

On the Origins of the Arrow of Time: Why There is Still a Puzzle about the Low Entropy Past

1. The most underrated discovery in the history of physics?

Late in the nineteenth century, physics noticed a puzzling conflict between the laws of physics and what actually happens. The laws make no distinction between past and future—if they allow a process to happen one way, they allow it in reverse.¹ But many familiar processes are in practice ‘irreversible’, common in one orientation but unknown ‘backwards’. Air leaks out of a punctured tyre, for example, but never leaks back in. Hot drinks cool down to room temperature, but never spontaneously heat up. Once we start looking, these examples are all around us—that’s why films shown in reverse often look odd. Hence the puzzle: What could be the source of this widespread temporal bias in the world, if the underlying laws are so even-handed?

Call this the *Puzzle of Temporal Bias*, or *PTB* for short. It’s an oft-told tale how other puzzles of the late nineteenth century soon led to the two most famous achievements of twentieth century physics, relativity and quantum mechanics. Progress on PTB was much slower, but late in the twentieth century cosmology provided a spectacular answer, or partial answer, to this deep puzzle. Because the phenomena at the heart of PTB are so familiar, so ubiquitous, and so crucial to our own existence, the achievement is one of the most important in the entire history of physics. Yet it is little-known and underrated, at least compared to the other twentieth century solutions to nineteenth century puzzles.

Why is it underrated? Partly because people underestimate the original puzzle, or misunderstand it, and so don’t see what a big part of it is addressed by the new cosmology. And partly for a deeper, more philosophical reason, connected with the view that we don’t need to explain initial conditions. This has two effects. First, people undervalue the job done so far by cosmology, in telling us something very surprising

¹In some rare cases discovered in the mid-twentieth century, the reversal also needs to replace matter with antimatter, but this makes no significant difference to anything discussed below.

about the early history of the universe—something that goes a long way towards explaining the old puzzle. And second, they don't see the importance of the remaining issues—the new issues thrown up by this story, about *why* the early universe is the way that modern cosmology reveals it to be.

I'm going to argue that the old philosophical view is mistaken. We should be interested in the project of explaining initial conditions, at least in this case. As a result, we should give due credit—a huge amount—to modern cosmology for what it has already achieved. And we should be interested in pushing further, in asking *why* the early universe is the way modern cosmology has revealed it to be.

To understand why these issues matter, we need to understand how what cosmology tells us about the early history of the universe turns out to be relevant to the puzzle of temporal bias. Let's begin in the nineteenth century, where the puzzle first comes to light.

2. PTB in the age of steam

In one sense, the temporal bias discovered by physics in the nineteenth century had never been hidden. It was always in full view, waiting for its significance to be noticed. Everybody had always known that hot things cooled down in a cooler environment, that moving objects tended to slow down rather than speed up, and so on. But two things changed in the nineteenth century. First, examples of this kind came to be seen as instances of a single general tendency or law—roughly, a tendency for concentrations of energy to become more dissipated, less ordered, and less available to do work. (The impetus for a lot of the relevant physics was the goal of extracting as much work as possible from a given lump of coal.) A measure of this disorder came to be called 'entropy', and the general principle then said that the entropy of a closed system never decreases. This is the famous *second law of thermodynamics*.

After formulating this general principle in the mid nineteenth century, physicists set about trying to explain it. Early work looked at the case of gases. Other properties of gases had already been successfully explained by assuming that gases were huge swarms of tiny particles, and applying statistical techniques to study the average

behaviour of these swarms. Building on earlier work by James Clerk Maxwell (1831–79), Ludwig Boltzmann (1844–1906) argued in the 1870s that the effect of collisions between randomly moving gas molecules was to ensure that the entropy of a gas would always increase, until it reached its maximum possible value. This is Boltzmann’s *H-theorem*.

This connection between thermodynamics and mechanics provides the crucial second ingredient needed to bring PTB into view. We’ve seen that PTB turns on the apparent conflict between two facts: the lack of temporal bias in the underlying laws, and the huge bias in what we can now call thermodynamic phenomena. In 1876, Boltzmann’s colleague Josef Loschmidt (1821–95) called attention to the first of these facts, noting that the Newtonian mechanics assumed to be guiding gas molecules is time-symmetric, in the sense that every process it allows to happen in one direction is also allowed in reverse. (Mathematically, all we need to do to change a description of one to a description of the other is to change the sign of all the velocities at a given instant.) Loschmidt’s immediate point was that if the second law rests on mechanics, it can’t be exceptionless. There are *possible* motions such that entropy decreases.

In response to Loschmidt, Boltzmann suggested a completely new way of thinking about second law. In place of the idea that collisions *cause* entropy increase, he offers us a new idea. Corresponding to any description of the gas in terms of its ‘macroscopic’ observable properties—temperature, pressure, etc.—there are many possible ‘microstates’—many possible configurations of molecules which all give the same macrostate. Boltzmann’s insight is that for non-equilibrium macrostates, the vast majority of these microstates are ones such that entropy increases in the future. It doesn’t have to do so, but usually it will. (It isn’t a trivial matter how to carve up the space of possible microstates, to get the measure right. An important part of Boltzmann’s contribution was to find the right way to do this.)

Still, Loschmidt’s argument turned on the realisation that thanks to the underlying time-symmetry of Newtonian mechanics, microstates come in pairs. For every possible microstate in which some process is occurring in one temporal direction,

there's another microstate in which the same process is occurring in the opposite temporal direction. So where does the asymmetry come in, on Boltzmann's new picture? To ask this question is to be struck by PTB.

This issue evidently occurred to Boltzmann at this time. His response to Loschmidt includes the following note:

I will mention here a peculiar consequence of Loschmidt's theorem, namely that when we follow the state of the world into the infinitely distant past, we are actually just as correct in taking it to be very probable that we would reach a state in which all temperature differences have disappeared, as we would be in following the state of the world into the distant future. (Boltzmann 1877, at p. 193 in translation in Brush 1966).

Thus Boltzmann seems to suggest that on the large scale, there is no temporal bias (and hence no PTB). But then why do we observe such striking asymmetry in our own region of space and time, if it doesn't exist on the large scale? And why is entropy so low now, given that according to Boltzmann's own way of counting possibilities, this is such an unlikely way to be?

