

Response

Recently I reviewed the book, *The Road to Maxwell's Demon: Conceptual Foundations of Statistical Mechanics*, by Meir Hemmo and Orly Shenker, in this journal (Hemmo and Shenker 2012a; Allori 2013). In that book, Hemmo and Shenker put forward their own account of statistical mechanics. They think that, since Boltzmann's explanation fails, we are in need of such an account. While I was enthusiastic to see a much needed book in the foundations of statistical mechanics, I also expressed some concerns about the main idea developed in it: I pointed out that Hemmo and Shenker's motivation was not very strong, because they seemed to misunderstand Boltzmann's work. I was aware that I might have been mistaken in interpreting their work, but having read their letter, I am still convinced they do not appreciate what Boltzmann did. In passing, let me say that I was disappointed by Hemmo and Shenker's attitude: the authors should have considered the possibility of not having been sufficiently clear in their presentation, and should have taken the opportunity to clarify their point of view in the response, instead of simply placing the blame for any misunderstanding on the reader's lack of sharpness.

Be that as it may, let us start from what Hemmo and Shenker write in their response. They claim that there is a contradiction between what we experience and what Boltzmann's explanation predicts. Their concern can be summarized along these lines:

- (1) Experience tells us that the dynamics will take a system through macrostates of gradually increasing size;
- (2) The probability of a macrostate is proportional to its size: thus the equilibrium macrostate, being the biggest macrostate, has the highest probability;
- (3) If (2), then we should expect the system to evolve directly into equilibrium from any macrostate;
- (4) (1) is in contradiction with (3): according to (1) the system evolves into equilibrium gradually, while according to (3) it directly jumps into it;
- (5) In order to solve the contradiction in (4), Hemmo and Shenker say, we should reject (2) and define the probability of a macrostate as depending on the

dynamics so that it changes with time and is not necessarily proportional to its size.

If we analyse each premise, we will see how, at the very best, they are contentious if not outright false.

Why accept (1)? It is unclear, since (1) is actually false: it is not a fact of experience that the macrostates that the system will cross are *gradually* increasing in size. In fact, when far from equilibrium the macrostates will (or certainly could) be of comparable sizes, so this gradual increase is not necessary at all. Indeed, the dynamics can lead a system into macrostates of smaller sizes; what instead is overwhelmingly likely to happen is that eventually the system will end up in the largest state of all, namely equilibrium state. The point is not the size of each single macrostate, but the ratio between the size of the average macrostate and the equilibrium state:

The most important fact about these macrostates, recognized by Boltzmann, concerns their sizes as measured using the natural volume measure on the phase space, Lebesgue or Liouville measure. It is, in fact, this natural volume measure that provides a sufficiently precise notion of ‘overwhelming majority’ for his typicality claim. ... For a macroscopic system, with particle number $N \sim 10^{20}$ or greater, this means that the overwhelming majority of the points of Γ_{eq} are in Γ_{feq} [Γ_{eq} and Γ_{feq} respectively indicate the energy hypersurface and the equilibrium macrostate—VA], the ratio of the size of a non-equilibrium macrostate to that of the equilibrium macrostate being ridiculously small, of order $10^{-10^{20}}$. (Goldstein 2012, 63–64)

This is the passage from which Hemmo and Shenker quote in their response to my review, but in overlooking the first part of it, they miss the crucial step.

Why accept (2)? Is the probability of a macrostate proportional to its size? Hemmo and Shenker claim that this is part of Boltzmann’s explanation, and they again quote Lebowitz and Goldstein (whom they thus recognize as the main proponents of Boltzmann’s analysis) to show this is the case. Unfortunately, again, their quotations are taken out of context: while it is true that Lebowitz uses the

word ‘probability’ in his passage, just before that he uses the word within quotation marks, presumably to suggest that it should not be taken literally:

To see how the explanation works, consider what happens when a wall dividing the box ... is removed at time t_a . The phase space volume available to the system now becomes fantastically enlarged; for 1 mole of fluid in a 1-liter container the volume ratio of the unconstrained region to the constrained one is of order 2^N or $10^{20^{10}}$. For the system ... this corresponds roughly to the ratio $|\Gamma_{M_d}/\Gamma_{M_a}|$ [Γ_{M_d} corresponds to the macrostate in which the gas is all spread out in the room, Γ_{M_a} to the one in which the gas fills just one half of the contained – VA]. We can then expect that when the constraint is removed the dynamical motion of the phase point X will with very high ‘probability’ move into the newly available regions of phase space, for which $|\Gamma_M|$ is large. This will continue until $X(t)$ reaches $\Gamma_{M_{eq}}$ [the equilibrium macrostate – VA]. (Lebowitz 1993, 35)

Moreover, the passage that immediately follows explains how we should understand this notion as typicality instead:

Boltzmann realized that for a macroscopic system the fraction of microstates for which the evolution leads to macrostates with larger S_B [Boltzmann’s entropy – VA] is so close to 1, in terms of their phase space volume, that such behavior is what should ‘always’ happen. In mathematical language, we say this behavior is ‘typical’. (Lebowitz 1993, 35)

Goldstein, whose article is entirely devoted to understanding the relation of probability to typicality, makes a similar point:

My main concern in this paper, however, is with typicality, a notion that, while extremely important for understanding probability, is *not really a notion of probability at all*. (Goldstein 2012, 59; emphasis added)

Like Lebowitz, Goldstein argues that the vast majority of microstates, in their wandering through phase space, will eventually end up in the equilibrium state. This is the typical behaviour: '[I]n 1877 Boltzmann argued that the evolution of a gas in accord with Boltzmann's equation, while not inevitable, is typical' (Goldstein 2012, 62).

