

# Three Concepts of Irreversibility and Three Versions of the Second Law

JOS UFFINK, UTRECHT

This paper aims to clarify the relation between the second law of thermodynamics and the notion of irreversibility by distinguishing three different meanings of the latter and to study how they figure in three versions of the second law of thermodynamics. A more extensive discussion is given in (Uffink 2001).

## 1. Three concepts of (ir)reversibility

Many physical theories employ a state space  $\Gamma$  containing all possible states  $s$  of a system. A process is then represented as a parameterised curve:

$$\mathcal{P} = \{s_t \in \Gamma : t_i \leq t \leq t_f\}.$$

Usually a theory allows only a subclass, say  $\mathcal{W}$ , of such processes (e.g. the solutions of the equations of motion). Let  $R$  be an involution (i.e.  $R^2s = s$ ) that turns state  $s$  into its ‘time reversal’  $Rs$ . In classical mechanics, for example,  $R$  is the transformation which reverses the sign of all momenta and magnetic fields. In a theory like classical thermodynamics, where the state does not contain velocity-like parameters, one may take  $R$  to be the identity. Further, define the time reversal  $\mathcal{P}^*$  of process  $\mathcal{P}$  by:

$$\mathcal{P}^* = \{(Rs)_{-t} : -t_f \leq t \leq -t_i\}. \quad (1)$$

The theory is called *time-reversal invariant* (TRI) if the class  $\mathcal{W}$  is closed under time reversal, i.e. iff:

$$\mathcal{P} \in \mathcal{W} \Rightarrow \mathcal{P}^* \in \mathcal{W}.$$

According to this definition the form of the laws (and a choice for  $R$ ) determines whether a theory is TRI or not. Note that it is irrelevant here whether the processes  $\mathcal{P}^*$  actually occur, but only that the theory allows them. Thus,

the fact that the sun never rises in the West does not mean that celestial mechanics is non-TRI.

Is time-reversal (non)invariance related to the second law? An application of the criterion to thermodynamics is not a matter of routine. In contrast to mechanics, thermodynamics does not possess equations of motion. This is because thermodynamical processes only occur after external intervention on the system (e.g.: removing a partition, pushing a piston, etc.) and do not reflect autonomous behaviour of a free system. This is not to say that time plays no role. Classical thermodynamics, in the formulation of Clausius, Kelvin or Planck, is concerned with processes, and its second law is clearly not TRI. However, in other formulations, e.g. by Gibbs, Carathéodory, or Lieb and Yngvason, this is less clear.

Now, the term '*(ir)reversible*' is usually attributed to processes rather than theories or laws. But in philosophy of physics, it is intimately connected with time-reversal invariance. Indeed, one calls a process  $\mathcal{P}$  allowed by a given theory irreversible iff the reversed process  $\mathcal{P}^*$  is excluded by this theory. Obviously, such a process  $\mathcal{P}$  exists only if the theory in question is not TRI. Conversely, every non-TRI theory admits irreversible processes in this sense. Therefore, discussions about (ir)reversibility and (non)-TRI in philosophy of physics mostly coincide.

However, the thermodynamics literature often uses the term 'irreversibility' to denote an aspect of processes which one might also call *irrecoverability*. In many processes, the transition from an initial state  $s_i$  to a final state  $s_f$ , cannot be fully 'undone', once the process has taken place. In other words, there is no process which starts off from state  $s_f$  and restores the initial state  $s_i$  completely. Wear and tear, erosion, etc. are the obvious examples. This meaning of the term goes back to Kelvin (1852), actually the first author to employ the term in a thermodynamical context. It is also this sense of irreversibility that Planck intended, when he called it the essence of the second law.

Many writers have emphasised irrecoverability in connection with the second law. Indeed, Eddington introduced his famous phrase of 'the arrow of time' while discussing the 'running-down of the universe', and illustrated it by examples involving 'irrevocable changes', including the case of Humpty-Dumpty who, allegedly, could not be put together again after his great fall. In retrospect, one might say that a better expression for this theme is the *ravages* of time rather than its arrow.

