

Justification in Statistical Mechanics.

Kevin Davey

Abstract

According to a standard view of the second law of thermodynamics, our belief in the second law can be justified by pointing out that low entropy macrostates are less probable than high entropy macrostates, and then noting that a system in an improbable state will tend to evolve toward a more probable state. I would like to argue that this justification of the second law of thermodynamics is fundamentally flawed, and will show that some puzzles sometimes associated with the second law are merely artifacts of this incorrect justification.

1 The Standard Story.

If we squirt some colored ink into a closed container of moving water, we are justified in expecting that the ink will eventually disperse itself evenly through the water. Indeed, if we prepare any system in a state of low entropy, and then isolate the system, we are justified in expecting that the system will experience an increase in entropy. In virtue of what are these beliefs justified? Such beliefs can, of course, be justified on purely inductive grounds. But one might also think that such beliefs ought to be able to be justified on the basis of more fundamental physical and mathematical principles. In this paper, I would like to discuss and criticize a common justification of this later sort.

The challenge posed in the last paragraph is to explain why we are justified in believing the second law of thermodynamics, given only more fundamental physical and mathematical principles. For our purposes, the second law of thermodynamics states that the entropy of an isolated system will increase until the system reaches its equilibrium state, which is a state of maximum entropy. Once in this equilibrium state, the system will remain there (or at least will remain there for a very long time.)

A standard (though somewhat schematic) argument for the second law is as follows:

Standard Story: Low entropy macrostates occupy a tiny portion of phase space, and so it is extremely improbable that a system will find itself in a low entropy macrostate. High entropy macrostates occupy essentially all of phase space, and so it is extremely probable that a system will find itself in a high entropy macrostate. If a system has an extremely probable and an extremely improbable state, then whenever the system is in the extremely improbable state, it will very likely find itself in the extremely probable state shortly thereafter, and will likely remain in the extremely probable state for some time. (That, after all, is just what it *means* to call a state extremely probable.) Thus, a system in a low entropy macrostate is overwhelmingly likely to evolve into a high entropy macrostate, and remain in that high entropy macrostate for some time.

The primary goal of this paper will be to argue that this story is deeply misguided.

In §2, I argue that the Standard Story rests critically on what I will call a *Probability Principle*. A Probability Principle is any claim of the form:

Probability Principle: If we know that a system is in macrostate M , then we are justified in describing the microstate of the system with the probability measure μ_M , where μ_M is defined in terms of M as ...

Different procedures for defining μ_M correspond to different Probability Principles.

In §3, I argue that no Probability Principle can be correct, and that the Standard Story is thus fundamentally flawed. In §4, I argue that this sheds light on some problems traditionally associated with statistical mechanics. For instance, one problem, sometimes grouped together with the Reversibility Objections, is that statistical mechanics tells us that an isolated system in a medium entropy state is much more likely to have had a high entropy past than a low entropy past, contrary to our experience. Some sort of maneuvering is therefore needed in order to salvage our ability to reason about the past using statistical mechanics. Precisely what sort of maneuvers are needed is a subject of some dispute. Even once such maneuvers are made, however, many accounts leave us having to say that the universe began its life in an exceedingly improbable state. We are then left wondering precisely what scientific obligation we are under to explain this initial state. My arguments, however, will show that there is no clear sense in which statistical mechanics gets the past wrong, and no clear sense in which we are forced to say that the universe began its life in an exceedingly improbable state. Indeed, there is no clear sense in which low entropy states are improbable, and high entropy states probable. Many associated problems then largely evaporate.

In §5, as an ‘appendix’ of sorts, I will discuss a few reasons why some may have felt tempted to adopt Probability Principles, and will explain why these reasons are misguided.

