Studies in History and Philosophy of Modern Physics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Studies in History and Philosophy of Modern Physics

journal homepage: www.elsevier.com/locate/shpsb

Histories, dynamical laws, and initial conditions –Invariance under time-reversibility and its failure in Markov processes, with application to the second law of thermodynamics and the past hypothesis

Elliott Sober

5185 Helen C. White Hall, Philosophy Department, University of Wisconsin, Madison, Madison, Wisconsin 53706, USA

ARTICLE INFO

Article history:
Received 5 March 2019
Received in revised form
19 August 2019
Accepted 22 August 2019
Available online xxx

Keywords:
Entropy
Markov process
Past hypothesis
Second law of thermodynamics
Time-reversal invariance
Time-translational invariance

ABSTRACT

A Markov process can be invariant under time reversal and it also can exhibit a failure of invariance that is "uniformly positive." I show how each of these possibilities contributes to the project of deciding when a temporal sequence of states has a higher probability than its mirror image. Neither suffices, but a distinct property of the Markov process completes the project, namely the unconditional probabilities of two possible states of the system at the start of the process. The concept of forward time-translational invariance plays a role in the analysis, but I discuss backward time-translational invariance as well. I argue that the Markov framework helps clarify how the Past Hypothesis (the hypothesis that the universe began in a very low entropy state) is related to the Second Law of Thermodynamics, and how each is relevant to explaining why histories that exhibit entropy increase have higher probabilities than histories that exhibit entropy decline. I argue that the Past Hypothesis, if true, helps explain this fact about histories, but a far weaker hypothesis about the universe's initial state suffices to do so.

1. Introduction

The idea of using Markov processes to clarify how the time-asymmetric second law of thermodynamics is compatible with time-symmetric laws of particle motion dates back at least to the famous dog/flea model of Ehrenfest and Ehrenfest-Afanassjewa (1907).[1] Abstracting away from the physical details of thermal processes, they consider n fleas (numbered 1, 2, …, n) and two dogs on which the fleas live. At each time step, a number between 1 and n is chosen at random, with the result that the flea with that number moves from the dog it is on to the other dog. This means that a flea's probability of changing dogs is 1/n regardless of whether it is on dog 1 or dog 2. However, if one dog has more fleas than the other at a given time, the dog with more fleas has a higher probability of losing a flea than the dog with fewer. As a result, the absolute value of the difference between the number of fleas on dog 1 and the number on dog 2 will, in expectation, decline with time, unless the two dogs house an equal number of fleas. This expected decline is due to there being a probabilistic asymmetry; the probability of going from a difference of d+1 to a difference of d (where $1 \leq d < n$) exceeds the probability of going from a difference of d to a difference of d+1. The dog/flea model provides a simple illustration of how a time-symmetric micro process (one governing what happens to individual fleas) can result in a time-asymmetric macro process (one concerning what happens to the difference in flea numbers between the two dogs).

Whereas the Ehrenfests (following Boltzmann) were interested in the relationship of micro to macro, my topic in this paper is resolutely macro. I will use Markov processes to analyze relationships that connect (i) the Second Law of Thermodynamics (**SLT**), (ii) the past hypothesis (which says that the universe began in a low-entropy state), and (iii) histories of entropy increase and their mirror images. Does the SLT suffice to explain why histories of entropy increase are more frequent than mirror-image histories of entropy decline? Does coupling the SLT with the past hypothesis explain this? And is the past hypothesis needed here, or can far weaker assumptions about initial conditions do the job? As with the Ehrenfests' discussion of dogs and fleas, I abstract away from the physics of heat and the property of entropy.

E-mail address: ersober@wisc.edu.

[1] See Kelly (1979, pp. 17–20) and Uffink (2007, Section 7.2) for discussion.

https://doi.org/10.1016/j.shpsb.2019.08.004
1355-2198/

2. The Markov property and (forward) time-translational invariance

Consider a probabilistic process in which a system can change from one of several mutually exclusive and collectively exhaustive states to another, repeatedly. The process can be described by conditional probabilities that have the form $\Pr(X(t_2) = b \mid X(t_1) = a)$; here $t_1 < t_2$, and $X(-)$ is a variable that maps times onto states. For convenience, I'll consider a sequence of discrete times. More substantively, I'll assume that the processes under discussion have the Markov property and are (forward) time-translationally invariant, meaning:

> A process has the **Markov property** for a partition[2] of states that the variable X might occupy if and only if, for any two states a and b in S, for any two times $t_1 < t_2$, and for any history $H(t_1)$ of the system before t_1 that is characterized by its sequence of X states during that temporal period, $\Pr(X(t_2) = b \mid X(t_1) = a) = \Pr(X(t_2) = b \mid X(t_1) = a \ \& \ H(t_1))$.
>
> A process is **(forward) time-translationally invariant** for a partition of states S if and only if, for any two states a and b in S, any two times t_1 and t_2, and any positive integer d, $\Pr(X(t_{1+d}) = b \mid X(t_1) = a) = \Pr(X(t_{2+d}) = b \mid X(t_2) = a)$.