Thus (2) is false. In fact, as already noted, the typicality measure is useful in order to define the size of a macrostate, since what is important is to establish that the equilibrium macrostate is much larger than the others. The point of both Lebowitz and Goldstein, thus, is that the size of a macrostate can help define the notion of typicality, which is not probability but can ground it:

It is via a law of large numbers kind of analysis using P [the measure for typicality – VA] that one can show that it typically happens that $\varrho_{emp} \approx \varrho_{th}$ [ϱ_{emp} and ϱ_{th} denote respectively the empirical and the theoretical distributions – VA] with typicality defined in terms of P . (Goldstein 2012, 66)

What depends on the size of the macrostate, by definition, is Boltzmann's entropy, but this has no direct relation with the notion of probability, as Lebowitz and Goldstein explain.

Finally, why accept (3)? Why should we accept that, at a given time, if there is a macrostate that has high probability, then every single other macrostate will immediately jump into it, regardless of whether it is accessible or not? It seems to me that the only reason for that could be that one *is immediately supposed to do* what is most probable. However, why is that? More to the point, what does it even mean? Clearly, it cannot mean that one *has to do* immediately what it is most probable to do; otherwise, there would be no need of probability. So maybe it could mean 'one most probably does immediately what most probably happens'. In this case, though, it is just an empty statement. Still, the contradiction arises only with the qualification 'immediately': there would not be any problem asserting that *eventually* the system will reach equilibrium; the problem is that, according to Hemmo and Shenker, the system will jump *immediately* into equilibrium, skipping all the gradually increasing macrostates in the middle we are supposed to have experience of.

Nevertheless, there seems to be no reason to believe that this should be the case unless we already had in mind what ‘probability’ meant. That is, we obtain what produces the contradiction in (4) only if we define the most probable state of affairs at time t_1 as the one in which the system at t_0 , immediately preceding t_1 , will go to. The point, though, is that we need an account of probabilities in statistical mechanics (and indeed this is what Boltzmann was trying to provide). Therefore, I think that Hemmo and Shenker are assuming a ‘folk’ notion of probability, according to which the most probable state of affair is most likely to happen *right away*, and they use it to find a problem in Boltzmann’s explanation. However, this approach seems to proceed backwards: Boltzmann’s analysis is supposed to ground the notion of probability (through typicality, in the words of Goldstein and Lebowitz). Thus, we should understand what probabilities in statistical mechanics are in terms of the ingredients provided by Boltzmann. Hemmo and Shenker, instead, do the opposite: they start from their folk notion, and since it does not perfectly match Boltzmann’s proposed explanation, they conclude that the latter is problematic. Naturally, the notion of probability so obtained should not conflict with the one that we use every day, but that does not mean that the latter should take precedence and be used as a premise against Boltzmann’s (or anyone’s) account.

In any case (even granting that the probability of a macrostate is proportional to its size, for the sake of the argument), it does not even follow from Boltzmann’s account that at any time, the system will immediately jump to the most probable state, namely equilibrium. In fact, if a system has, for instance, two accessible states, and one is more likely to happen than the other (namely, one is bigger than the other, in the assumption that I am for the moment conceding), then the system presumably will go into that one. However, if the bigger state is not accessible to the macrostate that the microstate is currently in, there is no reason to think the system will jump into it. In fact, while in a sense it is true that Boltzmann’s approach is not dynamical (that is, the entropy depends not on the dynamics but only on the size of the macrostate), there is another aspect in which it is. In fact, the dynamics will determine where the system will go from one instant to the next. That is, the dynamics will prevent this jump! The accessible macrostates will be the ones to be considered at the time immediately following the present: the dynamics will make the system go into one of these, and the

equilibrium state will be entirely irrelevant, if it is not directly accessible. This is exactly what I originally wrote in my review.

Hemmo and Shenker claim that I miss the point: it is exactly when the equilibrium state is not accessible that the contradiction arises. However, this is not the case: the contradiction in fact would arise only if (a) the probability of a state were proportional to its size, and (b) in Boltzmann's explanation the dynamics were irrelevant. As I pointed out above, though, both (a) and (b) are false: what matters in Boltzmann's explanation is the size of the equilibrium state with respect to all the other states (and not the relative sizes of the various non-equilibrium macrostates); and the dynamics 'guides' the microstates through different macrostates of various sizes (not necessarily of gradually growing sizes) until, eventually, the equilibrium is reached. In this way, it is clear why no jump as Hemmo and Shenker describe would ever happen in Boltzmann's account.

Having clarified all that, one can see that the contradiction in (4) does not arise, and there is no need to redefine the notion of probability of a macrostate as Hemmo and Shenker do. In fact, the main motivation for their account is that Boltzmann's analysis does not account for the dynamics of the system, but that is not the case, as we just saw.

To conclude on an optimistic note, I hope this exchange will be helpful to clarify certain crucial issues in the foundations of statistical mechanics, and, if not, at least to engage the reader in this exciting topic.

VALIA ALLORI
Department of Philosophy
Northern Illinois University