2 Justifying the Standard Story.

It is well known that fleshing out the details of the Standard Story requires some sort of ergodicity or mixing hypothesis.¹ For the sake of definiteness, we invoke the following:

Mixing Hypothesis: Let S be phase space, and let μ be the standard Lebesgue measure on S , normalized so that $\mu(S) = 1$. Let M and Q be measurable subsets of S , and let M_t be the t -second evolution of M . Then:

$$\lim_{t \rightarrow \infty} \mu(M_t \cap Q) = \mu(M)\mu(Q).$$

To see how the Mixing Hypothesis gives us the second law, let Q be the set of equilibrium states of the system (so that $\mu(Q) \approx 1$), and let M be an arbitrary macrostate (assumed to have non-zero measure.) Then:

$$\lim_{t \rightarrow \infty} \frac{\mu(M_t \cap Q)}{\mu(M)} \approx 1 \quad (*)$$

This is usually interpreted as saying that almost all microstates in M eventually end up in Q . It follows that a system in macrostate M is almost certain to end up in Q . Thus, we have established a version of the second law.

The Mixing Hypothesis (and its close relatives) have been criticized by Sklar [6], Earman and Redei [4], and others. These authors suggest that there is no evidence that such assumptions hold of realistic physical systems, and that there is perhaps even evidence that such assumptions fail. I take these criticisms to be persuasive.

I think, however, that there is a more basic error that occurs in this justification of the second law. Suppose we have a system for which the Mixing Hypothesis happens to hold. Does the Standard Story then give us good reason to suppose that the second law holds for this a system? I shall argue for a negative answer. Even in the case of a system known to obey the Mixing Hypothesis, the Standard Story fails to justify our belief in the second law.

To see the main problem, we ask: does equation (*) *really* entail that a system in the macrostate M is extremely likely to end up in Q at some point in its future? The argument for an affirmative answer proceeds as follows: suppose our system is in macrostate M at

¹See, e.g., Chapter 5 of Sklar [6].

$t = 0$. We choose a probability measure μ_M over the set of all possible microstates that is uniform over M , and which vanishes elsewhere. (That is to say, if $T \subset S$, then the probability that our system is in a microstate in T is given by:

$$\mu_M(T) = \mu(M \cap T) / \mu(M),$$

where μ is the standard Lebesgue measure.) Given any subset $T \subset S$, and any real r , let T_r be the set of microstates which, r seconds ago, were elements of T . Given this notation and this choice of probability measure, the probability that t seconds from now our system will lie in Q is:

$$\mu_M(Q_{-t}) = \mu(M \cap Q_{-t}) / \mu(M).$$

Using Liouville's Theorem, $\mu(M \cap Q_{-t}) = \mu(M_t \cap Q)$. Thus, the probability that t seconds from now our system will lie in Q is:

$$\mu(M_t \cap Q) / \mu(M).$$

The Mixing Hypothesis tells us that this quantity approaches a number close to 1 as t increases. Thus, almost all microstates in M eventually end up in Q at some point in their futures.

But the cogency of this argument depends on our choice of the measure μ_M as the relevant probability measure with which to describe the present state of the system. In particular, the following principle is being invoked:

Probability Principle 1: If we know a system is in a macrostate M , then we are justified in describing the present state of the system with the probability measure μ_M that is uniform over M , but which vanishes elsewhere.

Note that this is a principle about what we are *justified* in believing. Recall that the whole point of the Standard Story is to explain why we are *justified* in making predictions in accordance with the second law. If our explanation depends on one choice of probability measure over another (as we shall see it does), then it better be the case that we are *justified* in choosing that particular probability measure over the other, or else the Standard Story will fail to serve its justificatory purpose. This is why we need a principle that tells us that we are *justified* in our choice of a particular probability measure.

Without an assumption like Probability Principle 1, we are in no position to interpret (*) as stating that a system in the macrostate M is extremely likely to end up in Q in the future. To see why, let M be a macrostate, and let $x \in M$ be a microstate which never passes

through the equilibrium macrostate Q .² Define μ_x to be a probability measure on M which is supported only on x ; i.e., for any subset T of phase space,

$$\mu_x(T) = \begin{cases} 1 & \text{if } x \in T \\ 0 & \text{if } x \notin T \end{cases}$$

Using this measure of probability, a system in the macrostate M is *certain* to remain outside of Q forever, even though (*) holds. Thus, it is only if we draw some sort of connection between Lebesgue measure and the actual probability measure pertinent to the present state of the system, that we can take (*) to be telling us that a system in the macrostate M is extremely likely to end up in Q . This sort of connection is explicitly drawn in Probability Principle 1.