By "forward," I mean that the conditional probabilities in this second definition have the form Pr(later | earlier). I discuss backward time-translational invariance in the Appendix.

Many forward probabilities fail to be time-translationally invariant. For example, the probability that S will have an annual salary over $40,000 at age 32, given that S graduated from university in the United States at age 22, may have one value in this decade but a different one in the decade before. However, it is notable that familiar dynamical laws of nature are (forward) time-translationally invariant. Examples include the laws of quantum mechanics and those of population genetics; they hold true regardless of the time to which they are applied. I'll assume that the same is true of the second law of thermodynamics.

3. Invariance under time-reversal and a type of invariance failure

To begin, let's consider processes that are "invariant under time reversal." This means:

> (I) A process is **invariant under time reversal** for a partition of states S if and only if, for any two states a and b in S, and for any two times $t_1 < t_2$, $\Pr(X(t_2) = b \mid X(t_1) = a) = \Pr(X(t_2) = a \mid X(t_1) = b)$.

Notice that the two probabilities in **I** are forward-directed.

The definition just given is a bit nonstandard, at least for the physics literature. There, the more usual formulation is for states a and b to occur in the first probability in **I**, but a different, though related, pair of states, a* and b*, to occur in the second (Callender, 1995, p. 332; Earman, 2002, p. 247; Earman, 2006, p. 408). For example, consider a model in particle physics in which $\langle X(t), V(t)\rangle$ represents the position and instantaneous velocity of a particle at time t. If a particle moves from state $\langle X(t_1) = p, V(t_1) = v\rangle$ to $\langle X(t_2) = q, V(t_2) = v\rangle$, the time-reverse of this transition, in the sense of **I** (as liberalized in the physics literature), is moving from $\langle X(t_1) = q, V(t_1) = -v\rangle$ to $\langle X(t_2) = p, V(t_2) = -v\rangle$. This wrinkle won't matter to the ideas I'll present in what follows.[3]

A special case of **I** arises when there is a sequence of ordered states, and the system can evolve directly (i.e., in one time step) from one state to another only if the two states are adjacent. Each state is reachable from every other, but the transition to a nonadjacent state requires a series of changes. Two concepts of adjacency are needed here: states are adjacent according to an ordering of states, whereas times are adjacent according to an ordering of times. For example, Tuesday is adjacent to Wednesday but not to Thursday, and a temperature between 40 and 50° is adjacent to a temperature between 50 and 60, but not to a temperature between 60 and 70. Here's the definition, using "t_k" and "t_{k+1}" to denote adjacent times:

> **(I_{ord})**: A process is **invariant under time reversal for a partition S of ordered states**
> if and only if, for any two states a and b in S, and for any time t_k,
> $\Pr(X(t_{k+1}) = b \mid X(t_k) = a) = \Pr(X(t_{k+1}) = a \mid X(t_k) = b) > 0$ if a = b or a and b are adjacent, and
> $\Pr(X(t_{k+1}) = b \mid X(t_k) = a) = \Pr(X(t_{k+1}) = a \mid X(t_k) = b) = 0$ otherwise.

I want to contrast $\mathbf{I_{ord}}$, not with its negation, but with a special case of that negation, namely:

> **(IF_{upb})**: A process exhibits **invariance failure (with a uniform positive bias) for a partition S of n ordered states** if and only if, for any two states a and b, and for any time t_k, $\Pr(X(t_{k+1}) = b \mid X(t_k) = a) > \Pr(X(t_{k+1}) = a \mid X(t_k) = b)$, if a and b are adjacent and b>a, and
> $\Pr(X(t_{k+1}) = b \mid X(t_k) = a) = \Pr(X(t_{k+1}) = a \mid X(t_k) = b) = 0$ if a and b are not adjacent and $a \neq b$.

