How is it possible for the second law of thermodynamics to arise from time-symmetric fundamental laws?

There is universal consensus in the literature that we cannot *directly derive* the time-asymmetric second law of thermodynamics from the time-symmetric fundamental laws of physics (Price 1997, 18; Albert 2000, 70; Wallace 2011, 20). Here is why. First, the second law of thermodynamics is a macroscopic law. Second, the fundamental laws of physics are microscopic laws. Third, macroscopic laws must be derived from microscopic laws. Fourth, asymmetries cannot be directly derived from symmetries. Therefore, the second law of thermodynamics cannot be directly derived from the fundamental laws of physics.

Albert (2000, 71-77) thinks we can invert the result by adding two further premises (Wallace 2011, 1). First, the *entropic asymmetry* of the second law can be derived from the geometry of phase space[1] and simple postulates of probability (ie. from the time-symmetric fundamental laws of statistical mechanics (Wallace 2011, 1-22)) — call this the *probability premise*. Second, the *initial point* in phase space of the universe had much lower entropy than our current point — call this the *boundary premise*. Therefore, since the asymmetry in entropy is aligned with the asymmetry of time, the second law is time-asymmetric.

I think there is something fishy about this derivation. First, it lacks a plausible physical interpretation of the probabilities involved. I will adopt concerns from Sklar (1993, Chapter 3) and argue that none is forthcoming. Second, boundary points are not conceptually distinct from full trajectories in phase space. I shall argue that we therefore cannot non-trivially explain the second law in terms of time-symmetric fundamental laws.

My case will rest on the assumption that eternalism gets right the metaphysics of time. It is of course possible that eternalism is false. As such, what follows should be read as an informal reductio: eternalism *and* non-trivial derivations of the second law are incompatible.

---

[1] Phase space is a 6N dimensional space that represents the positions and velocities of N particles as a single point in the space. The trajectory of a point in phase space is the dynamical evolution, given a set of dynamical laws, of the system of particles that it represents. See Albert 2000 Chapter 3 for details.

The second law of thermodynamics purports to explain phenomena such as milk mixing into coffee, but never spontaneously "unmixing" from it. Its simplest formulation states that the entropy of an isolated system never decreases *over time* (Albert 2000, 49-50). Thus the second law has two components that in our world always go together: *entropic asymmetry* and *temporal asymmetry.*

Entropy is the measure of how many microstates in phase space can constitute the same macrostate.[2] The more microstates per macrostate, the higher the entropy. Higher entropy macrostates take up larger volumes in phase space. And if each point in phase space is assigned an *equal probability* (Wallace 2011, 28), then an overwhelming majority of trajectories in phase space will either move from lower to higher entropies, or stay at higher entropies. And so it is overwhelmingly probable that an isolated system that starts out with a low entropy will have a trajectory that increases in entropy; and overwhelmingly improbable that it will have a decreasing trajectory.

And this *seems* very close to the *entropic asymmetry* of the second law. If the universe at some point has a low entropy, then it is overwhelmingly improbable that entropy will ever decrease in any temporal direction. But does the entropic asymmetry really follow?

The literature is riddled with worries like the recurrence and reversibility objections (Wallace 2012, 3-9; Price 1997, 27-34). They share an intuition which will introduce my own complaints. It is this: the fact that it is overwhelmingly improbable that entropy decreases over time is compatible with the existence of possible worlds in which it does. But in our actual world entropy *never* decreases over time. To vindicate the derivation's explanatory power, we need to interpret the probabilities so as to say *why* the *entropic asymmetry* is non-trivially true in our, but not other, worlds.[3]

The standard view is that the probabilities must be *objective* (Wallace 2011, 22).[4] We should, along with Albert (2000, 81), think of probabilities "[...] as *supervening* in one way or another on the non-probabilistic facts of the world" and "[...] as having something or other to do [...] with *actual frequencies*".[5] But the derivation is based on the relative frequency of low-to-high and high-to-low entropy (and other irregular or stagnating) trajectories in phase space. Is this an *actual frequency*?

Our actual world forms one solitary trajectory in phase space and all unrealized trajectories in phase space belong to non-actual but possible worlds. And thus the relative frequency of the various possible trajectories has nothing to do with *actual frequencies* — only one trajectory is actual.[6] And it does not help to suppose that the

---

2 This is a simplification of Albert's presentation (2000, 76-77).
3 This is similar to the "construction constraint" from Wallace (2012, 9).
4 See Albert (2000; 62-70) for reasons to avoid subjective probabilities.
5 His emphasis.
6 Unless we follow Lewis (1986). However, what follows will apply to his view as well.

frequencies are located within our world, as we should then find that each point in phase space has either zero or one probability — at least if we presuppose eternalism, as I shall do.[7] In this context, eternalism is the view that each point in phase space exists equally in time. If we derive it from actual frequencies, entropic asymmetry is either trivial or false in our world.

Thus, suppose the probabilities are based on relative frequencies in *possible* worlds.[8] Call *strange* those possible worlds that share our fundamental laws, but whose trajectories do not always increase in entropy. Entropic asymmetry might seem to hold for segments of these strange trajectories, but not throughout. And while the probability premise shows that non-strange worlds vastly outnumber the strange, it is silent as to why our world is non-strange. But this is precisely what we want to have explained. The explanatory force must thus arise from the boundary premise (Price 1997, 40).

A potential remedy is to relocate the interpretation of objective probability, as in the properties of certain objects (dispositionalism), or some other objective feature of possible worlds. But whatever we choose, there will exist strange possible worlds with the same objective probability-property, but that nevertheless do not always exemplify entropic asymmetry. Because if these worlds were impossible, then the entropic asymmetry would be a necessary and trivial component of worlds with our fundamental laws. And once we have entropic asymmetry, we only need to suppose that it goes together with temporal asymmetry (the boundary premise). And thus the probability premise is redundant and devoid of explanatory force — no matter how we interpret the probabilities.

