Entropy

Full Paper

A Collision between Dynamics and Thermodynamics

Craig Callender

Department of Philosophy, UCSD, 9500 Gilman Drive, La Jolla, CA, 92093, USA. callender@ucsd.edu

Abstract. Philosophers of science have found the literature surrounding Maxwell's demon deeply problematic. This paper explains why, summarizing various philosophical complaints and adding to them. The first part of the paper critically evaluates attempts to exorcise Maxwell's demon; the second part raises foundational questions about some of the putative demons that are being summoned.

Introduction

In 1866 J.C. Maxwell thought he had discovered a Maxwellian demon—though not under that description, of course [1]. He believed that the temperature of a gas under gravity would vary inversely with the height of the column. From this he saw that it would then be possible to obtain energy for work from a cooling gas, a clear violation of Thompson's statement of the second law of thermodynamics. This upsetting conclusion made him worry that "there remains as far as I can see a collision between Dynamics and thermodynamics." Later, he derived the Maxwell-Boltzmann distribution law that made the temperature the same throughout the column. However, he continued to think about the relationship between dynamics and thermodynamics, and in 1867, he sent Tait a note with a puzzle for him to ponder. The puzzle was his famous "neat-fingered being" who could make a hot system hotter and a cold system colder without any work being done. Thompson in 1874 christened this being a "demon"; Maxwell unsuccessfully tried to rename it "valve." However named, the demon's point was to "show that the second law of thermodynamics has only a statistical validity."

Since that time a large physics literature has arisen that asks a question similar to one asked in theology, namely, does this devil exist?

Beginning with Popper, philosophers examining the literature on Maxwell's demon are typically surprised—even horrified [2,3,4,5,6,7]. As a philosopher writing for a physics audience, I want to explain why this is so. This paper is organized so as to offend everyone, believers and non-believers in demons. Apart from an agnostic middle section, the paper contains a section offending those who believe they have exorcized the demon and a section offending those who summon demons. Throughout the central idea will be to clearly distinguish the various second laws and the various demons. Since to every demon there is a relevant second law, and vice versa, my paper will therefore be a kind of map of the logical geography of the underworld of physics.

I. The First Demon

In the work of Clausius, thermodynamics was considered to be completely universal and applicable without limit. Just as Newton's law of gravitation applied to all matter in motion—atoms, apples, and planets—so too did thermodynamics apply to everything. Uffink [8] describes many difficulties in the arguments of Clausius and others in proving the universal applicability of the second law even before a consideration of mechanical theories, but still the idea that non-equilibrium states go to equilibrium states seems to have been applied to everything.

With the fall of caloric theory and the realization that heat was due to matter in motion, a natural question emerged about the relationship between classical mechanics and thermodynamics. Does thermodynamics apply to everything, even microscopic systems that aren't seen? The German physicists Clausius and (early) Boltzmann took what was perhaps the most natural path: thermodynamics still did apply to everything at every scale. With the theory so successful, induction suggested that there was no reason to doubt that thermodynamics applied at the microscopic level too. The two Germans thus provided arguments for believing that the laws of thermodynamics were corollaries of Hamiltonian mechanics. The H-Theorem was of course the most famous of these. But as people thought about the relationship between the two more, problems arose. (See [9] for an extensive philosophical discussion.)

First, Maxwell's demon challenged the second law. The object of Maxwell's attack is explicitly Clausius and Boltzmann. Thinking the phenomenological second law would fall out as a logical corollary of Hamiltonian mechanics, the Germans were living in "nephelococcygia" [cloud cuckoo land] according to Maxwell [1]. At the micro-level (which Maxwell himself never lived to see), he boldly predicted that we would see "violations of the second law more or less all the time" [1].