Where $a < b < c$ and the three states are adjacent, $\mathbf{IF_{upb}}$ says that the probability of going from a to b exceeds the probability of going from b to a, and the probability of going from b to c exceeds the probability of going from c to b, but the values of the two differences need not be the same. The process has a *uniform* positive bias, but the bias need not be *constant*.[4]

4. What do invariance and invariance failure entail about the probabilities of different histories?

The probabilities in $\mathbf{I_{ord}}$ and in $\mathbf{IF_{upb}}$ are all conditional. What do those dynamical laws say about the unconditional probabilities of the different histories the system might have? For example, if times $t_1 < t_2$ are adjacent and states $a < b$ are adjacent, does $\mathbf{I_{ord}}$ entail that

$$\Pr(X(t_1) = a \ \& \ X(t_2) = b) = \Pr(X(t_1) = b \ \& \ X(t_2) = a), \tag{1}$$

and does $\mathbf{IF_{upb}}$ entail that

[2] By a partition, I mean that the states are exclusive and exhaustive.

[3] Invariance under time-reversal is a different concept from the concept of detailed balance, which says that $\Pr(X(t_2) = b \mid X(t_1) = a)\Pr(X(t_1) = a) = \Pr(X(t_2) = a \mid X(t_1) = b)\Pr(X(t_1) = b)$, where the unconditional probabilities are equilibrium probabilities; see Kelly (1979, p. 5). Unfortunately, detailed balance is often called "reversibility" in the literature on Markov chains.

[4] $\mathbf{IF_{upb}}$ comes close to capturing how CP-symmetry (charge conjugation parity symmetry) is violated in particle physics (Christenson, Cronin, Fitch, & Turlay, 1964). Transitions of neutral kaon states violate time reversal, but they do so with different "strengths," including zero. Instead of the "uniformly positive" that I have in my definition of $\mathbf{IF_{upb}}$, what is needed for this physics example is "some positive and the rest zero." That modest change would not affect the claims I make in what follows. My thanks to Bryan Roberts for this point.

Please cite this article as: Sober, E., Histories, dynamical laws, and initial conditions –Invariance under time-reversibility and its failure in Markov processes, with application to the second law of thermodynamics and the past hypothesis, Studies in History and Philosophy of Modern Physics, https://doi.org/10.1016/j.shpsb.2019.08.004

$$\Pr(X(t_1)=a \,\&\, X(t_2)=b) > \Pr(X(t_1)=b \,\&\, X(t_2)=a)? \tag{2}$$

The answer to both questions is *no* (Ehrenfest-Afanassjewa, 1925). Assuming that $\Pr(X(t_1)=a)>0$ and $\Pr(X(t_1)=b)>0$, the definition of conditional probability entails that

$$\Pr(X(t_1)=a \,\&\, X(t_2)=b) = \Pr(X(t_2)=b \mid X(t_1)=a)\cdot\Pr(X(t_1)=a) \tag{3}$$

$$\Pr(X(t_1)=b \,\&\, X(t_2)=a) = \Pr(X(t_2)=a \mid X(t_1)=b)\cdot\Pr(X(t_1)=b). \tag{4}$$

It is obvious that the left-hand sides of (3) and (4) can be unequal, even if the first term on the right-hand side of (3) equals the first term on the right-hand side of (4). So $\mathbf{I_{ord}}$ does not entail that histories in which X increases must have the same probability as mirror-image histories in which X declines. It also is obvious that the left-hand sides of (3) and (4) can be equal, even if the first term on the right-hand side of (3) is greater than the first term on the right-hand side of (4). So $\mathbf{IF_{upb}}$ does not entail that histories in which X increases must be more probable than histories in which X declines.

I now want to consider the further question of how dynamical laws ($\mathbf{I_{ord}}$ and $\mathbf{IF_{upb}}$), coupled with equalities or inequalities between histories of increase and their mirror images constrain a third item – the unconditional probabilities that characterize initial conditions. To begin, let $t_1<t_2<t_3$ be three successive times and $a<b<c$ be three adjacent states. As just noted, $\mathbf{I_{ord}}$ does not entail the following:

$$\Pr(X(t_1)=a \,\&\, X(t_2)=b \,\&\, X(t_3)=c) = \Pr(X(t_1)=c \,\&\, X(t_2)=b \,\&\, X(t_3)=a), \tag{5}$$

but given the Markov property, (5) can be reformulate as:

$$\frac{\Pr(X(t_1)=a)}{\Pr(X(t_1)=c)} = \frac{\Pr(X(t_3)=a \mid X(t_2)=b)}{\Pr(X(t_2)=b \mid X(t_1)=a)}\cdot\frac{\Pr(X(t_2)=b \mid X(t_1)=c)}{\Pr(X(t_3)=c \mid X(t_2)=b)} \tag{5*}$$

It follows from $\mathbf{I_{ord}}$ and the assumption of (forward) time-translational invariance that the two ratios on the right-hand side of (5*) are both equal to unity, but that is not enough for (5*) to be true. One also needs the assumption that

$$\Pr(X(t_1)=a) = \Pr(X(t_1)=c). \tag{6}$$

More generally, if the process covers n consecutive times, with a steady monotonic increase from the start time in state x_1 to the end time in state x_n, $\mathbf{I_{ord}}$ entails that the probability of this history is the same as the probability of its mirror image precisely when the probability of the start time's being in the lowest state x_1 is the same as the probability of the start time's being in the highest state x_n.