Boltzmann doesn't seem to have asked these questions in the 1870s, and for the next twenty years PTB dropped back out of sight. It surfaced again in the 1890s, in a debate about the H-theorem initiated by E. P. Culverwell, of Trinity College, Dublin. As one contemporary commentator puts it, Culverwell's contribution was to ask in print "the question which so many [had] asked in secret, ... *'What is the H-theorem and what does it prove?'*" (Hall 1899, p. 685) This debate clarified some important issues, as we'll see in a moment. All the same, no one involved—not even Boltzmann—seems to have seen how much his new approach, formulated twenty years earlier in response to Loschmidt, had actually superseded the H-theorem.

This is an early manifestation of a confusion which has persisted in the subject ever since. To avoid it, we need to distinguish two different approaches to explaining the temporal bias of thermodynamics. As we'll see, both approaches face the question we're really interested in—Why is entropy so low, now and in the past?—but they have different conceptions of what else an explanation of the second law requires. On one

conception, the most interesting issue about time-asymmetry is somewhere else. This is one source of the tendency to undervalue the new contribution from cosmology, so it is important to draw distinctions carefully at this point.

3. How many asymmetries do we need?

What would it take to explain the temporal bias of thermodynamic phenomena? Since what needs to be explained is a time-asymmetry, it's a safe bet that an adequate explanation is going to contain some time-asymmetric ingredient. (Symmetry in, symmetry out, after all.) But there are two very different views about how many asymmetries we need. On some views we need only one, on others we need two.

The two-asymmetry approach

As the name suggests, the second law of thermodynamics was originally regarded as a physical *law*. Without delving into the philosophical issue about what this means, let's say that to think of the second law in this way is to think of it as having some kind of 'force' or necessity. In some sense, what the second law dictates is 'bound' to happen. The discovery that the second law is probabilistic rather than exceptionless doesn't necessarily undermine this conception. It simply means we need a constraint weaker than outright necessity—some kind of real 'propensity', for example.

Given this view of the second law, the task of explaining it in mechanical terms looks like the task of finding some mechanical factor which 'forces' or 'causes' entropy to increase (at least with high probability). The H-theorem itself is one such approach—it rests on the idea that the randomising effect of collisions between molecules causes entropy to increase. One of the major insights of the debate about the H-theorem in the 1890s was that if entropy is to increase, this causal mechanism must be time-asymmetric. If the H-theorem worked equally well in both directions, it would show that entropy is non-decreasing in both directions—which is only possible if it is constant.

How does this asymmetry get into the H-theorem? The first person to answer this question explicitly was Samuel Burbury (1831–1911), an English barrister who'd

turned to mathematical physics late in middle age, as loss of hearing curtailed his legal career. Burbury saw that the source of the asymmetry in the H-theorem is an assumption, roughly, that the motions of gas molecules are independent *before* they collide. He pointed out both that the H-theorem requires this assumption, and that if entropy is to increase, the assumption cannot hold *after* collisions. (See Burbury 1894, 1895. Burbury's argument is widely misinterpreted as showing that collisions *cause* correlations. In fact it shows no such thing. The correlations are simply those required by the assumption that entropy decreases towards the past, and are quite independent of whether the molecules collide at all. See Price 2002b.)

There are other causal approaches to the second law. One, called interventionism, attributes the increase in entropy to random and uncontrollable influences from a system's external environment. Another is a recent suggestion that a stochastic collapse mechanism proposed in certain extensions of quantum theory provides a randomising influence sufficient to ensure that entropy increases (see Albert 1994, 2000). Again, the point to keep in mind is that as in all such causal approaches, the mechanism needs to be time-asymmetric, if it is not to force entropy to be non-decreasing in both directions.

These causal approaches thus need two time asymmetries altogether. Why? Because a causal mechanism that ensures that entropy will not decrease won't by itself produce what we see. Entropy also needs to start low. If a system begins in equilibrium—i.e., with maximum possible entropy—such a mechanism will simply keep it there. There will be no observed *increase*. To get what we see, then, we need an asymmetric 'boundary condition' ensuring that entropy is low in the past, as well as an asymmetric mechanism to make it go up. What wasn't seen clearly in the 1890s, and has often been obscure since, is that this approach is thereby fundamentally different from the statistical approach suggested by Boltzmann in the 1870s, in response to Loschmidt. In Boltzmann's new approach, there is only one time-asymmetry—the only asymmetry is the low entropy boundary condition.

The one-asymmetry approach.

Think of a large number of gas molecules, isolated in a box with elastic walls. If the motion of the molecules is governed by deterministic laws, such as Newtonian mechanics, a specification of the microstate of the system at any one time uniquely determines its entire history (or ‘trajectory’). This means that Boltzmann’s assignment of probabilities to instantaneous microstates applies equally to whole trajectories. Accordingly, consider the set of all trajectories, with this Boltzmann measure. The key idea of Boltzmann’s statistical approach is that in the overwhelming majority of possible trajectories, the system spends the overwhelming majority of the time in a high entropy macrostate—i.e., among other things, a state in which the gas is dispersed throughout the container. And there is no temporal bias in this set of possible trajectories. Each possible trajectory is matched by its time-reversed twin, just as Loschmidt had pointed out.

Asymmetry comes in when we apply a low entropy condition at one end. For example, suppose we throw away all the possible trajectories *except* those in which the gas is completely confined to some small region R at the initial time T_0 . Restricted to the remaining trajectories, our original Boltzmann measure now provides a measure of the likelihood of the various possibilities consistent with this boundary condition—i.e., consistent with the gas’s being confined to R at T_0 . Almost all trajectories in this remaining set will be such that the gas becomes more dispersed after T_0 . The observed behaviour is thus predicted by the time-symmetric Boltzmann measure, once we ‘conditionalise’ in this way on the low entropy condition at T_0 .

On this view, then, there’s no asymmetric factor which ‘forces’ or ‘causes’ entropy to increase. This is simply the most likely thing to happen, given the combination of the time-symmetric Boltzmann probabilities and the single low entropy restriction in the past.

It’s worth noting that the correctness of the resulting probability judgements concerning the future behaviour implicitly *depends on the assumption* that that there is no corresponding low entropy restriction in that direction. So Boltzmann’s statistical

approach does not enable us to predict that entropy is unlikely ever to decrease, but only to draw a much weaker conclusion: entropy is unlikely to decrease, *unless there is the kind of constraint in the future which makes entropy low in the past*. The second law holds so long as there isn't a low entropy boundary condition in the future, but can't be used to exclude this possibility—*even probabilistically!*

As it stands, we know of no such condition in the future. The low entropy condition of our region seems to be associated entirely with a low entropy condition in our past. This condition is time-asymmetric, so far as we know, but this is the only time-asymmetry in play, according to Boltzmann's statistical approach.