Second, in making this point that the phenomenological second law was of limited validity,

Maxwell often used another thought experiment—one also used by Tait, Thompson and later made famous by Loschmidt. Noting that the time reversal invariance (TRI) of classical mechanics allowed one to in principle reverse the motions of all particles, he saw that such a reversal would bring about a decrease in entropy. Maxwell used the imagined reversibility of motions and demon interchangeably as illustrating the point that the second law was not universally valid.

Third, Zermelo noticed that Poincare recurrence or the quasi-periodicy of the solutions of Hamilton's equations also meant that entropy of isolated systems can and will go down, thus threatening the universal character of the second law.

These three theoretical arguments provide overwhelming reason to believe the second law's range is limited. To summarize the argument, assume:

a) Entropy S is a function of the dynamical variables $X(t)$ of an individual system
b) $S(X(t)) = S(X*(t))$, where '*' indicates a temporal reflection
c) The system is closed (its phase space is bounded).

If a, b, and c hold, then the TRI of Hamilton's equations implies S cannot increase monotonically for all initial conditions; and if a and c hold then the quasi-periodicity of the solutions to these equations implies S cannot increase monotonically for all time. If the system is really mechanical, S cannot exhibit monotonic behavior.

If these arguments were not enough, experiments in 1908 vindicated Maxwell's claim that at the microscopic level a strict (non-statistical) second law did not hold. Described by Einstein and Smoluchowski, Brownian motion theory's fluctuations contradicted a strict second law—yet they were experimentally confirmed by Perrin.

All of this is a trivial rehearsal of one of the great episodes of physics. The great puzzle to many philosophers, however, is why so many physicists say and develop theories that seem to deny the lesson of this great episode. Some seek to exorcise the demon, yet *Maxwell's* demon is a friend in this story. Maxwell's thought experiment and subsequent observation of fluctuations destroyed the idea that the second law of thermodynamics is universally valid at all levels. Philosophers also have serious questions about the particular rationale behind some of these exorcisms.

Example: Szilardian Exorcisms

The basic idea of Szilard and followers is that potential *perpetuum mobile* machines of the second kind do manage to do some work, yet there is an accounting error. Szilard referred to the failure to take into account the dissipation involved in measurement; alternatively, Bennett and

Landauer point to the failure to count the entropy cost of memory erasing. In both cases the idea is that an entropy-costing step in the demon's process compensates for any gain against entropy elsewhere. The dissipative step in the cycle guarantees that the net entropy change is either zero or positive.

But what is the argument for the claim that this dissipative step increases entropy? Szilard (quoted from [10], pp. 129, 132) writes:

> If we do not wish to admit that the second law has been violated, we must conclude
> that ...the measurement of x by y, must be accompanied by a production of entropy.

He concludes that he has shown demonic measurement to generate "exactly that quantity of entropy which is required by thermodynamics."

Beginning with Popper, philosophers have criticized Szilard-style exorcisms in two ways. First, they criticize Szilard's followers for apparently denying the basic lesson of the foregoing, that the second law is only statistically valid. Thus Popper writes, "This argument is unsatisfactory because the 'second law' here used as an independent authority is Planck's version—refuted since 1905... The argument is thus much worse than if it were merely circular (which it also is)" [2]. Szilard's followers often seem to assume the strict validity of phenomenological thermodynamics to get their conclusions. Jauch and Baron (reprinted in [10]), for instance, cite the piston in Szilard's cylinder as being inadmissible because it violates Gay-Lussac's law! But this is essentially to deny that fluctuations violate phenomenological thermodynamics. Zurek (reprinted in [10]) seems to admit that Jauch and Baron's point holds classically, but insists that a quantum treatment makes it admissible again. So the piston is viewed as admissible when classical mechanical, inadmissible when it conflicts with a *higher level* theory we know to be approximate, and then admissible again but only due to a *lower level* treatment! The reasoning in the literature often defies logical analysis; one can find many other examples, but as Feyerabend says about a related problem, "this kind of attack is more entertaining than enlightening" [3]. Note, incidentally, that if Szilard and his followers succeed in showing that entropy increase follows always, then they have succeeded too well. Entropy *never* decreasing is not compatible with the laws of statistical mechanics, for fluctuations will make (Boltzmann) entropy decrease sometimes: "a demon that *never* reduces entropy is no less problematic than one that *always* produces work" [6].