Similarly, if $a<b<c$ are three adjacent states, $\mathbf{IF_{upb}}$ does not entail inequality (7).

$$\Pr(X(t_1)=a \,\&\, X(t_2)=b \,\&\, X(t_3)=c) > \Pr(X(t_1)=c \,\&\, X(t_2)=b \,\&\, X(t_3)=a), \tag{7}$$

but again the Markov property allows (7) to be rewritten as

$$\frac{\Pr(X(t_1)=a)}{\Pr(X(t_1)=c)} > \frac{\Pr(X(t_3)=a \mid X(t_2)=b)}{\Pr(X(t_2)=b \mid X(t_1)=a)}\cdot\frac{\Pr(X(t_2)=b \mid X(t_1)=c)}{\Pr(X(t_3)=c \mid X(t_2)=b)} \tag{7*}$$

Given time-translational invariance, $\mathbf{IF_{upb}}$ entails that each of the two product terms on the right-hand side of (7*) is less than 1, but that isn't enough to make (7*) true. An additional assumption about the unconditional probabilities on the left-hand side is needed. A simple sufficient condition for (7*) is:

$$\Pr(X(t_1)=a) > \Pr(X(t_1)=c), \tag{8}$$

which can be true even if $\Pr(X(t_1)=a)$ is very small. Notice that the left-hand side of (7*) is a ratio between the smallest and the largest states mentioned.

There are two terms on the right-hand side of (7*), owing to the fact that the example at hand involves three times. This opens the door to a more general statement: if $\mathbf{IF_{upb}}$ is true and there are n times, the probability of a history of monotonic increase (from x_1 to x_n) is greater than the probability of its mirror image precisely when

$$\text{(R1)}\ \frac{\Pr(X(t_1)=x_1)}{\Pr(X(t_1)=x_n)}$$

is not too small. If the **R1** ratio is greater than or equal to 1, that suffices (but isn't necessary) for the generalization of (7*) to be true. As n increases, the right-hand side of the generalization of (7*) approaches zero. Thus, it gets easier and easier for the generalization of (7*) to be true as the number of states gets larger and larger.[5] As n increases, the constraint on the value of the **R1** ratio becomes more and more modest.

I so far have considered a history of monotonic increase and its mirror image, but there are other histories that have positive probability according to $\mathbf{IF_{upb}}$. For example, what does $\mathbf{IF_{upb}}$ say about the following inequality (where $a<b<c$ and they are adjacent)?

$$\begin{array}{l}\Pr(X(t_1)=a \,\&\, X(t_2)=b \,\&\, X(t_3)=c \,\&\, X(t_4)=b) > \\ \Pr(X(t_1)=b \,\&\, X(t_2)=c \,\&\, X(t_3)=b \,\&\, X(t_4)=a)\end{array}. \tag{9}$$

Is a history of two-ups-and-one-down more probable than a history of one-up-and-two-downs? Inequality (9) is true precisely when (9*) is true

$$\frac{\Pr(X(t_1)=a)}{\Pr(X(t_1)=b)} > \frac{\Pr(X(t_3)=b \mid X(t_2)=c)}{\Pr(X(t_3)=c \mid X(t_2)=b)}\cdot\frac{\Pr(X(t_2)=c \mid X(t_1)=b)}{\Pr(X(t_4)=b \mid X(t_3)=c)}\cdot\frac{\Pr(X(t_4)=a \mid X(t_3)=b)}{\Pr(X(t_2)=b \mid X(t_1)=a)}. \tag{9*}$$

Assuming (forward) time-translational invariance, $\mathbf{IF_{upb}}$ entails that the first and third product terms on the right side of (9*) are

[5] This result is compatible with Zermelo's use of the Poincare' recurrence theorem in criticism of Boltzmann's H theorem; see Uffink (2007) for discussion.