Thus we have two very different ways of trying to explain the observed temporal bias of thermodynamic phenomena. Our current interest is in what these approaches have in common, the fact that entropy is low now, and even lower in the past. On both approaches, the observed asymmetry depends on this fact—without it, we'd never see stars and hot cups of coffee cooling down, because we'd never see stars and hot coffee, full stop. The great discovery I mentioned at the beginning is a cosmological explanation of this crucial fact.

I've stressed the distinction between these approaches because although the low entropy boundary condition plays a crucial role even in the two-asymmetry approach, it plays second fiddle there to the supposed asymmetric cause. Looking for an elusive factor that forces entropy to increase, the two-asymmetry approach often pays little heed to what seems a mere boundary condition, the fact that entropy starts low. This is one source of the tendency to discount modern cosmology's contribution to the solution of PTB.

The two-asymmetry approach faces a serious problem. To say that some asymmetric mechanism *causes* entropy to increase, is to say that in the absence of that mechanism, entropy would not increase. Yet Boltzmann claims to have shown that for most possible initial microstates, entropy would increase anyway, without any such asymmetric mechanism. So friends of such mechanisms need to say that Boltzmann is

wrong—that the universe (probably) starts in a microstate such that without the mechanism, entropy would not increase. It’s hard to see what could justify such a claim. (I develop this objection in Price 2002a and 2002b.)

For present purposes, however, we needn’t try to adjudicate between the two views. Our interest is in the low entropy boundary condition, and in the issue as to whether it needs to be explained. So long as we keep in mind that even the two-asymmetry approach needs such a condition, we’ll be in no danger of thinking that the issue is optional—a product of a questionable conception of what an understanding of PTB requires.

4. Did it all happen by accident?

In the discussion of the mid-1890s, Boltzmann himself certainly saw the importance of the question as to why entropy is now so low—much lower than its theoretical maximum. In a letter to *Nature* in 1895, he offers a tentative answer, based on a suggestion he attributes to his assistant, Dr Schuetz. The proposal is in two parts. First, he notes that although low entropy states are very unlikely at any given time, they are very likely to occur eventually, given enough time. After all, if we toss a fair coin for long enough, we’re very likely to get a run of a billion heads eventually—we’ll just have to wait a very, very long time! So if the universe is extremely old, it’s likely to have had time to produce, simply by accident, the kind of low entropy region we find ourselves inhabiting. As Boltzmann puts it: “Assuming the universe great enough, the probability that such a small part of it as our world should be in its present state, is no longer small.” (Boltzmann 1895, p. 415)

Of course, it’s one thing to explain why the universe contains regions like ours, another to explain why we find ourselves in one. If they are so rare, isn’t it much more likely that we would find ourselves somewhere else? Answering this challenge is the job of the second part of Boltzmann’s proposal. Suppose that creatures like us simply couldn’t exist in the vast regions of ‘thin cold soup’ between the rare regions of low entropy. Then it’s really no surprise that we find ourselves in such an unlikely spot. All intelligent creatures will find themselves similarly located. As Boltzmann himself puts it,

“the ... H curve would form a representation of what takes place in the universe. The summits of the curve would represent the worlds where visible motion and life exist.”
(Boltzmann, 1895, p. 415)

[insert figure 1 about here]

Figure 1. Boltzmann’s entropy curve

Figure 1 shows what Boltzmann here calls the H curve, except that following modern convention, this diagram shows entropy on the vertical axis, rather than Boltzmann’s quantity H . Entropy is low when H is high, and vice versa, so the summits of Boltzmann’s H curve become the troughs of this entropy curve. More precisely, Figure 1 shows the entropic history of a single typical trajectory. Most of the time—*vastly* more of the time than this page makes it possible to show—the universe is very close to equilibrium. Very, very rarely, a significant fluctuation occurs, an apparently random rearrangement of matter that produces a state of low entropy. As the resulting disequilibrium state returns to equilibrium, an entropy gradient is produced, such as the one on which we apparently find ourselves, at a point such as A. If intelligent life depends on the existence of an entropy gradient, it only exists in the regions of these rare fluctuations.

Why do we find ourselves on an uphill rather than a downhill gradient, as at B? In a later paper, Boltzmann offers a remarkable proposal to explain this, too. Perhaps our perception of past and future depends on the entropy gradient, in such a way that we are bound to regard the future as lying ‘uphill’ on the entropy gradient.

In the universe ... one can find, here and there, relatively small regions on the scale of our stellar region ... that during the relatively short eons are far from equilibrium. What is more, there will be as many of these in which the probability of the state is increasing as decreasing. Thus, for the universe the two directions of time are indistinguishable, just as in space there is no up or down. But just as we, at a certain point on the surface of the Earth, regard the direction to the centre of the Earth as down, a living creature that at a certain time is present in one of these isolated worlds will regard the direction of time towards

the more improbable state as different from the opposite direction (calling the former the past, or beginning, and the latter the future, or end). Therefore, in these small regions that become isolated from the universe the ‘beginning’ will always be in an improbable state. (Boltzmann 1897, translation from Barbour 1999, p. 342)

This is perhaps the first time that anyone had challenged the objectivity of the perceived direction of time, and this alone makes Boltzmann’s hypothesis a brilliant and revolutionary idea. But the proposal also solves PTB in a beautiful way. It explains the apparent asymmetry of thermodynamics in terms of a cosmological hypothesis which is symmetric on the larger scale. So PTB simply goes away on the large scale—though without depriving us of an explanation for why we do find temporal bias locally. Boltzmann’s hypothesis is the kind of idea which simply deserves to be true!

5. The monkey wrench

Unfortunately it isn’t true, or at least there’s a huge spanner in the works. The problem stems directly from Boltzmann’s own link between entropy and probability. According to Boltzmann’s famous formula, $S = k \log W$, entropy is proportional to the logarithm of probability (the latter judged by the Boltzmann measure). In Figure 1, then, the vertical axis is a logarithmic probability scale. For every downward increment, dips in the curve of the corresponding depth become exponentially more improbable. So a dip of the depth of point A or point B is far more likely to occur in the form shown at point C—where the given depth is very close to the minimum of the fluctuation—than it is to occur in association with a much bigger dip, such as that associated with A and B. This implies that if we wish to accept that our own region is the product of ‘natural’ evolution from a state of even lower entropy, we must accept that our region is far more improbable than it needs to be, given its present entropy.