Second, and more important, Szilard only "saves" thermodynamics because he has assumed the truth of it in his argument! Even if all this talk of saving the strict second law were just loose talk, still the arguments seem blatantly circular. Popper was perhaps the first philosopher to make this point:

The present situation is logically highly unsatisfactory; for in many arguments within statistical mechanics (and also in information theory), use is made of the fact that certain events are impossible because otherwise the entropy law (in its phenomenological or Planckian form) would be violated. Clearly, all of these arguments are circular." [2]

Perhaps there is a more charitable interpretation of Szilard? Maybe he didn't really assume the truth of the second law but instead used it to help calculate the hidden entropy cost of measurement, *if* dissipation in measurement is what prevents the demon from operating. There are now two things to say. First, whatever Szilard's original intentions, it's clear that followers took him as actually "saving" the phenomenological second law (see Earman and Norton [3], 451ff). Now many people think of the second law as some puffed up transcendental law extending way past its original domain of applicability. Second, if Szilard is not assuming the truth of the second law in exorcizing the demon, then if the argument is to be at all interesting—not merely the calculation of what would have to be the case if the demon is to be stopped—there must be some argument that measurement *must* produce dissipation of some amount. The Szilard school points to principles in the lofty heights of information theory, principles which themselves typically rely on some version of the second law. (For example, Bennett relies on Landauer's thesis regarding the entropy of logically irreversible operations, but Landauer's thesis relies on the second law.)

What I have just described is more or less Earman and Norton's "sound versus profound" dilemma for exorcisms [4,5]: either the second law is assumed, in which case the argument is sound but circular, or the second law is not assumed but the argument relies on some new profound principle. Regarding the "profound" horn of the dilemma I want to express some skepticism about a lofty principle from information theory providing justification for an exorcism. What seems more pertinent would be an argument from the lowly principles of mechanics, classical or quantum—for these are the rules that govern the behavior of matter. Either the exorcism can be translated into and justified from accepted mechanical principles, in which case the detour through information theory may be a useful heuristic but ultimately not necessary, as Denbigh (in [10]) shows with Brillouin's account. Or the information theoretic exorcism cannot be translated into and justified from mechanics, in which case one is essentially proposing a new law of nature. But then it seems ad hoc to posit some otherwise unneeded restriction on the mechanically possible states of the universe.

Let me conclude section I by describing two other areas where physicists have at least sounded like they wanted to deny the statistical character of the second law.

Gibbs v. Boltzmann

Just as Maxwell's demon started as a friend and became a "paradox" that needed to be solved, so too have some viewed Loschmidt's and Zermelo's points about time-reversed trajectories and recurrence, respectively, as "paradoxes" that must be answered. Thus Rothstein [11] and Gross [12], for example, note that the "paradoxes" apply only to individual systems. They argue that since entropy is a function rightly applied to ensembles, the paradoxes do not threaten proofs of entropy increase. In other words, they deny premise a) of our above argument. That is, they are saying essentially that the Gibbs entropy, the entropy of ensembles, is the true entropy, not the Boltzmann entropy, the entropy of individual systems [13]. Recurrence and reversibility can be true of the individuals but not true of the ensemble, and the second law can be saved.