less than 1, while the second term is greater than 1. **IF$_{upb}$** says nothing further about the magnitudes of these three ratios, so it fails to provide the kind of simple criterion for this problem that it supplied for the case of monotonic increase versus decline. However, a logically stronger conception of invariance failure delivers the goods. Let's replace *uniform* positive bias with *constant* positive bias (**IF$_{cpb}$**). If i is the probability of increase in a unit time interval and d is the probability of decline, the result that

$$\frac{\Pr(X(t_1)=a)}{\Pr(X(t_1)=c)} > \frac{d^2 i}{i^2 d} = \frac{d}{i}. \tag{9**}$$

IF$_{cpb}$ says that $\frac{d}{i} < 1$; the ratio of unconditional probabilities on the left-hand side of (9**) must be bigger than that for (9**) to be true. Notice that the left-hand ratio in (9**) concerns the first and last states that the first history displays, not the highest state that history achieves, which in this example occurs in the middle of the history.

This distinction between the *highest* state a history occupies and the *last* state the history occupies was not present in the example of monotonic increase, but it does bear on how one should describe what all these examples have in common:

> Suppose s and e are any two states (adjacent or not) and history H runs from $X(t_1)=s$ to $X(t_n)=e$, so the mirror image history M runs from $X(t_1)=e$ to $X(t_n)=s$. Then whether H has a higher unconditional probability than M is settled by (i) the process model (for example, **I**, **IF$_{upb}$**, **IF$_{cpb}$**), (ii) the assumption of (forward) time-translational invariance, and (iii) the value of the ratio $\frac{\Pr(X(t_1)=s)}{\Pr(X(t_1)=e)}$.

Beyond that, it doesn't matter what the unconditional probabilities are of the other states that the start time t_1 might occupy, nor does it matter what the absolute values are of $\Pr(X(t_1)=s)$ and $\Pr(X(t_1)=e)$, and it also doesn't matter what the unconditional probabilities are for any of the states that the other times might occupy.

5. Application to thermodynamics

The second law of thermodynamics says that entropy has a high probability of increasing in closed systems, but it is often thought to be relevant to systems that aren't closed (like a cup of coffee that receives a drop of milk). The past hypothesis says that the whole universe had very low entropy at its birth.[6] This hypothesis is relevant to what happens subsequently in the universe writ large, but it isn't so clear how it pertains to the cup of coffee you now hold. The third item that needs to be considered is the fact that different histories in systems of a given type have different probabilities; for example, adding milk to coffee always results in a homogenous liquid, but milky coffee never separates into an island of milk in a sea of coffee. To get these three ducks in a row, I'll talk about relationships among three propositions that each are about systems "of type T":

> (**SLT**) **the second law of thermodynamics**: Entropy probably increases in systems of type T.
> **(P) the past hypothesis**: Systems of type T begin in a very low entropy state.
> **(THI) two-history inequality**: In systems of type T, any possible[7] history of monotonic increase in entropy has a higher probability than its mirror-image history of monotonic decline.

P is a nonprobabilistic statement and **THI** is a statement about unconditional probabilities, but what is the logical form of **SLT**? I take it to be a statement about a forward-directed conditional probability. When $t_1 < t_2$, **SLT** says (of systems of type T) that[8]

$$p(\text{Entropy}(t_2) > x \mid \text{Entropy}(t_1) = x) \text{ is high.} \tag{10}$$

This does not entail that

$$p(\text{Entropy}(t_1) < x \mid \text{Entropy}(t_2) = x) \text{ is high.} \tag{11}$$

Applied to the coffee example, **SLT** says that pouring milk into coffee has a high probability of yielding a homogeneous milky coffee, but the law does not say that homogeneous milky coffee has a high probability of tracing back to separate coffee and milk.

I use the phrase "systems of type T″ to bring **SLT**, **P**, and **THI** in contact with each other, but I won't try to provide a more informative description of what T should be. When the three propositions are formulated so that they describe different types of system, it is left open whether they have any bearing on each other. For example, if the past hypothesis is a claim about the early universe, is it relevant to explaining what happens when you pour milk into your coffee? Winsberg (2004) and Earman (2006) answer this question in the negative.

How do **SLT**, **P**, and **THI** relate to the preceding discussion of Markov processes? I stipulated at the start of the paper that I was considering Markov processes that involve a partition of discrete states, but proposition (10) makes it plain that **SLT** is about a continuous variable (entropy). In addition, my discussion of **I$_{ord}$** and **IF$_{upb}$** presupposes that there are adjacent states, but this isn't true for a continuous variable. These problems can be addressed by thinking about a discrete version of the **SLT** in which the continuous variable is chunked into a large number of small intervals. Another issue concerns whether the increase in entropy described by the **SLT** has the Markov property. Does present entropy screen-off past entropy from future entropy?[9] I'll assume here that it does, though I recognize that this assumption requires further investigation. And what does entropy mean in **SLT**? The entropy I'm thinking of is Boltzmann's.