To put this another way, if Boltzmann’s probabilities are our guide, then it is much easier to produce fake records and memories, than to produce the real events of which they purport to be records. Imagine, for example, that God chooses to fast-forward through a typical world-history, until he finds the complete works of

Shakespeare, in all their contemporary twenty-first century editions. It is *vastly* more likely that he will hit upon a world in which the texts occur as a spontaneous fluctuation of modern molecules, than that he'll find them produced by the Bard himself. (Editions typed by monkeys are probably somewhere in between, if the monkeys themselves are products of recent fluctuations.) In Boltzmann's terms, then, it is unlikely that Shakespeare existed, four hundred years ago. Someone like him exists somewhere in Boltzmann's universe, but he's very unlikely to be in *our* recent past. The same goes for the rest of what we take to be history. All our 'records' and 'memories' are almost certainly misleading.

There's another problem of a similar kind. Just as we should not expect the low entropy region to extend further back *in time* than it needs to in order to produce what we see, so we should not expect it to be any more extensive *in space* than we already know it to be. (Analogy: shuffle a deck of cards, and deal a hand of thirteen cards. The fact that the first six cards you turn over are spades does not give you reason to think that the rest of the hand are spades.) But we now observe vastly more of the universe than was possible in Boltzmann's day, and yet the order still extends as far as we can see.

Brilliant as it is, then, Boltzmann's hypothesis faces some devastating objections. Moreover, modern cosmology goes at least some way to providing us with an alternative. As I'll explain later, this may not mean that the hypothesis is completely dead—it might enjoy new life as part of an explanation of what modern cosmology tells us about the low entropy past. But for the moment, our focus needs to be on that cosmological story.

6. Initial smoothness

What boundary conditions, at what times, are needed to account for the time-asymmetry of observed thermodynamic phenomena? There seems no reason to expect a neat answer to this question. Low entropy just requires concentrations of energy, in useable forms. There are countless ways such stores of useable energy could exist, at some point in our past. Remarkably, however, it seems that a single simply-characterizable condition does

the trick. All the observed order seems attributable to a single characteristic of the universe soon after the big bang.

The crucial thing is that matter in the universe is distributed extremely smoothly, about one hundred thousand years after the big bang. It may seem puzzling that this should be a low entropy state. Isn't a homogeneous, widely dispersed arrangement of matter a disordered, high entropy arrangement? But it all depends on what forces are in charge. A normal gas tends to spread out, but that's because the dominant force—pressure—is repulsive. In a system dominated by an attractive force, such as gravity, a uniform distribution of matter is highly unstable. The natural behaviour of such matter is to clump together. Think of the behaviour of water on the surface of a waxy leaf, where the dominant force is surface tension—or of a huge collection of sticky polystyrene foam pellets, whose natural tendency is to stick together in large clusters.

To get a sense of how extraordinary it is that matter should be distributed uniformly near the big bang, keep in mind that we've found no reason to disagree with Boltzmann's suggestion that there's no objective distinction between past and future—no sense in which things 'really' happen in the direction we think of as past-to-future. Without any such objective distinction, we're equally entitled to regard the big bang as the end point of a gravitational collapse. For such a collapse to produce a very smooth distribution of matter is, to put it mildly, quite extraordinary, judged by our ordinary view about how gravitating matter should behave. (Imagine throwing trillions of sticky foam pellets into a tornado, and having them settle in a perfect sheet, one pellet thick, over every square centimetre of Kansas—that's an easy trick, by comparison!)

I stress that there are two very remarkable things about this feature of the early universe. One is that it happens at all, given that it is so staggeringly unlikely, in terms of our existing theory of how gravitating matter behaves. (Penrose 1989, ch. 7, estimates the unlikeliness of such a smooth arrangement of matter at 1 in $10^{10^{123}}$.) The other is that so far as we know, it is the *only* anomaly necessary to account for the vast range of low entropy systems we find in the universe. In effect, the smooth distribution of matter in the early universe provides a vast reservoir of low entropy, on which everything else

depends. The most important mechanism is the formation of stars and galaxies. Smoothness is necessary for galaxy and star formation, and most irreversible phenomena with which we are familiar owe their existence to the sun. (For more details see Penrose 1989, ch. 7.)

In my view, this discovery about the cosmological origins of low entropy is the most important achievement of late twentieth century physics. It is true that in one sense it simply moves the puzzle of temporal bias from one place to another. We now want to know *why* the early universe is so smooth. But as I've emphasised, it's an extraordinary fact that the puzzle turns out to be capable of being focussed in that place.

The puzzle of initial smoothness thus gives concrete form to the explanatory project which begins with the time-asymmetry of thermodynamics. If cosmology could explain initial smoothness, the project would be substantially complete, and PTB would be substantially solved. At the moment, however, it is very unclear what form a satisfactory explanation might take. I'll say a little more below about some of the possibilities.

However, my main task in the remainder of this chapter is to defend the claim that initial smoothness needs explaining. Some philosophers argue that it is inappropriate to ask for an explanation of such an initial condition. It would be nice if this were true, for it would imply that PTB has in large part been laid to rest—that most of the work that needs to be done, has been done. But I think these philosophers are mistaken, and hence that there is still a lot of work for cosmologists to do on this issue. Our role as philosophers is to help them to see the importance of the issue.

7. Should cosmologists be trying to explain initial smoothness?

In the light of late twentieth century cosmology, then, the late nineteenth century puzzle of low entropy takes a new concrete form. Why is the universe smooth, soon after the big bang? Should we be looking for an answer to this question? Some philosophers say not. For example, Craig Callender suggests that 'the whole enterprise of explaining global boundary conditions is suspect, for precisely the reasons Hume and Kant taught us, namely, that we can't obtain causal or probabilistic explanations of why the boundary

conditions are what they are.’ (Callender 1997, p. 69. See also Callender’s companion chapter in this volume, Callender 1998, pp. 149–50, and Sklar 1993, pp. 311–312, for similar concerns.)

There are a number of ways we might respond to this objection. We might argue that Hume and Kant are simply wrong. We might argue that there’s some different kind of explanation it is appropriate to seek for the smooth early universe, an explanation neither causal nor probabilistic. Or we might argue that the objection misses its target, because the smooth early universe isn’t a boundary condition in the relevant sense; but rather something else, something which does call for explanation. My strategy will be predominantly the third of these options, though with some elements of the first and second.