(Ir)recoverability differs from (non)-TRI in at least two respects. First, the only thing that matters here is the retrieval of the initial state  $s_i$ . It is not

necessary to specify a process  $\mathcal{P}^*$  retracing the intermediate stages of the original process in reverse order. A second difference is that we are dealing with a *complete* recovery. This means that all auxiliary systems that may have been used in the original process are also returned to their initial state.

A schematic expression of the idea is this. Let  $s$  be a state of the system and  $Z$  a (formal) state of its environment. Let  $\mathcal{P}$  be some process that brings about the transition:

$$\langle s_i, Z_i \rangle \xrightarrow{\mathcal{P}} \langle s_f, Z_f \rangle \quad (2)$$

Then  $\mathcal{P}$  is reversible in Planck's sense iff there exists another process  $\mathcal{P}'$  that produces

$$\langle s_f, Z_f \rangle \xrightarrow{\mathcal{P}'} \langle s_i, Z_i \rangle \quad (3)$$

The term 'reversible' is also used in a third sense, to denote processes which proceed so slowly that the system remains in equilibrium 'up to a negligible error' during the entire process. This is the meaning embraced by Clausius, and it appears to be the most common usage of the term in the physical-chemical literature; see e.g. Hollinger and Zenzen (1985), Denbigh (1989). A more apt name for this kind of processes is *quasi-static*. Of course, the above characterisation is vague, and has to be amended by criteria specifying what 'errors' are intended and when they are 'small'. These criteria take the form of a limiting procedure so that, strictly speaking, reversibility is here not an attribute of a particular process but of a series of processes.

Quasi-static processes need not be the same as those called reversible in the previous two senses. E.g., an ideal harmonic oscillator is reversible in Planck's sense, but not quasi-static. Conversely, the discharge of a condenser through a high resistance can be made to proceed quasi-statically, but even then it remains irreversible in Planck's sense.

Comparison with the notion of TRI is hampered by the fact that 'quasi-static' is not strictly a property of a process. Consider a process  $\mathcal{P}_N$  in which a system, originally at temperature  $\theta_1$  is consecutively placed in thermal contact with a sequence of  $N$  heat baths, each at a slightly higher temperature than the previous one, until it reaches a temperature  $\theta_2$ . By making  $N$  large, and the temperature steps small, such a process becomes quasi-static, and we can represent it by a curve in the space of equilibrium states. However, for any  $N$ , the time-reversal of the process is impossible.

Nevertheless, many authors call such a curve ‘reversible’, because one can consider another process  $\mathcal{Q}_N$ , in which the system, originally at temperature  $\theta_2$ , is placed into contact with a series of heat baths, each slightly *colder* than the previous one. Again, each process  $\mathcal{Q}_N$  is non-TRI. *A fortiori*, no  $\mathcal{Q}_N$  is the time reversal of any  $\mathcal{P}_N$ . Yet, if we now take the quasi-static limit, the state change of the system will follow the same curve in equilibrium space as in the previous case, traversed in the opposite direction. The point is, of course, that precisely because this curve is not itself a process, the notion of time reversal does not apply to it.

## 2. The second law according to Planck

Planck’s position has always been that the second law expresses irrecoverability of all processes in nature. However, it is not easy to analyse his arguments for this claim. Various editions of his book (Planck 1897) differ in many decisive details. Also, the English translation is unreliable. It uses the term ‘reversible’ indiscriminately, where Planck distinguishes between *umkehrbar*, which he uses in Clausius’ sense, i.e. meaning ‘quasi-static’, and *reversibel*, in the sense of Kelvin (1852) meaning ‘recoverable’. Moreover, he presented a completely different argument from the eighth edition onwards.