Even if the Mixing Hypothesis holds, the Standard Story allows us to infer the second law only if we assume something like Probability Principle 1. But is Probability Principle 1 correct? The answer is no. Let M be a non-equilibrium macrostate of a system, and suppose we know that our system is in M . Almost all microstates in M have a higher entropy past. If we are justified in describing the microstate of the system with the probability measure μ_M described by Probability Principle 1, then we are justified in believing that our system almost certainly had higher entropy in the recent past. But we have strong inductive grounds for believing that our system did not have higher entropy in the recent past. Therefore, Principle 1 is false. (Essentially this line of argument is developed in Chapter 4 of Albert [1].)

The natural thing to do at this point is modify Probability Principle 1. Here is a proposal that circumvents the problem:

Probability Principle 2: If we know a system is in a macrostate M , then we are justified in describing the microstate of the system with the probability measure μ_{M^*} that is uniform over M^* , but which vanishes elsewhere, where M^* is the set of microstates $x \in M$ such that x 's entropy was lower in the recent past.

(Here, by the 'recent past', we mean the interval of time $[-s, 0]$ over which we have good reason for believing that our system has been isolated.) Here is another proposal:

Probability Principle 3: Same as Probability Principle 2, except M^* is now

²For instance, in the case of hard spheres scattering in a closed rectangular container (one of the few cases for which the Mixing Hypothesis may be proved to hold), let x be a microstate in which all particles have velocities parallel to a fixed wall of the container, and in which no collisions between particles ever occur.

the set of microstates $x \in M$ compatible with the early universe being in some particular macrostate at some particular time.

And here is another:

Probability Principle 4: Same as Probability Principle 2, except M^* is now the set of microstates $x \in M$ that behave in a way that is compatible with all our justified expectations about the past and future of the system.

Regardless of which of these principles is chosen, we can argue from the Mixing Hypothesis that

$$\lim_{t \rightarrow \infty} \frac{\mu(M_t^* \cap Q)}{\mu(M^*)} \approx 1 \quad (**)$$

This tells us that almost all microstates in M^* eventually end up in Q . Using the relevant Probability Principle, it follows that a system in macrostate M is very likely to end up in Q . Thus, we are justified in predicting that our system will behave in accordance with the second law.

But are any of Probability Principles 2, 3 or 4 true? I claim that they are not. In fact, I claim that all principles of the form:

Probability Principle: If we know a system is in macrostate M , then we are justified in describing the microstate of the system with the probability measure μ_M , where μ_M is defined in terms of M as ...

are false.

But without a Probability Principle of this sort, it is difficult to see how the Standard Story can explain why we are justified in predicting that a system will obey the second law, even when it is known that the system obeys the Mixing Hypothesis. I move to my argument for the falsehood of Probability Principles now.

3 Against Probability Principles.

The main problem with any Probability Principle is that, according to such a principle, knowing the current macrostate M of a system is sufficient for us to have a justified belief about the probability with which our system lies in any sub-region of phase space. But it is perfectly possible – and indeed, almost always the case – for us to know the current macrostate of a system, without being in the position to form a justified belief about the

probability measure with which to describe the microstate of the system. Therefore, all Probability Principles are false.

Consider the following situation. Suppose a scientist brings in a glass box, which we observe to be filled with a gas in its equilibrium macrostate M . Suppose the scientist tells us that 10 minutes ago the gas was in equilibrium, and at that time he made a choice between:

Option A. Leaving the gas undisturbed, in its equilibrium macrostate, or

Option B. Forcing the gas out of its equilibrium macrostate, into a specific macrostate that typically takes 10 minutes to relax back into equilibrium.

The scientist will not tell us which option he chose, and knowing that the system is presently in its equilibrium macrostate does not help us deduce which choice was made. The scientist does tell us, however, that 1 hour ago he selected a real number α with $0 < \alpha < 1$ (by some unknown procedure) and that he then devised some random procedure which guaranteed that he chose Option A with probability α , and Option B with probability $1 - \alpha$.

For no value of α are we justified in believing that α was the probability chosen by the scientist. We may assume that we know nothing about the scientist, and nothing about the procedure by which he chose α . We do not even whether his selection of α was truly random, so we are not entitled to claim a uniform probability distribution for α .