SLT instantiates **IF$_{upb}$** since it entails (when $e > 0$, e is small, and $t_1 < t_2$) that

$$\begin{aligned} &p(\text{Entropy}(t_2) = x + e \mid \text{Entropy}(t_1) = x) > \\ &p(\text{Entropy}(t_2) = x \mid \text{Entropy}(t_1) = x + e). \end{aligned} \tag{12}$$

The upshot is that the previous discussion of Markov processes can be used to clarify how **SLT**, **P**, and **THI** are related to each other.

First, **P** and **SLT** are logically independent. Second, **SLT** does not entail **THI** and neither does **P**, but the conjunction **SLT&P** does entail **THI** if (forward) time-translational invariance is true. This second point leads to the third: **P** is much stronger than it needs to be. True, for **SLT** to entail **THI**, *some* additional assumption is needed, but it needn't be the assumption that $\Pr(X(t_{start}) = x_{lowest}) = 1$. Here "$x_{lowest}$" denotes a disjunction of possible entropy point values that comprise, say, the lowest 1%. Nor is it required that $\Pr(X(t_{start}) = x_{lowest}) > \Pr(X(t_{start}) = x_{highest})$. The point about

[6] Albert (2000) and Loewer (2007) argue that the past hypothesis does important work in explaining various time-asymmetries, including thermodynamic time-asymmetries.

[7] I add "possible," since some conceivable histories of monotonic increase are in fact impossible – for example, because systems in a given state can evolve directly only to states that are adjacent.

[8] The lower-case "p(–)" represents a probability density. Also, a qualification is needed here: x isn't the maximum possible value.

[9] This question is distinct from the following: "Does present entropy screen-off the past history of the system's micro states from its future entropy?" I take it that the answer to the latter question is *no*.

R1 in the previous section shows that what is required is just that the ratio of $\Pr(X(t_{start}) = x_{lowest})$ to $\Pr(X(t_{start}) = x_{highest})$ not be too small.

These facts about entailment suggest some claims about explanation and evidence. **P** doesn't explain **SLT**, nor is the converse true. **SLT&P**, if true, would explain **THI** (assuming (forward) time-translational invariance), but that isn't much of an argument for **P**, since a considerably weaker assumption can be conjoined with **SLT** to entail (and thus explain) **THI**.

This skeptical comment about **P** (now understood as a claim about the early universe) leaves it open that there are observations, distinct from **THI**, that provide evidence for **P**. Indeed, an observation of the present entropy of the universe may do the trick.[10] As already noted, if y is the present entropy of the universe, then **SLT** does not entail that an earlier state *probably* had an entropy lower than y. However, **SLT** does entail that the *maximum likelihood estimate* of the entropy of an earlier state is lower than y, given that y is the entropy of the present state.[11] That is, the point value of x that maximizes the conditional probability density

$$p(\text{Entropy}(\text{now}) = y \mid \text{Entropy}(\text{birth}) = x) \tag{13}$$

is such that $x < y$. A step further can be taken. If you know the age of the universe and the expected rate of entropy increase in the universe, you can determine which point value of x maximizes the value of conditional density (13). You then can say that this point value of the universe's entropy at its birth, coupled with the expected rate of entropy increase, explains the point value of the universe's present entropy. Notice, however, that the proposition thereby explained is not a proposition about time asymmetry. Notice also that I am treating proposition **P** as posing an *estimation* problem; I am suspicious of introducing it as a *postulate* that is said to be justified on the grounds that it would explain something if it were true. The latter approach is suspect because it fails to consider whether alternative postulates might also explain the thing in question.

The strategy pursued in this paper contrasts sharply with what might loosely be termed "the Boltzmann program," one part of which is to explain time asymmetries in entropy histories by postulating that higher entropy macro-states have higher unconditional probabilities than lower entropy macro-states possess. The Boltzmann program focuses on unconditional probabilities; I've mentioned these too, but they play second fiddle to the conditional probabilities that I have foregrounded. As mentioned earlier, the Boltzmann program is to develop a two-level picture, wherein the micro-dynamics obey **I** (invariance under time-reversal) while the macro-dynamics, since they obey **SLT**, must obey $\mathbf{IF_{upb}}$. This contrasts with the one-level story I have told here about the relationship of **SLT**, $\mathbf{IF_{upb}}$, and **P**.