I shall only mention Planck’s latter argument, published first in Planck (1926). He starts from the statement that “friction is an *irreversibel* process”, which he considers to be an expression of Kelvin’s principle.<sup>1</sup> He then considers an adiabatically isolated fluid capable of exchanging energy with its environment by means of a weight at height  $h$ . Planck asks whether it is possible to bring about a transition from an initial state  $s$  of this system to a final state  $s'$ , in a process which brings about no changes in the environment other than the displacement of the weight. If  $Z$  denotes the state of the environment and  $h$  the height of the weight, the desired transition can be represented as

$$(s, Z, h) \xrightarrow{?} (s', Z, h').$$

---

1 This may need some explanation because, at first sight, this statement does not concern cyclic processes or the *perpetuum mobile* at all. But for Planck, the statement means that there exists no process which ‘undoes’ the consequences of friction, i.e., a process which produces no other effect than cooling a reservoir and doing work. The condition ‘no other effect’ here allows for the operation of any type of auxiliary system that operates in a cycle.

He argues that, by means of ‘*reversibel-adiabatic*’<sup>2</sup> processes, one can always achieve a transition from the initial state  $s$  to an intermediary state  $s^*$  in which the volume equals that of state  $s'$  and the entropy equals that of  $s$ . That is, one can realise a transition

$$(s, Z, h) \rightarrow (s^*, Z, h^*), \quad \text{with} \quad V(s^*) = V(s') \quad \text{and} \quad S(s^*) = S(s).$$

Whether the desired final state  $s'$  can now be reached from the intermediate state  $s^*$  depends on the value of the only independent variable in which  $s^*$  and  $s'$  differ. For this variable one can choose the energy  $U$ .

There are three cases:

- (1)  $h^* = h'$ . In this case, energy conservation implies  $U(s^*) = U(s')$ . Because the coordinates  $U$  and  $V$  determine the state of the fluid completely,  $s^*$  and  $s'$  must coincide.
- (2)  $h^* > h'$ . In this case,  $U(s^*) < U(s')$ , and the state  $s'$  can be reached from  $s^*$  by letting the weight perform work on the system, e.g. by means of friction, until the weight has dropped to height  $h'$ . According to the above formulation of Kelvin’s principle, this process is irreversible (i.e. irrecoverable).
- (3)  $h^* < h'$  and  $U(s^*) > U(s')$ . In this case the desired transition is impossible. It would be the reversal of the irreversible process just mentioned, and thus realise a *perpetuum mobile* of the second kind. Now, Planck argues that in all three cases, one can also achieve a transition from  $s^*$  to  $s'$  by means of heat exchange in an *umkehrbar* (i.e. quasi-static) process in which the volume remains fixed. For such a process he writes

$$dU = TdS. \tag{4}$$

Using the assumption that  $T > 0$ , it follows that, in the three cases above,  $U$  must vary in the same sense as  $S$ . That is, the cases  $U(s^*) < U(s')$ ,  $U(s^*) = U(s')$  or  $U(s^*) > U(s')$ , can also be characterised as  $S(s^*) < S(s')$ ,  $S(s^*) = S(s')$  and  $S(s^*) > S(s')$  respectively.

An analogous argument can be constructed for a system consisting of several fluids. Just as in earlier editions of his book, Planck generalises the

---

2 Apparently, Planck’s pen slipped here. He means: *umkehrbar-adiabatic*.

conclusion (without a shred of proof) to arbitrary systems and arbitrary physical/chemical processes:

Every process occurring in nature proceeds in the sense in which the sum of the entropies of all bodies taking part in the process is increased. In the limiting case, for reversible processes this sum remains unchanged. [...] This provides an exhaustive formulation of the content of the second law of thermodynamics (Planck 1926, p. 463)

Planck's argument can hardly be regarded as satisfactory for the bold and universal formulation of the second law. It applies only to systems consisting of fluids, and relies on several implicit assumptions which can be questioned outside of this context. In particular, this holds for the assumption that there always exist functions  $S$  and  $T$  (with  $T > 0$ ) such that  $\bar{d}Q = TdS$ , where  $\bar{d}$  denotes an inexact differential; and the assumption of a rather generous supply of quasi-static processes.

