Review of Earth System Dynamics (ESD) Manuscript #2019-6

Title: Thermodynamic optimality in Earth sciences. The missing constraints in modeling Earth system dynamics?

Author: Westhoff et al

Review

The manuscript describes recent experience in using thermodynamic optimality principles (TOP) in the earth sciences.

The manuscript asserts that while TOP has been applied previously, that there has been much scepticism about the application of TOP because of [quoted from p. 3, lines 7-14]:
"*1) inconsistencies in the use of thermodynamic concepts and terminology; 2) the lack of a theory postulating a priori the respective ranges of applicability of the different TOPs Martyushev and Seleznev (2014) identified this issue as one of the main reasons for criticism on TOPs – and listed several restrictions of the maximum entropy production principle; 3) apparently arbitrary choices of processes and boundaries considered in systems with many interlinked processes producing power and entropy (Volk and Pauluis, 2010). 4) a lack of a widely accepted physical foundation that explains why a system should evolve to a maximum/minimum; and 5) claims that the MEP principle is an inference principle, making it a hypothesis that cannot be rejected (because you can always argue that a rejection is caused by a missing constraint: see e.g. Ross et al., 2012).*"

I agree with the authors that this is a reasonable summary of the current situation with the TOP field and the existing scepticism. However, on reading the article it was not clear what the overall approach was to address that scepticism.

On the one hand, the article begins with section 2 that describes the foundations of the topic. Figure 1 (but see below) was a potentially good start as was the preparation of Table 1 listing a number of previous studies and highlighting the foundations of those studies. However, on reading later sections (2, 3, 4), the level of synthesis did not emerge (for me at least) and these later parts tended to focus on a comprehensive description of what had been done. The description is too comprehensive and to this reader, I ended up being just as confused as when I started reading the manuscript.

In that respect, given the previous criticisms (e.g. Goody 2007, Ross et al 2012, J of Physical Chemistry, and others) I do not think that the current manuscript has adequately addressed its first principle aim of clarifying terminology and concepts.

To me at least, the problem here is that the authors have to decide whether (i) to produce a comprehensive list of all previous applications, or (ii) attempt to synthesise just a few of the previous studies into a coherent framework. I would strongly advocate for the latter option (ii) since I do not see the value of another comprehensive review given the previous high profile criticisms/scepticisms (e.g. Goody 2007, Ross et al 2012). On the latter synthesis approach advocated above, what is critical here is to have a strong foundation of language and formalism and to use that foundation to re-examine just a few key (previous) applications and attempt to cast those into a common intellectual framework. In particular, I was looking to Figure 1 for that formalism but after repeatedly reading the relevant text and studying Fig. 1, I realised that I did not understand either the figure or the related text.