There is no conflict between these approaches if they aim at solving different problems. I've focused on explaining asymmetries in the probabilities of *histories* and their mirror images, where **SLT** provides the explanation once it is supplemented with modest assumptions about the probabilities attaching to two possible states that the start time might occupy. Boltzmannians want to explain why the **SLT** is true at the macro-level by giving a plausible model of the probabilities that apply at the micro-level. I have offered no objection to this project. I note, however, that explaining asymmetries in histories and explaining the **SLT** are distinct undertakings. My only beef with the Boltzmann program is the suggestion that **P** (the Past Hypothesis) is needed to explain time asymmetries in entropic histories. It isn't.

6. Concluding comment

The application just described of the Markov process idea of $\mathbf{IF_{upb}}$ to **SLT**, **P**, and **THI** has nothing special to do with the fact that those three propositions are about the thermodynamic concept of entropy, but I nonetheless hope that abstracting away from that physical detail throws light on how the three propositions are related. If the problem of "explaining thermodynamic time asymmetries" is the problem of explaining why histories of entropy increase are more common than histories of entropy decline, then **SLT** is plainly relevant, but the relevance of **P** is less clear. Given **SLT**, propositions much weaker than **P** about the system's initial state suffice to explain those asymmetries. And if the problem is to explain **SLT**, **P** does not do that, either.

The strategy of abstracting away from physical details may also be useful in connection with the more general problem of time asymmetry. Time asymmetries often involve properties that are not discussed in the laws of physics (Barrett & Sober, 1994). For example, natural selection sometimes instantiates the $\mathbf{IF_{upb}}$ model, but fitness and heritability are not discussed in physical theories. And within physics itself, there are time asymmetries that aren't thermodynamic (Earman, 2006). The claim that statistical mechanics solves "the" problem of time asymmetry embodies an untenable reductionism.

Acknowledgements

I am grateful to Martin Barrett, Jeremy Butterfield, Craig Callender, Eddy Chen, Daniel Hausman, Jenann Ismael, Bryan Roberts, William Roche, Mike Steel, Peter Vranas, Wayne Myrvold, and the two anonymous referees of this journal for helpful comments, and to Julie Larson-Guenette for translating Ehrenfest-Afanassjewa (1925) from the Dutch.

Appendix. (Backwards) time-translational invariance and its implications

I earlier defined the concept of forward time-translational invariance. Here's a definition for backward:

> A process is **(backward) time-translationally invariant** for a partition S of states if and only if, for any two states a and b in S, for any two times t_1 and t_2, and for any positive integer d, $\Pr(X(t_1) = b \mid X(t_{1+d}) = a) = \Pr(X(t_2) = b \mid X(t_{2+d}) = a)$.

Notice that the probabilities in this definition are of the form Pr(earlier | later).

One can't blithely assume both sorts of time-translational invariance, since that entails that the probability distribution of states never changes (Sober, 1993). This can be seen by applying the odds formulation of Bayes's theorem to three adjacent times $t_1 < t_2 < t_3$ and two adjacent states (a and b):

$$\frac{\Pr(X(t_2) = b \mid X(t_1) = a)}{\Pr(X(t_3) = b \mid X(t_2) = a)} = \frac{\Pr(X(t_1) = a \mid X(t_2) = b)}{\Pr(X(t_2) = a \mid X(t_3) = b)} \cdot \frac{\Pr(X(t_2) = b)}{\Pr(X(t_3) = b)}. \tag{14}$$

If (forward) time-translational invariance holds, the ratio on the left-hand side equals 1. If (backward) time-translational invariance

[10] Here I'm assuming for the sake of argument that it makes sense to attribute an entropy to the whole universe, both in its present state and at its birth. Earman (2006) doubts that it is physically meaningful to attribute a low Boltzmann entropy to the early universe, but see Wallace (2010).

[11] For an analog of this inference problem in evolutionary biology, see Sober (2015, p. 183).

holds, the first ratio on the right-hand size equals 1. So, if both invariances hold, the second ratio on the right-hand side equals 1. This applies to all the possible states the system might occupy, and to all subsequent times. So for each state, the unconditional probability of the system's occupying that state doesn't change with time after t_2. This means that forward and backward time-translational invariance are incompatible for a system that has its states change probability with time.