### 3. The second law according to Carathéodory

Carathéodory (Carathéodory 1909) construed thermodynamics as a theory of equilibrium states rather than (cyclic) processes. A thermodynamical system is described by a state space  $\Gamma$ , represented as a (subset of a)  $n$ -dimensional manifold with the state variables serving as coordinates. He assumes that  $\Gamma$  is equipped with the standard Euclidean topology. But metrical properties do not play a role, and there is no preference for a particular system of coordinates. The fundamental concept is a relation called *adiabatic accessibility*, which represents whether state  $t$  can be reached from state  $s$  in an adiabatic process,<sup>3</sup> and the second law is formulated as follows:

**CARATHÉODORY'S PRINCIPLE:** In every open neighbourhood  $U_s \subset \Gamma$  of every state  $s$  there are states  $t$  such that for some open neighbourhood  $U_t$  of  $t$ : all states  $r$  within  $U_t$  cannot be reached adiabatically from  $s$ .

---

3 Carathéodory calls a container adiabatic if the system contained in it remains in equilibrium, regardless of what occurs in the environment, as long as the container is not moved nor changes its shape. Thus, the only way of inducing a process of a system in an adiabatic container is by deformation of the walls. (E.g. a change of volume or stirring.) A process is said to be adiabatic if it takes place while the system is in an adiabatic container.

He then introduces the so-called ‘simple systems’ (defined by four additional conditions) and obtains

CARATHÉODORY’S THEOREM: For simple systems, Carathéodory’s principle is equivalent to the proposition that the differential form<sup>4</sup>  $\delta Q := dU - \delta W$  possesses an integrable divisor, i.e. there exist functions  $S$  and  $T$  on the state space  $\Gamma$  such that

$$\delta Q = TdS. \quad (5)$$

Thus, for simple systems, every equilibrium state can be assigned a value for entropy and absolute temperature. Curves representing quasi-static adiabatic changes of state are characterised by the differential equation  $\delta Q = 0$ , and by virtue of (5) one can conclude that (if  $T \neq 0$ ) entropy remains constant. Obviously  $S$  and  $T$  are not uniquely determined by the relation (5). Carathéodory discusses further conditions to determine the choice of  $T$  and  $S$  up to a constant of proportionality.

I want to mention a number of strong and weak points of the approach. An advantage is that it provides a mathematical formalism for thermodynamics, comparable to relativity theory. There, Einstein’s original approach, which starts from empirical principles (the light postulate and relativity principle), was replaced by an abstract Minkowski spacetime, where the empirical principles are incorporated in local properties of the metric. Similarly, Carathéodory constructs an abstract state space where an empirical statement of the second law is converted into a local topological property. Furthermore, all coordinate systems are treated on the same footing (as long as there is only one thermal coordinate, and they generate the same topology). Note further that the environment of the system is never mentioned explicitly in his treatment. This too is a conceptual advantage.

But Carathéodory’s work has also provoked objections. Many complain that the absence of an explicit reference to a *perpetuum mobile* obscures the physical content of the second law. Other problems in Carathéodory’s approach concern the additional assumptions needed to obtain the result (5), i.e. the restriction to simple systems. Further, Carathéodory’s proof actually establishes merely the *local* existence of functions  $S$  and  $T$  obeying (5). This

---

4 Here the notation  $\delta$  is used to denote that  $\delta Q$  and  $\delta W$  are not exact differentials, in contrast to  $dU$  and  $dS$ .

does not guarantee the existence of a pair of functions, defined globally on  $\Gamma$ , that obey (5).

For the present purpose, we need to investigate whether and how this work relates to ‘irreversibility’. Carathéodory conceives of thermodynamics as a theory of equilibrium states, rather than processes. But his concept of ‘adiabatic accessibility’ does refer to processes between equilibrium states.

In order to judge the time-reversal invariance of the theory of Carathéodory one must specify a time reversal transformation  $R$ . It seems natural to choose this in such a way that  $R_s = s$  and  $R(\prec) = \succ$ . Then Carathéodory’s principle is *not* TRI. Indeed, the principle forbids that  $\Gamma$  contains a state  $s$  from which one can reach all states in some neighbourhood of  $s$ . It allows models where a state  $s$  exists from which one can reach no other state in some neighbourhood. Time reversal of such a model violates Carathéodory’s principle. However, this non-invariance manifests itself only in rather pathological cases. If we exclude them, Carathéodory’s theory becomes TRI.