I’ll proceed as follows. First, I’ll appeal to your intuitions. I’ll ask you to imagine a discovery that cosmology might have made about the universe, a case in which it seems intuitively clear that we would seek further explanation. I’ll then argue that we have no grounds for taking this imaginary case to be different from the actual case. (On the contrary, I claim, it is the actual case, but described in a non-standard way.) Next, I’ll clarify the status of the low entropy ‘boundary condition’, and in particular, call attention to a sense in which its status seems necessarily to be more than that of a mere boundary condition. It is ‘lawlike’ rather than ‘factlike’ in nature, in a sense I’ll make more precise. (I’ll argue that unless it has this status, we have no defence against the kind of sceptical challenge which proved so devastating to Boltzmann’s proposed explanation of the low entropy past.) Finally, I’ll respond briefly to Callender’s elucidation of the Humean objection to the project of explaining initial smoothness, in the light of this clarification of its status.

First a note on terminology. In recent years, a number of writers have taken to calling the supposition that the early universe has low entropy the ‘Past Hypothesis’. This phrase was introduced by David Albert, who takes it from a passage in Richard Feynman’s *The Character of Physical Law* in which Feynman says ‘I think it necessary to add to the physical laws the hypothesis that in the past the universe was

more ordered ... than it is today.’ (Feynman 1967, p. 116) In this formulation, however, there is no special mention of cosmology. But Albert’s most explicit formulation of what he means by the term refers explicitly to cosmology. According to Albert, the Past Hypothesis ‘is that the world first came into being in whatever particular low-entropy highly condensed big-bang sort of macrocondition it is that the normal inferential procedures of cosmology will eventually present to us.’ (Albert 2000, p. 96).

This terminological point is important. Taken in Feynman’s original non-specific form, the Past Hypothesis is certainly capable of further explanation—a fit topic for cosmological investigation (as Feynman (1967, p. 116) himself notes, saying that although this hypothesis is now “considered to be astronomical history”, “perhaps someday it will also be a part of physical law.”). And after all, this is precisely what’s happened. The abstractly characterised fact has now been explained by the smoothness of the early universe, and the issue is simply whether cosmologists should be trying to take things a stage further. Nothing in Feynman’s proposal suggests that they should not.

Taken in Albert’s form, however, the Past Hypothesis is by definition the final deliverance of cosmology on the matter. While it is then analytic that the Past Hypothesis itself will not be further explained, we have no way of knowing whether current cosmology is final cosmology—experience certainly suggests not! So the trivial semantic fact that the Past Hypothesis (so defined) cannot be further explained provides no reason not to try to explain the smooth early universe. Even if it that were the Past Hypothesis, we wouldn’t find that out, presumably, until we’d tried to explain it further, and become convinced by persistent failure that it was the final theory. I think that similar remarks apply to the term ‘Past State’, which Callender’s uses in Chapter 12—it, too, can be read in either of two ways. So to side-step these terminological confusions, I’ll to avoid using these terms, and concentrate on what we actually have from cosmology, viz., the smooth early universe. Is this something we should be trying to explain?

8. What's special about *initial conditions*?

Part of the usual resistance to the idea of explaining initial conditions is associated with the thought that we normally explain events in terms of *earlier* events. By definition, there is nothing earlier than the initial conditions.

In the present context, however, this preference is on shaky ground. Here's a way to make this vivid. Imagine that in recent years physics had discovered that the matter in the universe is collapsing towards a big crunch, fifteen billion years or so in our future—and that as it does so, something very peculiar is happening. The motions of the individual pieces of matter in the universe are somehow conspiring to defeat gravity's overwhelming tendency to pull things together. Somehow, by some extraordinary feat of cooperation, the various forces are balancing out, so that by the time of the big crunch, matter will have spread itself out with great uniformity. A molecule out of place, and the whole house of cards would surely collapse! Why? Because as Albert (2000, 151), puts it, 'the property of being an *abnormal* [i.e., entropy-reducing] microstate is extraordinarily *unstable* under small perturbations.' By the lights of the Boltzmann measure, then, the tiniest disturbance to our imagined entropy-reducing universe would be expected to yield an entropy-increasing universe.

As a combination of significance and sheer improbability—the latter judged by well-grounded conceptions of how matter is expected to behave—this discovery would surely trump anything else ever discovered by physics. Should physicists sit on their hands, and not even try to explain it? (They might fail, of course, but that's always on the cards—the issue is whether it is appropriate to try.) If this discovery didn't call for explanation, what conceivable discovery ever would?

In my view, however, this state of affairs is *exactly* what physics has discovered! I've merely taken advantage, once again, of the fact that if there is no objective sense in which what we call the future is *really* the 'positive' direction of time, then we can equally well describe the world by reversing the usual temporal labelling. Relabelled in this way, the familiar expansion from a smooth big bang becomes a contraction to a smooth big crunch, with the extraordinary characteristics just described. And surely if it

is a proper matter for explanation described one way, it is a proper matter for explanation described the other?

Both steps in this argument could conceivably be challenged. The first relies, as I said, on the view that there is no objective distinction between past and future, no difference between our world and a world in which exactly the same things happen, but in the opposite order. This claim is contentious. One prominent writer who rejects it is John Earman. In a classic (1974) paper on the direction of time, Earman suggests—correctly, in my view—that someone who endorses this view about time would have no grounds to reject an analogous view about spatial parity; and would thus be committed to the view that there is no objective difference between a possible world and its mirror-reversed twin. I agree, and to me, this seems the right view in that case, too.

It would take us too far afield to try to settle this issue here. For the moment, the important thing is that someone who wants to say that my imagined physical discovery is different from the actual discovery made by cosmology, and that this accounts for the fact that it would call for explanation in a way which the smooth early universe does not, faces an uphill battle. First, they owe us an account of the objective difference between past and future. Second, they need to explain how this difference makes a difference to what needs explaining. And third, they need to explain how they know they've got things the right way round—how they know that we live in the world where the smooth extremity does not need explaining, rather than the temporal mirror world, in which it does.

Absent such arguments, I take the lesson of this example to be as follows. Our ordinary intuitions about what needs explaining involve a strong temporal bias, a temporal bias we should eliminate, if we want our physical explanations to show reasonable invariance under trivial redescription of the phenomena in question. In particular, our tendency simply to take initial conditions for granted is unreliable, because the same conditions can equally well be regarded as final conditions.

It might be objected that this doesn't necessarily show that the smooth early universe calls for explanation. Perhaps the argument actually cuts the other way,

showing that our intuitions about the redescribed case—the smooth ‘late’ universe—are unreliable. Perhaps we would be wrong to try to explain a smooth big crunch.