Suppose, however, that some Probability Principle tells us that we are justified in describing the present state of our system with a probability measure μ^* . Let N be the set of microstates which, 10 minutes ago, were in the particular out of equilibrium macrostate referred to in Option B. From the fact that we are justified in using the probability measure μ^* , it follows that we are justified in inferring that $\alpha = \mu^*(N)$. But we are not justified in believing that $\alpha = \mu^*(N)$. Therefore, the Probability Principle is false. And so no Probability Principle can be correct – and it is worth noting that this is so even when the macrostate M is an equilibrium macrostate.³

The general lesson here is that the correct probability measure with which to describe the present state of a system is, in general, determined by all sorts of facts about the past of the system – including, for instance, the way in which the system was prepared. Such facts about the past of the system need not be known by us, nor need we be in a position

³One can easily generate similar examples in which the final macrostate M is not an equilibrium macrostate. For instance, perhaps the scientist presents us with a glass of water with a thin film of ice on the surface, and tells us that 1 hour ago, he either placed 6 ice cubes in a glass of water at room temperature (with probability α) or 7 slightly less cold ice cubes in a glass of water at room temperature (with probability $1 - \alpha$.)

to form justified beliefs about such facts. Because of this, we will generally not be justified in making claims about the correct probability measure with which to describe the present state of the system.⁴

Of course, this is not to deny that there might be cases in which we *can* have a justified belief about which probability measure describes the current state of a system⁵; nor it is to deny that we can always make an *unjustified* claim or conjecture about the probability measure describing the current state of the system. All I intend to deny is that there is some sort of effective procedure that can be used to reliably move from knowledge of the present macrostate of a system to a justified probability measure describing the state of the system.

One might think that the problem here is that we need to adopt a principle that includes information about the past of the system. For instance, consider:

Historic Probability Principle: Suppose we know that a system has been isolated during the interval of time from $t = -r$ to $t = 0$. Suppose also that at any time s in this range, the system is known to have been in macrostate M_s . Then we are justified in describing the microstate of the system with the probability measure μ , where μ is defined in terms of $\{M_t | t \in [-r, 0]\}$ as follows ...

This is not a Probability Principle as defined earlier, nevertheless, it is worth mentioning two of its serious defects. First, in general we do not know the past macroscopic history of systems of interest, and so relying on a Historic Probability Principle severely limits the sorts of situations in which we are justified in expecting a system to obey the second law. It is for this reason that we have focused on Probability Principles having the form described earlier.

But second, let us suppose the Historic Probability Principle tells us that we are justified in describing the present microstate of a system with probability measure μ , where we take the system to have been isolated over the period of time $[-r, 0]$. Let μ_{-r} be the probability measure that evolves into μ over an r second period. We are then justified in describing the state of the system at $t = -r$ with the probability measure μ_{-r} . But in general, we

⁴This phenomenon also arises in the case of non-Markovian dynamics. Here, the future behaviour of a system depends not only on the present state of the system, but also on the history of the system. If we know the present macrostate of such a system, but do not know its history, we will generally not be in a position to specify a probability measure over all possible present microstates.

⁵For instance, we might have strong inductive grounds for describing the state of a coin tossed in the air with a probability measure that is uniform over the set of all possible angular orientations.

need not be in a position to form a justified belief as to which probability measure correctly describes the initial microstate of the system. Just as before, the correct probability measure may depend on things completely unknown to us, such as the way the system was prepared at the moment of isolation. We should therefore be suspicious of any Historical Probability Principle. Although a more detailed analysis of Historical Probability Principles would be welcome, we will not pursue this here, but will focus on the class of Probability Principles defined earlier.

4 Consequences

4.1 Entropy and Probability

What effect does all this have on our understanding of the second law, and in particular, the Standard Story? According to the Standard Story, because low entropy macrostates occupy a tiny portion of phase space, they should be assigned low probability. But it is difficult to see how this claim can be justified, other than with a Probability Principle. Because Probability Principles are incorrect, the Standard Story falls into jeopardy.

Furthermore, once we free ourselves from the shackles of incorrect Probability Principles, the situation for the Standard Story gets even worse. For I claim that the low entropy states actually found in nature need not be improbable at all. If low entropy states are not necessarily improbable, then it cannot be correct to say that the second law of thermodynamics is just an expression of the fact that a system will tend to move from an improbable to a more probable state. The key intuition behind the Standard Story thus disintegrates.