It is easy to show that Carathéodory’s approach fails to capture the content of the second law à la Planck, namely by exhibiting models of his formalism in which this version of the second law is invalid. An example is obtained by swapping the meaning of terms in each of the three pairs ‘heat/work’, ‘thermal/deformation coordinate’ and ‘adiabatic’/‘without any exchange of work’. The validity of Carathéodory’s formalism is invariant under this operation for fluids. Indeed, we obtain, as a direct analog of (5):  $dW = p dV$  for all quasistatic processes of a fluid. Thus pressure and volume here play the role of temperature and entropy respectively. Further, irreversibility makes sense here too. For fluids with positive pressure, one can increase the volume of a fluid without doing work, but one cannot decrease volume without doing work. But still, the analog of the principles of Clausius or Kelvin are false: A fluid with low pressure can very well do positive work on another fluid with high pressure by means of a lever or some hydraulic mechanism. And, thus, the sum of all volumes of a composite system can very well decrease, even when no external work is provided.

#### 4. The second law according to Lieb and Yngvason

Lieb and Yngvason (1999) have provided a rigorous approach to the second law. I can only sketch those main ideas that are relevant to my topic.



Formally, their work builds upon the approach of Carathéodory (1909) and Giles (1964). (In its physical interpretation, however, it is more closely related to Planck.) A system is represented by a state space  $\Gamma$  on which a relation  $\prec$  of adiabatic accessibility is defined. Further, one may combine two systems in state  $s$  and  $t$  into a composite system in state  $(s, t)$ , and there is an operation of ‘scaling’, i.e. the construction of a copy in which all its extensive quantities are increased by a factor  $\alpha$ . This is denoted by a multiplying the state with  $\alpha$ . The main axioms read:

A1. REFLEXIVITY:  $s \prec s$

A2. TRANSITIVITY:  $s \prec t$  and  $t \prec r$  imply  $s \prec r$

A3. CONSISTENCY:  $s \prec s'$  and  $t \prec t'$  implies  $(s, t) \prec (s', t')$

A4. SCALE INVARIANCE: If  $s \prec t$  then  $\alpha s \prec \alpha t$  for all  $\alpha > 0$

A5. SPLITTING AND RECOMBINATION:

For all  $0 < \alpha < 1$  :  $s \prec (\alpha s, (1 - \alpha)s) \prec s$

A6. STABILITY: If there are states  $t_0$  and  $t_1$  such that  $(s, \varepsilon t_0) \prec (r, \varepsilon t_1)$  holds for a sequence of  $\varepsilon$ 's converging to zero, then  $s \prec r$ .

7. COMPARABILITY HYPOTHESIS:

For all states  $s, t$  in the same  $\Gamma$ :  $s \prec t$  or  $t \prec s$ .<sup>5</sup>

The comparability hypothesis is intended to characterise a particular type of ‘simple’ systems.<sup>6</sup> A substantial part of their paper is devoted to to derive this hypothesis from further axioms. I will, however, not go into this.

The central aim is to derive the following result, which Lieb and Yngvason call

5 The clause ‘in the same  $\Gamma$ ’ means that the hypothesis is not intended for the comparison of states of scaled systems. Thus, it is not demanded that we can either adiabatically transform a state of 1 mole of oxygen into one of 2 moles of oxygen or conversely.

6 Beware that the present meaning of the term does not coincide with that of Carathéodory. For simple systems in Carathéodory’s sense the comparability hypothesis need not hold.

THE ENTROPY PRINCIPLE: There exists a function  $S$  defined on all states of all systems such that when  $s$  and  $t$  are comparable then

$$s \prec t \text{ iff } S(s) \leq S(t). \quad (6)$$

The question whether this result actually follows is somewhat involved. They show that the entropy principle follows from axioms A1–A6 and the comparability hypothesis under some special conditions which, physically speaking, exclude mixing and chemical reactions. To extend the result, an additional ten axioms are needed. And even then, only a weak form of the above entropy principle is actually obtained, where ‘iff’ in (6) is replaced by ‘implies’.