(Callender’s companion chapter suggests that he would take this view.) For my part, I find it hard to make sense of this possibility. As I said, if the imagined discovery did not strike us as calling for explanation (in the light of our pre-existing expectations about how gravitating matter ought to behave), then it is hard to see what discovery ever would call for explanation. However, it would be nice to do better than simply trading intuitions on this point, and for this, I think, we need some additional guidelines. In particular, we need to pay closer attention to the theoretical role of the ‘boundary condition’ in question. I’ll approach this issue by considering another objection to the project of explaining the low entropy past.

9. The just good luck objection

The objection in question is close to one expressed by D. H. Mellor, in a recent response to John Leslie. Mellor describes the following example from Leslie:

Suppose you are facing a firing squad. Fifty marksmen take aim, but they all miss. If they hadn't all missed, you wouldn't have survived to ponder the matter. But you wouldn't leave it at that: you'd still be baffled, and you'd seek some further reason for your luck. (2002, p. 227)

Mellor then writes,

Well, maybe you would; but only because you thought the ability of the firing squad, the accuracy of their weapons, and their intention to kill you made their firing together a mechanism that gave your death a very high physical probability. So now suppose there is no such mechanism. Imagine ... that our universe (including all our memories and other present traces of the past) started five minutes ago, with these fifty bullets coming past you, but with no prior mechanism to give their trajectories any physical probability, high or low. Suppose in other words that these trajectories really were among the *initial* conditions of our universe. If you thought that, should you really be baffled and seek some further reason for your luck? (2002, p. 227).

It might be argued—and Callender’s companion chapter suggests that he would be sympathetic to this idea—that the smooth early universe is like this imagined case, in requiring no mechanism to bring it about. Isn’t the smooth early universe just a matter of luck, like the trajectories of the bullets, in Mellor’s example?

But let’s think some more about Mellor’s example. Let’s imagine ourselves in Mellor’s world, being told that another 50, or 500, or 5000 bullets are yet to arrive. Should we expect our luck to continue? In my view, to think it’s an accident that the first 50 bullets missed us just *is* to have no expectation that the pattern will continue in new cases. Perhaps we think something else about new cases, or perhaps we’re simply agnostic, but either way, we don’t ‘project’ from the initial 50 cases. If the pattern does continue, say for another 500 cases, we might go on attributing it purely to luck. But we can’t both expect it to continue indefinitely, and attribute that in advance merely to luck. For to take the generalisation to be projectible *is* to treat it as something more than merely an accident—as something *lawlike*.

Similarly, if we think that the smooth early universe is just a matter of luck, then we have no reason to expect that the luck will continue, when we encounter new regions of the universe—regions previously too far away to see, for example. Again, perhaps we’ll think it won’t continue, or perhaps we’ll be agnostic. But either way, we won’t think that it will continue.

This argument is very similar to one version of the objection we encountered in §5 to the Boltzmann hypothesis. There, the spatial version of the objection was that if the low entropy past is just a statistical fluctuation, we shouldn’t expect more of it than we’ve already discovered—we shouldn’t expect to see more order, as we look further out into space. Similarly in the present case: if the smooth early universe is just a piece of luck, we shouldn’t expect our luck to continue.

As actually used in contemporary cosmology, the hypothesis of the smooth early universe is not like this. It is taken to be projectible, in the sense that everyone expects it and its consequences (the existence of galaxies, for example) to continue to hold, as we

look further and further out into space. The hypothesis is thus being accorded a lawlike status, rather than treated as something that ‘just happens’.

This argument was analogous to the spatial version of the objection to the Boltzmann hypothesis. The more striking temporal version of that objection also carries over to the present case, I think. For suppose we did think of the smooth early universe as a lucky accident. The essence of the temporal objection to Boltzmann was that there are many lucky accidents compatible with what we see—almost all of them far more likely than the smooth big bang, in terms of the Boltzmann measure. So why should we think that the actual accident was a smooth early universe, rather than one of those other possibilities? The upshot is that the belief that the smooth big bang is a lucky accident seems (all but) incompatible with the belief that it actually happened!

In my view, the present state of play is this. Modern cosmology is implicitly committed to the view that the smooth big bang is not merely a lucky accident. But we don’t yet understand how this can be the case. This puzzle is the twentieth century’s legacy to the twenty-first—its transformation of the original nineteenth century puzzle of temporal bias. It is not an exaggeration, in my view, to say that this is one of *the* great puzzles of contemporary physics (even if a puzzle whose importance is easily underrated, for the reasons we’ve already canvassed). At this point, philosophers should not be encouraging physicists to rest on their laurels (or laureates). On the contrary, we should be helping them to see the full significance of this new puzzle, and encouraging them to get to work on it!

In what directions should they be looking? I’ll say a little about this issue in a moment, but before that, I want to respond briefly to another aspect of the challenge from Hume, Kant, and Callender to the project of explaining initial conditions.

10. The only one universe objection

In his companion chapter, Callender cites Hume’s famous objection to the project of explaining ‘the generation of a universe’, as Hume puts it. As Callender says, Hume’s point ‘is that since the cosmos happens only once, we cannot hope to gain knowledge of any regularities in how it is created.’

I offer three responses to this objection. First, the required boundary condition is not necessarily unique, because the universe may contain other relevantly similar singularities. Second, even if it were unique, there is an important and familiar sense in which its components provide generality. And third, explanations of unique states of affairs are perfectly normal—in fact, unavoidable—at least in the case of laws.

Not necessarily unique

It is far from clear that the required boundary condition is unique. The expanding universe may eventually recollapse again, in which case there will be a big crunch in our future, as well as a big bang in our past. It is true that the trend of recent astronomical evidence has been against this possibility, but we are here canvassing possibilities, and should certainly leave this one on the table. In any case, even if whole universe doesn't recollapse, it is thought that parts of it will, as large accumulations of matter form black holes. As writers such as Hawking (1985, p. 2491) and Penrose (1979, pp. 597–8) have pointed out in this context, this process is very much like a miniature version of collapse of the entire universe. In some respects, then, the big bang is one of a general class of events, of which the universe may contain many examples. The big bang may have special significance, but it far from clear that its properties could not be derivable from some general theory of singularities, a theory testable in principle by observation of multiple instances of the phenomena it describes.

Even one case provides generality

If matter in the universe as a whole is smoothly distributed after the big bang, this implies that the following is true of the matter in every individual region of the universe. As we follow the matter in that region backwards in time, towards the big bang, we find irregular accumulations of matter disappearing. Somehow, the particular chunk of the matter in the region in question manages to spread itself out—interacting with other chunks as it does so, but not presumably with all other chunks, since some of them are too far away. As aliens from another dimension, as it were, we could select chunks at

random, and discover that this same behaviour was characteristic of all of them.