How might a low entropy state turn out *not* to be improbable? In order to address this question, let us first consider the following counter-argument:

We should think of low entropy states as improbable, because it is exceedingly unlikely that an isolated system, in a high entropy state, will fluctuate out of that state and into a lower entropy state shortly thereafter.

There is, I think, something profoundly irrelevant about the central observation of this argument; for although it is possible for a system in equilibrium to fluctuate out of equilibrium and into a lower entropy state, very few systems we actually find in low entropy states are in such states as a result of such a fluctuation. Most of the world's glasses of ice-water, for instance, are the products of very deliberate interactions in which ice and water end up

mixed, and are not the result of random thermal fluctuations that begin with an isolated glass of water. In what sense, then, is the glass of ice-water that we actually find in nature in an improbable state? There is simply no straightforward sense in which this is so, as the following examples help to make clear.

Suppose we live in a world in which, by government fiat, all good citizens must do everything they can to keep all glasses of water half full of ice, and that the citizenry is largely successful at this. Suppose we find a glass of water with no ice in it. We will be justified in believing that one hour ago, the glass of water probably had ice in it, because experience suggests that most glasses of water almost always have ice in them. The state of the glass of water in which it contains ice is thus highly probable, even though it has low entropy.⁶ In this world, when a glass of ice-water melts, it moves from a probable state to an improbable state.

By contrast, consider a world in which, by government fiat, all good citizens must do everything they can to keep all glasses free of ice, and that the citizenry is largely successful at this. Suppose we find a glass of water with no ice in it. We will be justified in believing that one hour ago, the glass of water probably had no ice in it, because most glasses of water almost never have ice in them. The state of the glass of water in which it contains ice is thus highly improbable.⁷ In this case, when a glass of ice-water melts, it moves from an improbable state to a probable state.

The lesson here is that whether a low entropy state of a system has low probability, and whether a high entropy state of a system has high probability, depends on the environment in which the system exists, and the sorts of interactions the system is likely to undergo. There is no straightforward, a-priori connection between low entropy and low probability. It cannot then be the case that the second law of thermodynamics is just an expression of the fact that a system will tend to move from an improbable to a more probable state.

To all this, a critic might reply that the universe is destined to undergo heat death, and so in the long run, low entropy states end up being very rare, and hence improbable. But this point is surely irrelevant. If we want to inductively justify statements about the way that sub-systems of the universe will behave when the universe is in a particular far-from-equilibrium state, then we will focus on our experience of the universe while it is in that particular far-from-equilibrium state. The fact that, on cosmic time scales, that particular

⁶A principle such as Probability Principle 1 gets this case wrong.

⁷A principle such as Probability Principle 2 gets this case wrong.

far-from-equilibrium state occupies the mere blink of an eye is irrelevant to the justification of statements about what happens when the eye is in fact blinking.

An interesting corollary should be noted. There has been some literature (see, for instance, [2]) on the question of whether an especially improbable initial state of the universe is the sort of thing that requires explanation. This discussion is motivated by the fact that the universe is thought to have started in an extremely low entropy state. If we take extremely low entropy states to be extremely improbable states, we are then forced to say that the universe started in an extremely improbable state; and we must decide how to react to this. But once we abandon the idea that low entropy means low probability, there remains no clear sense in which the universe began in an extremely improbable state. Indeed, once we abandon all Probability Principles, we are no longer even required to take probabilistic claims about the universe's initial state to be *meaningful*. This is not to deny that such claims could be meaningful, nor is it to deny that the universe may turn out to have originated in an extremely improbable state. However, neither the meaningfulness nor the truth of the claim that the universe originated in an improbable state follow from the mere fact that the universe started in an extremely low entropy state. The burden rests on someone who nevertheless believes that the universe started in an extremely improbable state to demonstrate otherwise.

4.2 Does Statistical Mechanics Get the Past Wrong?

Another conceptual problem with statistical mechanics is that it can appear to give us information about the past at odds with our experience. For instance, suppose we adopt Probability Principle 1. We are then forced to claim that a system in a medium entropy state most likely had higher entropy in the recent past. However, we surely have strong inductive grounds for thinking that the system had lower entropy in the recent past, and so we have a problem.