I already described what $\mathbf{IF_{upb}}$ entails about the probabilities of monotonic increase and monotonic decline when the forward version of time-translational invariance is assumed. If you shift to the backward version, there are parallel consequences for $\mathbf{IF_{upb}}$. The (7*) inequality, which uses forward-directed probabilities, can be rewritten in a logically equivalent form that uses backward-directed probabilities, as follows (where $a<b<c$ and the three states are adjacent):

$$\Pr(X(t_3)=c)\bullet\Pr(X(t_2)=b\,|\,X(t_3)=c)\bullet\Pr(X(t_1)=a\,|\,X(t_2)=b) > \Pr(X(t_3)=a)\bullet\Pr(X(t_2)=b\,|\,X(t_3)=a)\bullet\Pr(X(t_1)=c\,|\,X(t_2)=b). \quad (15)$$

A rearrangement of (15) that parallels what I did earlier for (7*) yields a necessary and sufficient condition for the truth of (15), namely that

$$\frac{\Pr(X(t_3)=c)}{\Pr(X(t_3)=a)} \quad (16)$$

not be too tiny. This ratio needn't be greater than one, though that suffices.

Moving from the three times in our running example to n times, it gets easier and easier for a history of monotonic increase to be more probable than a history of monotonic decline as n increases. Assuming (backward) time-translational invariance, the criterion for $\mathbf{IF_{upb}}$ to entail **THI** in a history that spans n times is that

$$(\mathbf{R2})\ \frac{\Pr(X(t_n)=x_n)}{\Pr(X(t_n)=x_1)}$$

not be too small. This requirement gets easier and easier to satisfy as n increases. Notice that **R2** describes states of the end time, not states of the start time, and the biggest value for $X(t_n)$ appears in the numerator, not the denominator.

In summary, if $\mathbf{IF_{upb}}$ is true, then **THI** is true under each of two assumptions: (i) there is (forward) time-translational invariance and the **R1** ratio isn't too small, and (ii) there is (backward) time-translational invariance and the **R2** ratio isn't too small.

Declaration of interest

None.

Funding sources

The author acknowledges support from the William F. Vilas Trust of the University of Wisconsin, Madison, USA.

References

Albert, D. Z. (2000). *Time and chance*. Harvard University Press.

Barrett, M., & Sober, E. (1994). The second law of probability dynamics. *The British Journal for the Philosophy of Science, 45*, 941–953.

Callender, C. (1995). The metaphysics of time reversal. *The British Journal for the Philosophy of Science, 46*, 331–340.

Christenson, H., Cronin, J., Fitch, L., & Turlay, R. (1964). “Evidence for the 2π decay of the K_2^0 meson. *Physics Review Letters, 13*, 138.

Earman, J. (2002). What time reversal invariance is and why it matters. *International Studies in the Philosophy of Science, 16*, 245–264.

Earman, J. (2006). The past hypothesis – not even false. *Studies in History and Philosophy of Modern Physics, 37*, 399–430.

Ehrenfest-Afanassjewa, T. (1925). *Over een misverstand betreffende de toepassing van de waarschijnlijkheidstheorie op de leer van de irreversibiliteit der natuurverschijnselen.” (“About a misunderstanding concerning the application of probability theory to the doctrine of the irreversibility of natural phenomena”) Verslag van de gewone vergaderingen der Afdeeling Natuurkunde van de KNAW (Report of the regular meetings of the Physics Department of the Royal Netherlands Academy of Arts and Sciences)* (Vol. 34, pp. 598–600).

Ehrenfest, P., & Ehrenfest-Afanassjewa, T. (1907). Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem. *Physikalische Zeitschrift, 8*, 311–314.

Kelly, F. (1979). *Reversibility and stochastic networks*. Wiley.

Loewer, B. (2007). Counterfactuals and the second law. In H. Price, & R. Corry (Eds.), *Causation, physics, and the Constitution of reality: Russell's republic revisited*. Oxford University Press.

Sober, E. (1993). Temporally oriented laws. *Synthese, 94*, 171–189.

Sober, E. (2015). *Ockham's razors – a user's manual*. Cambridge University Press.

Uffink, J. (2007). Compendium to the foundations of classical statistical physics. In J. Butterfield, & J. Earman (Eds.), *Handbook for the philosophy of physics* (pp. 924–1074). Elsevier.

Wallace, D. (2010). Gravity, entropy, and cosmology – in search of clarity. *The British Journal for the Philosophy of Science, 61*, 513–540.

Winsberg, E. (2004). Can conditioning on the ‘past hypothesis’ militate against the reversibility objections? *Philosophy of Science, 71*, 489–504.

Please cite this article as: Sober, E., Histories, dynamical laws, and initial conditions –Invariance under time-reversibility and its failure in Markov processes, with application to the second law of thermodynamics and the past hypothesis, Studies in History and Philosophy of Modern Physics, https://doi.org/10.1016/j.shpsb.2019.08.004