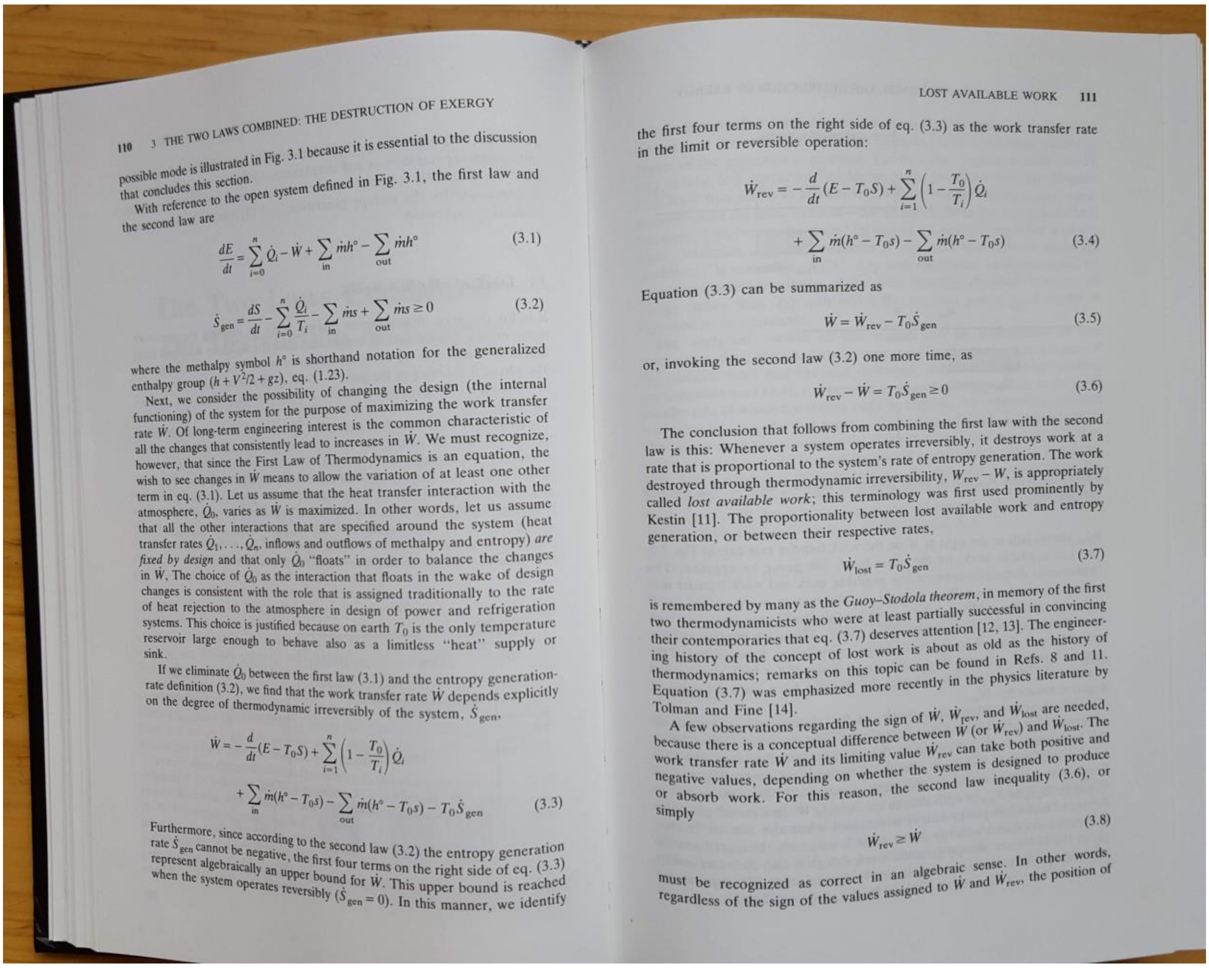
To see my difficulty, let us start with Fig. 1a and compare that with Fig. 1b. In Fig. 1a we have three heat fluxes **Jin**, **J** (connecting the two boxes) and **Jout**. I originally thought that all would presumably have units of J (i.e., Joules) as those are the conventional units for heat (and work and energy change). However, there is no work shown on figure 1a (or on figure 1b). Now go to Figure 1b. We have a finite gap where "power" is ?done? (instead of work done) and this power is then dissipated to heat that is re-absorbed and leaves the entire systems+surroundings at a fixed temperature. While not given, the units for the power term (denoted with symbol G) are presumably J s^{-1}, i.e., power is the rate of doing work and so it has a time component. I imagine the dissipation (symbol used = D) must have the same units as those for G (J s^{-1}). But those units are different from the heat fluxes (which all have units of J and not J s^{-1}). Nowhere does either figure state where the work was done? Critically work must be done at the boundary and any work transfer is not associated with entropy change.

The same problem applies to the other two panels (Fig. 1cd, 1ef) of Fig. 1.

(An alternative interpretation is that the units of heat are J s^{-1}, and the diagrams are dimensionally correct. But on that interpretation this is not necessarily a carnot cycle!)

Now compare that with the conventional thermodynamic basis that many readers will have learnt in their earlier studies. The carnot cycle is a cycle combining 4 separate but reversible processes, that take a system from a defined state through a progression of different states (i.e., through a quasi-static set of changes) before returning to the starting state. The reversible processes are accompanied by flows of heat and work and when summed over the entire cycle they perform the maximum work possible such that any other sequence of states that return to the system to the start will have done less work. Of course the total work done plus the heat absorbed must equal zero around the cycle – that is how it is defined. Again, as I am sure the authors well know, that is all very interesting but the reversible processes are done in an infinitely long amount of time so this maximum work is not actually available and not very useful. Again, as the authors well know, any real process will be irreversible (e.g. friction examples given) and will produce less than the maximum possible work (of a reversible process) but the irreversible process can be completed in finite time. Hence in any real process we have to consider time if we want to consider irreversibility. Further, for an irreversible process, the change in entropy of the system plus surroundings must be greater than zero. However, this does not automatically mean that the entropy of the system will change in a particular way – for example, if heat flowed out of a closed system then the entropy will decrease (while the entropy of the surroundings increases). Importantly, entropy is associated with heat transfers but not with work transfers. This is the basic understanding that many potential readers will begin with. My suggestion is that if you want to start with a carnot cycle then you have to relate what you are trying to do to this classical understanding of a carnot cycle. Here I am talking about a series of quasi-static processes with constant mass over one cycle. Without establishing that underlying link to the readership, you risk losing (and ultimately confusing) people (like me) even more. What is the point of that?