Now, it's one thing to think that the Gibbs entropy is immune to the argument. That may be, depending in part on whether one uses classical or quantum mechanics (in quantum mechanics the whole ensemble still recurs). But it is another thing to think that a) is *false* and as a result there is no tension between microscopic reversibility and macroscopic irreversibility. To use an entropy function, in this context, that is not sensitive to the behavior of individual systems makes this entropy function *irrelevant*. It may be fine for some other use, but for present purposes—reconciling thermodynamics with mechanics—it is of no use since the thermodynamic entropy is applicable to individual systems. My coffee in the thermos has an objective thermodynamic entropy as a property. And it is also a quantum system, which if left to itself, will in time recur to its present macrostate. To "save" the second law by saying that some other function, unrelated to this one, can still evince monotonic behavior is beside the point (see [14,15,16]). Like Maxwell's demon, Loschmidt's and Zermelo's points are not "paradoxes" that need to be resolved; rather they are good true arguments that ought to be accepted—and can be if only we accept the trivially obvious: namely, that the second law is not universally valid.

Subjective v. Objective Probabilities

Here I cannot go deeply into all the subtleties with this topic. But the formal similarity between the Shannon, Gibbs and Boltzmann entropies, plus Szilard-style attempts at exorcising the demon, plus work in computability, plus Gibbs paradox, and other claims have encouraged a subjective reading of the probabilities in statistical mechanics. On this school of thought, the entropy is a function of how much knowledge one has of the microstate of the system. Thinking of the second law this way, given the formal similarity between Gibbs and Shannon entropies, fosters the impression that one can keep the second law universally valid.

Again, it's one thing to say this of the Shannon entropy, which really is a feature of one's knowledge and useful in many fields; but it's another thing to say this of any entropy that is supposed to

be the microphysical counterpart of the thermodynamic entropy. On the subjective interpretation, the entropy of a system goes to zero if one knows the exact microstate of the system. Suppose we place a hot iron bar in contact with a cold iron bar. Suppose also that a different demon, Laplace's, informs you of the exact microstate of this joint system. Does this mean that heat won't flow from the hot bar to the cold bar? No! Trees in the forest tend to equilibrium even when no one is looking. For criticism of the subjective interpretation, see [17].

II. Statistical demons

As Popper (and Smoluchowski) clearly saw [2], one might agree with the foregoing, but still want to save a new version of the second law:

> If we give up Planck's law of the excluded perpetual motion machine of second order, without restating it in some form, then there is nothing left in our physical laws to tell us that we cannot build such a perpetual motion machine.

That is, we can admit that it's in principle physically possible for a Maxwell demon, but be impressed by the fact that we can't use Brownian movement or other fluctuations to reliably exploit these entropy decreases to do useful work. If it's physically possible, why can't we do this? It's still a striking fact about our universe that there aren't any *perpetuum mobiles* of the second kind in it. Is this merely an "accident" of our universe or is there some deeper explanation?

Popper, Smoluchowski, and others have formulated various replacements, and one can imagine better and worse ones. What is clear, however, is that there are no uncontroversial statistical mechanical (as opposed to purely thermodynamical) *perpetuum mobiles*. Entropy may momentarily be decreasing in fluctuation phenomena, but this process doesn't reliably do work. Can a device do better than Brownian motion can do? Can it harness the fluctuations to reliably do work? If statistical demons are possible in principle, then the second law is not even statistically valid with any necessity. Rather the second law just follows from the fact that it happens to be that our universe has few or no demons in it; but we could just as well have lived in a universe with the very same dynamical laws and had plenty of demons violating statistical second laws.

Recent philosophical work on demons [6,7] has, I think, explained in very general terms why ratchet-and-pawl type demons as envisioned by Smoluchowski and Feynman can't work. But some of it [7] has also argued that demons are possible once one relaxes the common assumption that the demon must return to its *exact* starting macrostate for the weaker claim that the demon need only return to *some* macrostate whose entropy is not greater that of its initial macrocondition.