Wouldn't this count as a generalisation, if anything does?

Here's an analogy. As I mentioned a moment ago, recent observational evidence suggests that the expansion of the universe is actually accelerating. We don't think it inappropriate to seek to explain why this should be so. On the contrary, this is widely regarded as a fascinating project, likely to require new physical theories. But in one sense, this expansion too is just one unique case. There's just one universe, and it has to behave some way, so why not this way? Again, part of a proper response to this challenge seems to be that we find the same thing happening in many parts of the universe, and that this suggests some unifying underlying explanation. That seems to me to be precisely what we find in the case of the smooth early universe, too.

Even unique things get explained

The laws of nature are unique, in the sense that there is only one world of which they are the laws. Yet we often think it proper to explain laws, by showing that they follow from more fundamental laws. It isn't always clear where this is appropriate or needed, but there's certainly a good deal of consensus on these things. I've argued above that the low entropy early universe needs to be regarded as a lawlike hypothesis, if we are to avoid objections analogous to those that afflict the Boltzmann hypothesis. It seems a reasonable project to seek some deeper understanding of this hypothesis—to hope to show how it follows from something more fundamental. (Again, we may fail, but the question is whether we should try.)

11. What might explanations look like?

I want to finish by mentioning some strategies for seeking to explain the smooth early universe. I don't think any of these strategies is unproblematic as it stands, but they do give some sense of both the options and the problems facing this important theoretical task.

The appeal to inflation

The first approach stems from what cosmologists call the inflationary model. This model is a kind of front end to the standard big bang model, describing what might have happened to the universe in its extremely early stages. The proposal is that when the universe is extremely small—perhaps simply the product of some quantum fluctuation—the physical forces in play are different from those with which we are familiar. In particular, gravity is repulsive, rather than attractive, and the effect is that the universe experiences a period of exponential expansion. As it grows it cools, and at a certain point undergoes a ‘phase transition’. The forces change, gravity becomes attractive, and the universe settles into the more sedate expansion of the ‘ordinary’ big bang. (For an introduction, see Linde 1987.)

Since first proposed in the 1980s, one of the main attractions of the inflationary model has been that it seems to explain a features of the early universe which the standard big bang model simply has to take for granted. One of these features, it is claimed, is the smoothness of the universe after the big bang. However, the argument the inflation explains smoothness is essentially statistical. The crucial idea is that during the inflationary phase the repulsive gravity in will tend to ‘iron out’ inhomogeneities, leaving a smooth universe at the time of the transition to the classical big bang. Presenting the argument in *Nature* in 1983, Paul Davies concludes that

the Universe ... began in an arbitrary, rather than remarkably specific, state. This is precisely what one would expect if the Universe is to be explained as a spontaneous random quantum fluctuation from nothing. (1983, p. 398)

But this argument illustrates the temporal double standard that commonly appears in discussions of these problems. After all, we know that we might equally well view the problem in reverse, as a gravitational collapse towards a big crunch. In statistical terms, this collapse may be expected to produce *inhomogeneities* at the time of any transition to an inflationary phase. Unless one temporal direction is already privileged, the statistical reasoning is as good in one direction as the other. Hence in the absence of a justification

for the double standard—a reason to apply the statistical argument in one direction rather than the other—the appeal to inflation doesn't seem to do the work required of it.

Davies misses this point. Indeed, he also argues that

a recontracting Universe arriving at the big crunch would not undergo 'deflation,' for this would require an exceedingly improbable conspiracy of quantum coherence to reverse-tunnel through the phase transition. There is thus a distinct and fundamental asymmetry between the beginning and the end of a recontracting Universe. (1983, p. 399)

But as Page (1983) points out, this conflicts with the argument he has given us concerning the other end of the universe. Viewed in reverse, the transition from the ordinary big bang to the inflationary phase involves exactly this kind of 'improbable conspiracy.' If deflation is unlikely at one end, then inflation is unlikely at the other.

For these reasons, amongst others, it is far from clear that the inflationary approach works as it stands to explain the smooth early universe. Nevertheless, it illustrates a possible strategy for doing so—an approach that involves making early smoothness probable, by showing that under plausible constraints, all or most possible universes compatible with those constraints have the feature in question.

The anthropic strategy

Perhaps the reason that the universe looks so unusual to us is simply that we can only exist in very unusual bits of it. We depend on the entropy gradient, and could not survive in a region in thermodynamic equilibrium. Could this explain why we find ourselves in a low entropy region?

This is the anthropic approach, already encountered in the form of the Boltzmann hypothesis. As in that case, the idea is interesting, but faces severe difficulties. For one thing, it depends on there being a genuine multiplicity of actual bits of a much larger universe, of which our bit is simply some small corner. It is no use relying on other merely *possible* worlds, since that would leave us without an explanation for why ours turned out to be the real world. (If it hadn't turn out this way, we wouldn't have been

around to think about it, but this doesn't explain why it did turn out this way.) So the anthropic solution is very costly in ontological terms. It requires that there be vastly more 'out there' than we otherwise expect.

All the same, this would not be a disadvantage if the cost was one we were committed to bearing anyway. It turns out that according to some versions of the inflationary model, universes in the normal sense are just bubbles in some vast foam of universes. So there might be independent reason to believe that reality is vastly more inclusive than it seems. In this case, the anthropic view does not necessarily make things any worse.

The second difficulty is that as Penrose (1979, p. 634) emphasizes, there may well be much cheaper ways to generate a sufficient entropy gradient to support life. The observed universe seems vastly more unlikely than intelligent life requires. Again, this is close to an objection to Boltzmann's view. We noted that Boltzmann's suggestion implies that at any given stage, we should not expect to find more order than we have previously observed. The same seems to apply to the contemporary argument. Life as we know it doesn't seem to require an early universe which is smooth everywhere, but only one which is smooth in a sufficiently large area to allow a galaxy or two to form (and to remain relatively undisturbed while intelligent life evolves). This would be much cheaper in entropy terms than global smoothness.

However, the inflationary model might leave a loophole here, too. If the inflationary theory could show that a smooth universe of the size of ours is an all or nothing matter, then the anthropic argument would be back on track. The quantum preconditions for inflation might be extremely rare, but this would not matter, so long as there is enough time in some background grand universe for them to be likely to occur eventually, and it is guaranteed that when they do occur a universe of our sort arises, complete with its smooth boundary.

So the anthropic strategy cannot be excluded altogether. It depends heavily on the right sort of assistance from cosmological theory, but if that were forthcoming, this approach might explain why we find ourselves in a universe with a low entropy history.