The key point here is that the change in entropy during a reversible process becomes the reference against which the actual change in entropy in a real irreversible process can be compared. The entropy beyond that for a reversible process is the entropy related to irreversibility that is the foundation of this topic. The best explanation I have read on this topic is in the advanced graduate level textbook by A. Bejan (1997, Advanced Engineering Thermodynamics, 2nd ed., John Wiley and Sons). I attach a photograph of p.110-111 that deals with several key points:



110 3 THE TWO LAWS COMBINED: THE DESTRUCTION OF EXERGY

possible mode is illustrated in Fig. 3.1 because it is essential to the discussion that concludes this section.

With reference to the open system defined in Fig. 3.1, the first law and the second law are

$$\frac{dE}{dt} = \sum_{i=0}^{n} \dot{Q}_i - \dot{W} + \sum_{\text{in}} \dot{m}h^\circ - \sum_{\text{out}} \dot{m}h^\circ \qquad (3.1)$$

$$\dot{S}_{\text{gen}} = \frac{dS}{dt} - \sum_{i=0}^{n} \frac{\dot{Q}_i}{T_i} - \sum_{\text{in}} \dot{m}s + \sum_{\text{out}} \dot{m}s \geq 0 \qquad (3.2)$$

where the methalpy symbol h° is shorthand notation for the generalized enthalpy group $(h + V^2/2 + gz)$, eq. (1.23).

Next, we consider the possibility of changing the design (the internal functioning) of the system for the purpose of maximizing the work transfer rate $\dot{W}$. Of long-term engineering interest is the common characteristic of all the changes that consistently lead to increases in $\dot{W}$. We must recognize, however, that since the First Law of Thermodynamics is an equation, the wish to see changes in $\dot{W}$ means to allow the variation of at least one other term in eq. (3.1). Let us assume that the heat transfer interaction with the atmosphere, $\dot{Q}_0$, varies as $\dot{W}$ is maximized. In other words, let us assume that all the other interactions that are specified around the system (heat transfer rates $\dot{Q}_1, \ldots, \dot{Q}_n$, inflows and outflows of methalpy and entropy) *are fixed by design* and that only $\dot{Q}_0$ "floats" in order to balance the changes in $\dot{W}$, The choice of $\dot{Q}_0$ as the interaction that floats in the wake of design changes is consistent with the role that is assigned traditionally to the rate of heat rejection to the atmosphere in design of power and refrigeration systems. This choice is justified because on earth T_0 is the only temperature reservoir large enough to behave also as a limitless "heat" supply or sink.

If we eliminate $\dot{Q}_0$ between the first law (3.1) and the entropy generation-rate definition (3.2), we find that the work transfer rate $\dot{W}$ depends explicitly on the degree of thermodynamic irreversibly of the system, $\dot{S}_{\text{gen}}$,

$$\dot{W} = -\frac{d}{dt}(E - T_0 S) + \sum_{i=1}^{n} \left(1 - \frac{T_0}{T_i}\right) \dot{Q}_i + \sum_{\text{in}} \dot{m}(h^\circ - T_0 s) - \sum_{\text{out}} \dot{m}(h^\circ - T_0 s) - T_0 \dot{S}_{\text{gen}} \qquad (3.3)$$

Furthermore, since according to the second law (3.2) the entropy generation rate $\dot{S}_{\text{gen}}$ cannot be negative, the first four terms on the right side of eq. (3.3) represent algebraically an upper bound for $\dot{W}$. This upper bound is reached when the system operates reversibly ($\dot{S}_{\text{gen}} = 0$). In this manner, we identify

LOST AVAILABLE WORK 111

the first four terms on the right side of eq. (3.3) as the work transfer rate in the limit or reversible operation:

$$\dot{W}_{\text{rev}} = -\frac{d}{dt}(E - T_0 S) + \sum_{i=1}^{n} \left(1 - \frac{T_0}{T_i}\right) \dot{Q}_i + \sum_{\text{in}} \dot{m}(h^\circ - T_0 s) - \sum_{\text{out}} \dot{m}(h^\circ - T_0 s) \qquad (3.4)$$

Equation (3.3) can be summarized as

$$\dot{W} = \dot{W}_{\text{rev}} - T_0 \dot{S}_{\text{gen}} \qquad (3.5)$$

or, invoking the second law (3.2) one more time, as

$$\dot{W}_{\text{rev}} - \dot{W} = T_0 \dot{S}_{\text{gen}} \geq 0 \qquad (3.6)$$