III. Interrogating Devils

Many researchers agree that there could be a Maxwell's demon; in fact, many in this issue cautiously believe they have found one. A quick survey of the literature [18] prompts me to ask the question of whether Maxwell would have been worried by some of these demons. Some of the demons to be discussed seem not to be aimed at either demons of the first or second kind. These demons apparently attack a second law that is neither Maxwell's nor Boltzmann's. Alternatively, maybe they are best not characterized as demons at all. Instead they might be fascinating systems in their own right, without challenging the second law properly conceived. If these systems could lessen mankind's dependence on fossil fuels, they would be no less interesting if it turned out that in some legalistic sense they didn't "count" as second law violators. (I would, however, be sympathetic to the reply that perhaps then we ought to rewrite the second law to prevent it from being so toothless.) I now want to raise a few simple-minded queries about the demons to follow.

Nonequilibrium Steady State Maxwell's demons?

Some of the demons to be discussed are in fact non-equilibrium steady state demons. To a student of classical thermodynamics this is surprising, for one of course expects that as the system relaxes to equilibrium its entropy will rise. Do these non-equilibrium demons decrease more entropy than we expect to increase by the system going to equilibrium? These demons are additionally surprising because most, if not all, classical thermodynamic properties are essentially tied to equilibrium. The entropy of a state A, S(A), for instance, is defined as the integral from B to A over dQ/T through a reversible transformation, where B is some arbitrary fixed state. For S to be a state function, the transformation between B and A must be in principle quasi-static, i.e., a succession of equilibrium states. Continuity considerations then imply that the initial and final states B and A must also be equilibrium states. For non-equilibrium states, therefore, the concepts of entropy, internal energy, etc., simply don't apply. To talk of the entropy of the gas while it passes between equilibrium states in (say) Joule's free expansion experiment is, from the perspective of classical thermodynamics, a misuse of the concepts. Thermodynamics, or thermostatics, has little to say about the system until it settles to a new equilibrium.

But of course, there is no principled reason why science should be impotent in the face of non-equilibrium situations. Non-equilibrium thermodynamics of systems both near and far from equilibrium have been developed. But if we are to have demons challenging non-equilibrium second laws we need non-equilibrium second laws. For this it seems we need a non-equilibrium concept of

entropy. But extending entropy to non-equilibrium is "surprisingly difficult", and it's not even clear it can be done [19].

Presumably the hope is that there will be *local versions* of the second law: entropy flowing out of the boundaries of a small region is less than the entropy generated in that region. And hopefully, the argument for this will not merely appeal to the truth of the equilibrium second law. Indeed, as many know far better than I, the question of Maxwell's demon will be very tricky here; for, since work is being done upon the system by nonconservative forces to get the system into a stationary state, the entropy of the system (as opposed to its surroundings) *decreases*.

Non-Fundamental demons?

Another possibly curious feature of some demons to be discussed in this issue is that it's not clear that they are described with fundamental physics. I don't mean classical as opposed to quantum, for we may suppose that quantum mechanics yields classical mechanics in a certain limit. I mean instead that the proposed demon may not be described with what we take to be fundamental physics as opposed to approximation. That is, if the demon is described—perhaps of practical necessity due to its complexity—as only arising in various limits with various approximations, etc., it may be that the demon is summoned only by the approximation technique.

Approximations may be obvious, as when one uses a standard technique such as the WKB method, or approximations may be more subtle. Suppose I said that classical mechanics is not reversible for it contains friction forces that are irreversible, i.e., non-conservative velocity dependent forces. I might then create a demon using these irreversible forces to create a kind of one-way valve. But of course I would be wrong: the system would be nonconservative, whereas we believe that fundamentally the system, if isolated, should be conservative. The frictional forces are really non-fundamental forces arising as an approximation to the complex interaction of fundamental forces. My demon would not work, but rather would arise from approximation—even if the equations were solved exactly without numerical simulation or other approximation techniques. Just as "interventionists" use phase volume non-conserving approximations (often with stochastic 'kick' terms added to the Hamiltonian) to prove that entropy must *increase* [20], I'm worried that such approximations might also be used unwittingly to prove that entropy *decreases*. Is it clear, for instance, that the trapping probabilities of the Baule-Weinberg-Merrill model in the "gravitator" demon of [21] are consistent with the underlying fundamental dynamics?