Note that the theorem is obtained without appealing to Carathéodory’s principle. In fact the axioms and hypothesis mentioned above allow models which violate the principle of Carathéodory (Lieb and Yngvason 1999, p. 91). For my purpose, the question is what connection with irreversibility is in this formulation of the second law. As before, there are two aspects to this question: irrecoverability and time-reversal (in)variance. We have seen that Lieb and Yngvason interpret the relation (6) as saying that entropy must increase in irreversible processes. At first sight, this interpretation is curious. Adiabatic accessibility is not the same thing as irreversibility. So how can the above axioms have implications for irreversible processes?

This puzzle is resolved by looking at the physical interpretation Lieb and Yngvason propose for  $\prec$ :

ADIABATIC ACCESSIBILITY: A state  $t$  is adiabatically accessible from a state  $s$ , in symbols  $s \prec t$ , if it is possible to change the state from  $s$  to  $t$  by means of an interaction with some device (which may consist of mechanical and electric parts as well as auxiliary thermodynamic systems) and a weight, in such a way that the auxiliary system returns to its initial state at the end of the process whereas the weight may have changed its position in a gravitational field (Lieb and Yngvason 1999, p. 17).

This view differs from Carathéodory’s, or indeed, anybody else’s: clearly, this term is not intended to refer to processes occurring in a thermos flask. Even processes in which the system is *heated* are adiabatic, in the present sense, when this heat is generated by an electrical current from a dynamo driven by a descending weight. Actually, the condition that the auxiliary systems

return to their initial state in the present concept is strongly reminiscent of Planck's concept of 'reversible'!

This is not to say, of course, that they are identical. As we have seen before, a process  $\mathcal{P}$  involving a system, an environment and a weight at height  $h$ , which produces the transition  $\langle s, Z, h \rangle \xrightarrow{\mathcal{P}} \langle s', Z', h' \rangle$  is reversible for Planck iff there exists a 'recovery' process  $\mathcal{P}'$  which produces  $\langle s', Z', h' \rangle \xrightarrow{\mathcal{P}'} \langle s, Z, h \rangle$ . Here, the states  $Z$  and  $Z'$  may differ from each other. For Lieb and Yngvason, a process  $\langle s, Z, h \rangle \xrightarrow{\mathcal{P}} \langle s', Z', h' \rangle$  is adiabatic iff  $Z = Z'$ . But in all his discussions, Planck always restricted himself to such reversible processes 'which leave no changes in other bodies', i.e. obey the additional requirement  $Z = Z'$ . These processes are adiabatic in the present sense.

A crucial consequence is that, in the present sense, it follows that if a process  $\mathcal{P}$  as considered above is adiabatic, any recovery process  $\mathcal{P}'$  is automatically adiabatic too. Thus, we may conclude that if an adiabatic process is accompanied by an entropy increase, it cannot be undone, i.e., it is irreversible in Planck's sense. This explains why the result (6) is seen as a formulation of a principle of entropy increase. Thus we obtain the conclusion implying the existence of irrecoverable processes by means of a satisfactory argument!

However, note that this conclusion is obtained from the comparability hypothesis and by excluding mixing and chemical processes. The weak version of the entropy principle, which is derived when we drop the latter restriction, does not justify this conclusion. Moreover, note that it would be incorrect to construe (6) as a characterisation of *processes*. The relation  $\prec$  is interpreted in terms of the *possibility* of processes. Thus, when  $S(s) < S(t)$  for comparable states, this does not mean that *all* processes from  $s$  to  $t$  are irreversible, but only that there exists an adiabatic irreversible process between these states. So the entropy principle here is not the universal proposition of Planck. The fact that it is not necessary to introduce time-reversal non-invariance into the formalism to obtain the second law, is very remarkable.

However, there remains one problematical aspect of the proposed physical interpretation. It refers to the state of auxiliary systems in the environment of the system. Thus, we are again confronted by the old question, when shall we say that the state of such auxiliary systems has changed, and when are we fully satisfied that their initial state is restored. This question remains intractable from the point of view of thermodynamics, as long as the states of arbitrary auxiliary systems (e.g. living beings) are not representable in the

thermodynamical formalism. Thus, the question when the relation  $\prec$  holds cannot be decided in thermodynamical terms.