The problem, of course, is with our choice of μ . Suppose a probability distribution μ that we use to describe the present state of a system entails some proposition X about the past of our system, where X is at odds with our experience. The correct conclusion to draw is simply that we are not justified in using μ to describe the present state of the system – for if we were, we would then be justified in believing X , which we are not.

In general, any choice of probability measure to describe the present state of a system must be justifiable on inductive grounds. So, for instance, if we are justified in describing

the present state of a system with a probability measure μ for which $\mu(X) \approx 1$ (for some proposition X), then it must be the case that we have strong inductive grounds for believing that X is true. The only case in which statistical mechanics can tell us to believe something about the past that is inconsistent with our best inductively formed expectations is if we use an unjustified probability measure μ . Only in this relatively uninteresting sense is statistical mechanics capable of getting the past wrong.⁸

It will be instructive to apply these considerations to the ‘skeptical catastrophe’ discussed by Albert in Chapter 6 of [1].⁹ The problem is this: suppose I only have knowledge of the present macrostate of a small portion of the universe, and that I am presented with a 50 year old photograph of my grandmother. Experience suggests that this photograph was probably formed in a less ragged state (i.e., in a lower entropy state) some time ago, as the result of an interaction of a camera with my grandmother. If, however, we describe the microstate of a closed subsystem of the universe with a probability measure that is uniform over its present macrostate, then it turns out to be exceedingly unlikely that the photograph had such a past. Instead, it is much more likely that the photograph is the result of some sort of fluctuation from equilibrium of a piece of photographic paper. There is nothing special about photographs here – essentially the same point can be made about *all* records of the past, including our memories. And so we must conclude that pretty much all our knowledge about the past, insofar as it rests on records, memories and relics, is unreliable at best, and false at worst. This is a skeptical catastrophe.

Albert resolves this catastrophe by adopting his ‘Past Hypothesis’, but a much simpler solution is available. On one hand, the Probability Principle on which the argument is based is false. The skeptic, however, may reply that once we place ourselves in the position in which we only have knowledge of the present, we then have *no grounds* for thinking that the relevant Probability Principle is false. Ignoring the question of whether there is an unfair shift in the burden of proof here, this reply quickly gets the skeptic in even more trouble. For suppose we only have knowledge of the present macrostate of some chunk of the universe. A probability measure can only be justified on inductive grounds, and so if all our knowledge is knowledge of the present, then we have no good inductive grounds with which to justify *any* probability measure. But without a justified probability measure, statistical mechanics yields no *justified* beliefs about the past. And if statistical mechanics

⁸Although Albert appears to draw a similar conclusion, he then goes on to adopt a Probability Principle of the sort criticised earlier.

⁹The phrase ‘skeptical catastrophe’ is taken from p. 116 of [1].

yields no justified beliefs about the past, then statistical mechanics has no consequences capable of conflicting with common sense. Thus, there is no skeptical catastrophe.

5 Defenses of Probability Principles.

Before finishing, it will be of interest to consider some arguments that might be given for various Probability Principles. We will consider arguments based on the Principle of Indifference, arguments based on the stipulation that a probability measure be stationary, and arguments based on ergodicity/ frequentist considerations.

5.1 The Principle of Indifference.

What sorts of arguments might be given to support Probability Principle 1? Perhaps the most straightforward way to justify this choice of probability distribution is to invoke the Principle of Indifference, according to which we should choose a uniform probability measure over a set of possibilities whenever we lack a reason to choose a non-uniform probability measure over that set of possibilities. This sort of reasoning can also be used to justify Probability Principles 2, 3 and 4, and even certain Historical Probability Principles.

The Principle of Indifference as thus stated has been much criticized; see especially Chapter 12 of Van Fraassen's [7].¹⁰ Some critics have pointed out that the Principle of Indifference fails to deliver a unique probability distribution; this has become known as 'Bertrand's Paradox', and is a serious problem for the Principle. An equally compelling, but separate, concern is that the absence of a reason for choosing a non-uniform probability measure surely does not *justify* the choice of *any* probability measure, let alone a uniform probability measure. A similar point has also been made by Sklar; see pp. 118-120 of [6].