Some of the demons proposed in the literature live in open quantum systems. As every student first learns, the second law applies only to isolated systems. Is openness essential to the demon's life? One worry is that openness function like friction did a moment ago. It is commonly said that the

formalism of open quantum systems—systems interacting with their environment—is fundamentally irreversible and dissipative. But if the system according to the fundamental laws of physics is reversible and conservative, do we really have a demon?

This point about openness is perhaps also of relevance to the gravity demons. It is not clear that classical thermodynamics operates very well outside the idealization that there are no significant long-range forces present. Pippard [22], for example, states that the notion of adiabatic isolation is applicable only when gravity is excluded. So if the second law is restricted to adiabatically isolated systems, as Clausius assumed, and if Pippard is right, then it's not clear that a gravity demon meets the strict requirements of a closed isolated system. Further discussion of this and related questions regarding the range of the Maxwell-Boltzmann distribution is needed.

It is also possible that some of the demons proposed are indeed fully compatible with fundamental physics. But then the relevant question, especially for the quantum demons, is whether these are systems that can actually do some work. After all, an electron orbiting the nucleus of a hydrogen atom is in some sense a *perpetuum mobile*, but the trick is to get such a process to do work for you.

The systems considered in this issue are much more complicated than Maxwell's. Hopefully we'll find out whether the devil lives in the detail or merely in the approximation details.

References

1. Garber, E., S. G. Brush and C.W.F. Everitt, *Maxwell on Heat and Statistical Mechanics*, Lehigh University Press, Bethlehem, PA, 1995, 105-120 on gravity; 187-188, 192-193, 205 on demons.
2. Popper, K. *British Journal for the Philosophy of Science* **8**, 151-155 (1957).
3. Feyerabend, P. K., in *Mind, Matter and Method: Essays in Philosophy and Science in Honor of Herbert Feigl*, University of Minnesota Press, Minneapolis, 409-412, 1966.
4. Earman, J. and Norton, J., *Studies in History and Philosophy of Modern Physics* **29**, 435-471 (1998).
5. Earman, J. and Norton, J., *Studies in History and Philosophy of Modern Physics* **30**, 1-40 (1999).
6. Shenker, O., *Studies in History and Philosophy of Modern Physics* **30**, 347-372 (1999).
7. Albert, D., *Time and Chance*, Harvard, Cambridge, 2000.
8. Uffink, J. *Studies in History and Philosophy of Physics* **32**, 305-394 (2001).
9. Sklar, L. *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics*. Cambridge University press, NY, 1993.
10. Leff, H.S., and Rex, A. F., editors, *Maxwell's demon: Entropy, Information, Computing*, Princeton, New Jersey, 1990.
11. Rothstein, J. *Foundations of Physics* **4**, 83-89 (1974).
12. Gross, D.H.E. cond-mat/0105313 (2001).

13. Lebowitz, J.L. *Physica A***194**, 1-27 (1993).
14. Callender, C. *Studies in History and Philosophy of Physics* **32**, 539-553 (2001).
15. Callender, C. *Journal of Philosophy*, XCVI, 348-373 (1999).
16. Shenker, O. *British Journal for Philosophy of Science*, **50**, 33-58 (1999).
17. Loewer, B. *Studies in History and Philosophy of Modern Physics* **32***, 609-620 (2001).
18. Weiss, P. *Science News*, **158**, 234. (2000)
19. Gallavotti, G. *Statistical Mechanics: A Short Treatise*, Springer, Berlin, 1999.
20. Bergmann, P. and Lebowitz, J.L. *Physical Review* **99**, 578-87 (1955).
21. Sheehan, D.P., Glick, J., and Means, J.D. *Foundations of Physics*, **30**, 1227-1256 (2000).
22. Pippard, A. B., 1966, *Elements of Classical thermodynamics for Advanced Students of Physics*, Cambridge University Press, 1964, p. 5.