## 5. Discussion

There is a large variety in the views on irreversibility and the second law. On one end, there is Planck's view that the second law expresses the irreversibility of all processes in Nature. A convincing derivation of this bold claim has, however, never been given. On the other extreme is Gibbs' approach, which completely avoids any connection with time.

But even for approaches in the middle ground, the term 'reversible' is used in various meanings: time-reversal invariant, recoverable, and quasi-static. In the debate on the question how the second law relates to statistical mechanics, however, most authors have taken irreversibility in the sense of time-reversal non-invariance. The point that in thermodynamics the term usually means something very different has been almost completely overlooked.

The formal approaches by Carathéodory and Lieb and Yngvason show that it is possible to build up a precise formulation of the second law without introducing a non-TRI element. The resulting formalism implies only that an entropy function can be constructed consistently, i.e. as either increasing between adiabatically accessible states of *all* simple systems, or decreasing. At the same time, the Lieb-Yngvason approach does imply that entropy increasing processes between comparable states are irreversible in Planck's sense. This shows once more the independence of the two notions.

Finally, I would like to point out an analogy between the axiomatisation of thermodynamics in the Carathéodory and Lieb-Yngvason approach and that of special relativity in the approach of Robb (1921). In both cases, we start out with a particular relationship  $\prec$  which is assumed to exist between points of a certain space. In relativity, this is the relation of connectability by a causal signal. In both cases, it is postulated that this relation forms a pre-order. In both cases, important partial results show that the forward sectors  $\mathcal{C}_s = \{t : s \prec t\}$  are convex and nested and that  $s$  is on the boundary of  $\mathcal{C}_s$ . And in both cases the aim is to show that the space is 'orientable' (Earman 1967) and admits a global function which increases in the forward sector. If this analogy is taken seriously, the Lieb-Yngvason entropy principle has just as much to do with TRI as the fact that Minkowski space-time admits a global time coordinate.

## References

- Carathéodory, Constantin 1909 “Untersuchungen über die Grundlagen der Thermodynamik”, *Mathematische Annalen*, 67, 355–386, English translation by Joseph Kestin, “Investigation into the foundations of thermodynamics” in: Joseph Kestin (ed.), 1976 *The second law of thermodynamics*, Stroudsburg, PA: Dowden, Hutchinson and Ross, 229–256. (This translation is not quite accurate.)
- Denbigh, Kenneth G. 1989 “The Many Faces of Irreversibility”, *British Journal for Philosophy of Science*, 40, 501–518.
- Earman, John 1967 “Irreversibility and temporal asymmetry”, *Journal of Philosophy*, 64, 543–549.
- Eddington, Arthur 1935 *The Nature of the Physical World*, London: J.M. Dent & Sons.
- Giles, Robin 1964 *Mathematical Foundations of Thermodynamics*, Oxford: Pergamon.
- Hollinger, Henry B, and Zenzen, Michael J. 1985 *The Nature of Irreversibility*, Dordrecht: Reidel, 1985.
- Kelvin 1852 “On a universal tendency in nature to the dissipation of mechanical energy”, in Joseph Kestin (ed.), 1976 *The second law of thermodynamics*, op.cit., 194–198.
- Lieb, Elliott and Yngvason, Jakob 1999 “The physics and mathematics of the second law of thermodynamics”, *Physics Reports*, 310, 1–96, erratum, 314, (1999), 669.
- Planck, Max 1897 *Vorlesungen über Thermodynamik*, Leipzig: Veit & Comp. — 1926 “Über die Begründung des zweiten Hauptsatzes der Thermodynamik”, *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, 453–463.
- Robb, Alfred A. 1921 *The Absolute Relations of Time and Space*, Cambridge: Cambridge University Press.
- Uffink, Jos 2001 “Bluff your way in the second law of thermodynamics”. *Studies in History and Philosophy of Modern Physics*, 32, 305–394.