5.2 Stationarity Arguments.

Let us focus on the case in which our system is in equilibrium. Let μ be the probability measure which is uniform over all phase space. Then μ has the property that it is 'stationary'; that is to say, it remains invariant under evolution in time. Furthermore, if the system is ergodic, then μ is the *only* stationary probability distribution that assigns probability 0 to sets of measure 0.¹¹

¹⁰For a defense of the Principle of Indifference, see [3].

¹¹For a discussion of this point, see pp. 159-161 of Sklar [6].

Given that we expect the properties of a system in equilibrium to remain constant over time, and given the reasonableness of assigning probability 0 to sets of measure 0, we should conclude that when an ergodic system is in equilibrium, then μ is the probability measure with which to describe the microstate of the system.

This Probability Principle is quite modest, for it is only intended to apply to an ergodic system in equilibrium. Nevertheless, it is false. While we have strong inductive grounds for thinking that the properties of an isolated system in equilibrium will not change in the future, we have no grounds for thinking that such a system's properties did not change in the past. Many isolated systems presently in equilibrium were isolated in an out-of-equilibrium state. If we know that an isolated system was prepared in an out-of-equilibrium state, or if we are not in a position to judge whether a system was out-of-equilibrium when first isolated, then we are not justified in employing a stationary probability measure to describe the present microstate of the system.¹²

5.3 Ergodicity and Frequentism.

Finally, suppose a system is ergodic. Fix a sub-region $R \subset S$ of phase space S , and let $\alpha = \mu(R)/\mu(S)$. Then, for almost all points x of phase space, the proportion of time that a system beginning in microstate x will spend in R is just α . If we identify the probability of a system being in R with the proportion of time it spends in R , as a frequency interpretation of probability might suggest, then we must conclude that the probability of our system being found in R is α . Thus, sub-regions of phase space with equal volume must be assigned equal probability. And this means that, given no additional information about the system, we ought to describe the present microstate of the system with a probability distribution that is uniform over all of phase space.

But similar problems arise. Typically, we do *not* know nothing about a system; the additional knowledge we have about a system will generally allow us to argue that the system is *more* likely to be in one sub-region of phase space than another equally sized subregion. A simple example goes back to Reichenbach [5] – if the weather has been hot for the last few days, it is more likely that the weather will be hot tomorrow than that it will be cold, even if it is assumed that, in the long run, hot and cold temperatures are equally frequent. In this way, information about the recent past can have a bearing on our choice of

¹²Another serious problem is that it is not clear why we should assign probability 0 to a set of measure 0. See pp. 182-188 of Sklar [6] for a discussion of this point.

probability measures for the present state of a system.¹³

This is not to concede, however, that we are justified in using a uniform probability measure in cases in which we have absolutely no grounds for making any claims about the past of a system. For if we truly have no grounds for making any claims about the past of a system, then we are surely not justified in describing the present state of the system with *any* probability measure.

5.4 Conclusion

Although there are undoubtedly other arguments for Probability Principles, I conjecture that they all fail for similar reasons. There is no way to account for the variety of knowledge we might or might not have about the past of a system with a simple, a-priori principle that focuses solely on the present macrostate of the system. Once this is realized, some (though not all) of the conceptual problems associated with statistical mechanics disappear.

References

- [1] Albert, D. *Time and Chance*, Harvard University Press, 2000.
- [2] Callendar, C., ‘Measures, Explanation and the Past: Should Special Initial Conditions Be Explained?’, *British Journal for the Philosophy of Science*, 55 (2004), 195-217.
- [3] Castell, P., ‘A Consistent Restriction of the Principle of Indifference’, *Brit. J. Phil. Sci.*, 49 (1998), 387-395.
- [4] Earman, J. and Redei, M., ‘Why Ergodic Theory Does Not Explain the Success of Equilibrium Statistical Mechanics’, *British Journal for the Philosophy of Science*, 47 (1996), 63-78.
- [5] Reichenbach, H., *The Direction of Time*, 2000. (Originally University of California Press, 1956.)
- [6] Sklar, L., *Physics and Chance*, Cambridge University Press, 1993.
- [7] Van Fraassen, B., *Laws and Symmetry*, Oxford University Press, 1990.

¹³This objection seems closely related to point (3) on p. 159 of [6].