The conclusion that follows from combining the first law with the second law is this: Whenever a system operates irreversibly, it destroys work at a rate that is proportional to the system's rate of entropy generation. The work destroyed through thermodynamic irreversibility, $W_{\text{rev}} - W$, is appropriately called *lost available work*; this terminology was first used prominently by Kestin [11]. The proportionality between lost available work and entropy generation, or between their respective rates,

$$\dot{W}_{\text{lost}} = T_0 \dot{S}_{\text{gen}} \qquad (3.7)$$

is remembered by many as the *Guoy–Stodola theorem*, in memory of the first two thermodynamicists who were at least partially successful in convincing their contemporaries that eq. (3.7) deserves attention [12, 13]. The engineering history of the concept of lost work is about as old as the history of thermodynamics; remarks on this topic can be found in Refs. 8 and 11. Equation (3.7) was emphasized more recently in the physics literature by Tolman and Fine [14].

A few observations regarding the sign of $\dot{W}$, $\dot{W}_{\text{rev}}$, and $\dot{W}_{\text{lost}}$ are needed, because there is a conceptual difference between $\dot{W}$ (or $\dot{W}_{\text{rev}}$) and $\dot{W}_{\text{lost}}$. The work transfer rate $\dot{W}$ and its limiting value $\dot{W}_{\text{rev}}$ can take both positive and negative values, depending on whether the system is designed to produce or absorb work. For this reason, the second law inequality (3.6), or simply

$$\dot{W}_{\text{rev}} \geq \dot{W} \qquad (3.8)$$

must be recognized as correct in an algebraic sense. In other words, regardless of the sign of the values assigned to $\dot{W}$ and $\dot{W}_{\text{rev}}$, the position of

Figure. Photo of p110-111 in Bejan (1997).

Hopefully the scanned page is clear enough to be read.

Now this formulism did not begin with Bejan – it was established much earlier and can trace its roots back to the first paper by Willard Gibbs (in 1873) that introduced the temperature-entropy diagram (as an alternative to the pressure-volume diagram. (Although the ideas are old, the presentation of those ideas by Bejan is the best description that I have read.)

With a foundation like that it should be possible to start with a set of diagrams (energy, entropy) for closed and open systems, and then go through a few selected applications and identify the relevant terms (and ideally identify those terms which do not matter for the application in question).

In fact the above-noted framework can, and has been used for problems like flow with friction (Bejan, 1997, p. 136-138) or mixing (Bejan 1997, p. 138-141) or heat transfer across a finite gap (Bejan, 1997, p. 134-136) and these are all applications described in the current manuscript.

In summary, my overall suggestion is to change the basic approach of the manuscript. Instead of listing all the previous studies and approaches, etc., just take a few key studies and cast those in a consistent framework that allows people to see the link back to the carnot cycle idea you began with while reformulating the carnot cycle to conventional understanding. With that approach you will address the first aim, i.e., about the inconsistency of thermodynamic concepts and terminology (p.3, lines 7-8).

I had many other comments but given my suggestion of a complete re-write then I did not give all details. The ones below were a few that I note here.

Other Comments

1. p. 2, line 16, "In the final steady state …." Do you mean at rest ? If so then please say it like that.

2. P. 2, line 17. Asset ?? What does this mean? Also, in the same sentence you can remove the full stop, i.e., … conversions, and such limits ……

3. P. 3, line 19. TYPO. several.

4. P. 8, footnote. I did not understand the point of the footnote? Why not incorporate the text of the footnote into the main text.

5. P. 14, line 10. ?TYPO? … physics **that** are required

6. P. 15, line 4. ?distinguished? I think you might mean distinctive here.

7. P. 15, lines 20-22. I did not understand this on several levels. Work transfers are not associated with entropy production. So if you minimise work then why does that also minimise entropy production? (Intuitively, if work is zero then all of the energy change is due to heat flow (in a closed system) and there is a maximum possibility of generating entropy. Further, with that objection aside, how is all this realised by work being performed uniformly along the channel.

8. P. 20, line 30-31. I did not understand why a study based on observations means that it was not possible to assess how far from an optimal state the study was.

9. P. 22, lines 12-15. I did not understand this example. If the system cannot transfer energy then it cannot transfer heat or do work. Why would this system evolve to a state of equal temperature in the absence of heat and work transfers?

10. P. 25, lines 27-30. I did not understand this point. Power is the rate of doing work. Why is any optimisation restricted to those fluxes?