathéodory, C. (1909) 'Untersuchungen über die Grundlagen der Thermodynamik', *Mathematische Annalen* **67**, 355-386. English translation in Kestin (1976), pp. 229-256.

Carathéodory, C. (1925) 'Über die Bestimmung der Energie und die absoluten Temperatur mit Hilfe von reversiblen Prozessen', *Sitzungsberichte der Preussischen Akademie der Wissenschaften* pp. 39-47.

Davies, P. (1999) 'Is the flow of time an illusion?', *International Framtider* **9**, 4-8.

Dugdale, J.S. (1996) *Entropy and its Physical Meaning* (London: Taylor and Francis Ltd.)

Earman, J. (1974) 'An attempt to add a little direction to "the problem of the direction of time" ', *Philosophy of Science* **41**, 15-47.

Einstein, A. (1919) 'My theory', *The London Times, November 28*, p. 13. Reprinted as 'What is the Theory of Relativity?', in A. Einstein *Ideas and Opinions* (New York: Crown Publishers, Inc.), pp. 227-232.

Einstein, A. (1969) 'Autobiographical Notes', in P.A. Schilpp (ed.) *Albert Einstein: Philosopher-Scientist*, Vol 1 (Illinois: Open Court), pp. 1-94.

Halliwell, J.J., Pérez-Mercador, J. and Zurek, W.H. (eds.) (1994) *Physical Origins of Time Asymmetry* (Cambridge: Cambridge University Press).

Hawking, S. (1994), 'The No Boundary Condition And The Arrow Of Time', in Halliwell *et al.* (1994), pp. 346-357.

Hatsopoulos G.N. and Keenan J.H. (1962) 'A Single Axiom for Classical Thermodynamics', *Journal of Applied Mechanics* **29**, 193-199.

Hatsopoulos G.N. and Keenan J.H. (1965) *Principles of General Thermodynamics* (New York: John Wiley & Sons Inc.).

Kestin, J. (1963) Contribution to 'Discussion: A single axiom for classical thermodynamics', *Journal of Applied Mathematics* **30**, 313-314.

Kestin, J. (1976) *The Second Law of Thermodynamics* (Stroudsberg, Pennsylvania: Dowden, Hutchison and Ross, Inc.)

Kestin, J. (1979) *A Course in Thermodynamics*. Vols. I, II (New York: McGraw-Hill).

Lebowitz, J. (1993) 'Macroscopic laws and microscopic dynamics' *Physica* **A194**, 1-27.

Lebowitz, J. (1994), 'Time's Arrow and Boltzmann's Entropy', in Halliwell *et al.* (1994), pp. 131-146.

Lieb, E. and Yngvason, J. (1998) 'A Guide to Entropy and the Second Law of Thermodynamics', *Notices of the American Mathematical Society* **45** (5), 571--581.

Lieb, E. and Yngvason, J. (1999) 'The physics and mathematics of the second law of thermodynamics', *Physics Reports* **310**, 1-96, erratum **314** (1999), 669. e-Print: cond-mat/9708200.

Lieb, E. and Yngvason, J. (2000) 'A Fresh Look at Entropy and the Second Law of Thermodynamics', *Physics Today*, April issue; e-Print: math-ph/0003028.

Lith, J. van, (1999) 'Reconsidering the concept of equilibrium in classical statistical mechanics', *Philosophy of science* **66** (Proceedings), S107-S118.

Price, H. (1996) *Time's Arrow and Archimedes' Point* (New York, Oxford University Press).

Ridderbos, T. (2000) 'The Arrow of Time in Statistical Mechanics and Quantum Cosmology', Ph.D. thesis, Cambridge University.

Schrödinger, E. (1933) ' Über den zweiten Haptsatz der Thermodynamik', *Sitzingsberichte de Preussischen Akademie der Wissenschaften Physikalisch-mathematische Klasse*, 165.

Schrödinger, E. (1950) 'Irreversibility', *Proceedings of the Royal Irish Academy* **53**, 189-195.

Schulman, L.S, (1997) *Time's arrows and quantum measurement* (Cambridge: Cambridge University Press).

Sklar, L. (1993) *Physics and Chance* (Cambridge: Cambridge University Press).

Uffink, J (2001) 'Bluff your way in the Second Law of Thermodynamics', *Studies in the History and Philosophy of Modern Physics*, to appear.

Uhlenbeck, G.E. and Ford, G.W. (1963) 'Lectures in Statistical Mechanics', in *Lectures in Applied Mathematics,* vol I, Mark Kac (ed.) (Providence, Rhode Island: American Mathematical Society).

Zeh, H.-D. (1992) *The Physical Basis of the Direction of Time* (New York: Springer-Verlag).

Zemansky M.W. (1957) *Heat and Thermodynamics* (Tokyo: Kogakusha Co. Ltd.; New York: McGraw-Hill Book Company Inc.)