But of course it was never claimed that the probability premise should do the explanatory work. Price (1997, 40) also argues that it cannot. He accepts (unlike me) that probabilities can explain *entropic asymmetry*, but denies that they can explain *temporal asymmetry*. Probabilities are neutral to directions in time. So the result of the probability premise would be, in contrast to the second law, that entropy never decreases — neither towards the past nor towards the future.

This conclusion follows from Price's demand for an atemporal viewpoint. Since we are trying to explain the asymmetry of time, we cannot presuppose it. We must treat the reasons for asymmetry as invertible in time. Thus, any reason for one end of a trajectory to be the start is also a reason for the other end to be the start. It further follows that, without further justification, the boundary premise cannot plausibly maintain that the initial point of the world's trajectory had low entropy without

---

7 See Saunders (2002) and Balashov & Janssen (2003) for arguments in favor of eternalism, and Price (1997, 12-16) for its relevance to the present argument.

8 The following is inspired by Sklar (1993, Chapter 3).

maintaining the same about the final point (Price 1997, 111-112; Callender 2004, 196). I believe, along with Callender (2004, 209), that this objection is most likely insurmountable. But I also think that *even if it should be answered*, it would not suffice to explain the second law.

My argument begins with a detail of phase space. Given a fixed set of *deterministic* (Albert 2000, 41) fundamental laws, no two trajectories in phase space intersect.[9] Thus, the specification of deterministic fundamental laws and *any* single point in phase space is *sufficient* to specify a full trajectory. And so the conceptual distinction between boundary points and other points is must fall apart: specifying a boundary point is analytically identical to specifying a different point on that same trajectory and to specifying the full trajectory. Moreover, a world (actual or possible) is given a full (physical) description by its trajectory in phase space.[10] Thus, all possible worlds that share the same fundamental laws will differ *only* in what trajectory they take in phase space.

Consider what the boundary premise will say in light of the above. It originally stated that the initial point of our trajectory in phase space had lower entropy than our current point. But once we see that initial points are not distinct from full trajectories, one of three interpretations of the premise follows.

First, we can take it to say that our trajectory has *some* low entropy boundary point. This is most faithful to the original statement, but in runs straight back into the problems of the previous section: it allows the boundary point to belong to a strange trajectory. Thus, it relies upon the probability premise to sieve the strange trajectories from the non-strange — a task to which it cannot rise.

Second, we can take it to say that our trajectory is entropically asymmetric in virtue of having *the right kind of* low entropy boundary point[11] — and then define our direction of time along its direction of increase. But this blatantly begs the question, as it is both necessary and sufficient for the second law that we inhabit such a trajectory.

Third, we can take it to be a rejection of the postulated determinism of the fundamental laws. This would allow for intersecting trajectories, so that the specification of a single point would no longer entail the full trajectory. But this merely moves the problem down the line, If it presupposes the second law in order to justify the rejection of determinism, then it begs the question. Thus, it must find indeterminism plausible for entirely orthogonal reasons. But if we do, it seems that we can postulate a boundary point without begging the question.

---

9 Wallace (2012, 3) argues that this conclusion holds for both classical and quantum cases.

10 Supposing phase space can be generalized to the correct fundamental laws, of course.

11 At least for some significant segment, from what we take to be the Big Bang to the present day and (presumably) beyond.

But suppose we can find such reasons in favor of indeterminism. The indeterministic fundamental laws must either permit or prohibit deviations from the second law given a low entropy boundary point. If they prohibit deviations, then the specification of a boundary point yet again begs the question.

But if they permit deviations, then we will yet again find ourselves incapable of sifting the strange worlds from the non-strange. For the only possible reason that our world is non-strange, on this view, is that it is a result of a series of indeterministic events. So the second law true in our world because, by mere happenstance, the fundamental laws indeterministically led from lower to higher entropy at every juncture. No more can be said about why our world is non-strange, which is exactly what we want to explain.

Regardless of how we interpret the derivation of the second law (from the probability premise and the boundary premise) it will be deficient.[12] It cannot show that the second law arises solely and non-trivially from time-symmetric fundamental laws. It can only show that it arises trivially from facts about our world's trajectory in phase space. But here the explaining is done exclusively by the trajectory, leaving nothing for the derivation to illuminate. It is not necessarily that the second law cannot be explained by anything at all — it is just that it cannot be explained by this form of derivation.

References.

Albert, David. (2000). Time and Chance.

Balashov, Y., & Janssen, M. (2003). Presentism and relativity. The British journal for the philosophy of science, 54(2), 327-346.

Callender, Craig. (2004). Measures, Explanations, and the Past: Should 'Special' Initial Conditions be Explained?. Brit. J. Phil. Sci. 55. 195-217.

Lewis, D. (1986). On the plurality of worlds (Vol. 322). Oxford.

Price, Huw. (1997). Time's Arrow & Archimedes' Point: New Directions for the Physics of Time.

Sklar, L. (1995). Physics and chance: Philosophical issues in the foundations of statistical mechanics. Cambridge University Press.

Saunders, S. (2002). How relativity contradicts presentism. Royal Institute of Philosophy Supplements, 50, 277-292.

Wald, Robert M. (2006). The arrow of time and the initial conditions of the universe. Studies in History and Philosophy of Modern Physics 37 394-398.

---

[12] See Callender (2004, 209-213) for a different argument for a very similar conclusion.

Wallace, David. (2011). The Logic of the Past Hypothesis.
Wallace, David. (2012). The Arrow of Time in Physics.