If so, however, then there is hugely more to reality than we currently imagine, and the concerns of contemporary astronomy pale into insignificance in comparison.

Penrose's Weyl hypothesis

The writer who has done most to call attention to the importance and specialness of the smooth big bang is Roger Penrose. Penrose himself proposes that there must be an additional law of nature, to the effect that the initial extremities of the universe obey what amounts to a smoothness constraint. In technical terms, his hypothesis is that the so-called Weyl curvature of spacetime approaches zero in this initial region. (See Penrose 1979, and particularly Penrose 1989, ch. 7.)

We might object to the use of the term 'initial' here. Isn't Penrose presupposing an objective distinction between past and future? But the difficulty is superficial. Penrose's claim need only be that it is a physical law that there is one temporal direction in which the Weyl curvature always approaches zero towards the universe's extremities. The fact that conscious observers inevitably regard that direction as the past will then follow from the sort of argument already made by Boltzmann.

Another objection might be that Penrose's proposal does little more than simply redescribe the smooth early universe. There is some justice in this comment, but the proposal might nevertheless constitute theoretical progress. By characterising what needs to be explained in terms of the Weyl curvature, it might provide the right focus for further and deeper theoretical explanation, say from quantum cosmology.

One important issue is whether, as Penrose thinks, his proposal needs to be time-asymmetric, or whether the Weyl curvature might approach zero towards the extremities of the universe in both directions. This alternative would do just as well at explaining the smoothness of the big bang, and have the advantage of not introducing a new time-asymmetry into physics. However, it would imply that entropy would eventually decrease if the universe recontracts in the distant future. This is an unpopular view, but it turns out that most of the arguments that cosmologists give for rejecting it are rather weak. They amount to pointing out that such an outcome would be unlikely—as indeed it would be, by ordinary Boltzmann lights. But as in the case of the past low entropy

condition, the whole point of the extra condition would be that the Boltzmann measure is not the last word, once cosmological factors are taken into account. So there seems little reason to prefer Penrose's asymmetric hypothesis to the symmetric version of the same thing. (One interesting question is whether a future low entropy condition would have observable present effects. For more this and related issues, see Price 1996, ch. 4.)

12. Conclusion

The above examples give some sense of how physics might come to regard the smooth early universe as a consequence of something more basic. It is true that the project remains rather vague, but isn't this what we should expect? Looking back into the history of modern physics, we can see that the recent discovery of the smooth early universe represents a huge advance in our understanding of a puzzle which was only coming dimly into view a century ago. The size of the advance ought to remind us how hard it is to look forwards, and predict the course of future physics. True, we know that in some way or other, future physics will incorporate much of current physics (much of which is surely right, so far as it goes). But we almost nothing about the form the incorporation will take, or the nature and extent of the novelty—the new framework, within which the incorporation of the old will take place.

Concerning the smooth early universe itself, we can be reasonably confident that it, or something like it, will remain a part of future physics—and an important part, given its centrality to explanation of something so crucial as the temporal bias of thermodynamics to the nature of the world in which we find ourselves. But whether it will remain fundamental in its own right, something not further explained elsewhere in our new theories, we simply don't know. The best we can do is to trust our intuitions, and see what we can find.

I've argued that these intuitions benefit from some philosophical therapy, to prevent us from taking too seriously some old concerns about the project of explaining initial conditions. With the benefit of such therapy, I think that most physicists' intuitions, like mine, will be that there is an important explanatory project here, that there

is likely to be something interesting to find. As to whether we're right, of course, only time itself will tell—but only if we try!

HUW PRICE

REFERENCES

- Albert, D. (1994). The foundations of quantum mechanics and the approach to thermodynamic equilibrium. *British Journal for the Philosophy of Science*, 45, 669–677.
- (2000). *Time and chance*. Cambridge, Mass.: Harvard University Press.
- Barbour, J. (1999). *The end of time*. London: Weidenfeld and Nicholson.
- Boltzmann, L. (1877). Über die Beziehung zwischen des zweiten Hauptsatze der mechanischen der Wärmetheorie. *Sitzungsberichte, K. Akademie der Wissenschaften in Wien, Math.-Naturwiss*, 75, 67–73. Reprinted in translation in Brush (1966).
- (1895). On certain questions of the theory of gases. *Nature*, 51, 413–15.
- (1897). Zu Hrñ Zermelo's Abhandlung "Über die mechanische Erklärung irreversibler Vorgänge" *Wied. Ann.* 60, 392-398.
- Brush, S. (1966). *Kinetic theory, vol. 2: irreversible processes*. Oxford: Pergamon Press.
- Burbury, S. H. (1894). Boltzmann's minimum function, *Nature*, 51, 78.
- (1895). Boltzmann's minimum function, *Nature*, 51, 320.
- Callender, C. (1997). Review of H. Price, *Time's Arrow and Archimedes' Point*, *Metascience*, 11, 68-71.
- (1998). The view from no-when, *British Journal for the Philosophy of Science*, 49, 135–59.
- Davies, P. (1983). Inflation and Time Asymmetry in the Universe. *Nature*, 301, 398–400.

- Earman, J. (1974). An attempt to add a little direction to the problem of the direction of time. *Philosophy of Science*, 41, 15-47.
- Feynman, R. (1967). *The character of physical law*. Cambridge, Mass.: MIT Press.
- Hall, E. H. (1899). Review of S. H. Burbury, *The Kinetic Theory of Gases* (Cambridge: Cambridge University Press, 1899). *Science, New Series*, 10, 685-688.
- Hawking, S. W. (1985). Arrow of time in cosmology. *Physical Review, D*, 33, 2489–2495.
- Mellor, D. H. (2002). Too many universes. In N. A. Manson (Ed). *God and design: the teleological argument and modern science* (pp. 221-228). London: Routledge.
- Linde, A. (1987). Inflation and quantum cosmology. In S. W. Hawking, and W. Israel (Eds). *Three hundred years of gravitation* (pp. 604–630). Cambridge: Cambridge University Press,
- Page, D. (1983). Inflation does not explain time asymmetry. *Nature*, 304, 39–41.
- Penrose, R. (1979). Singularities and time-asymmetry. In S. W. Hawking, and W. Israel (Eds). *General relativity: an Einstein centenary* (pp. 581–638). Cambridge: Cambridge University Press.
- (1989). *The Emperor's New Mind*, Oxford: Oxford University Press.
- Price, H. (1996). *Time's arrow and Archimedes' point: new directions for the physics of time*, New York: Oxford University Press.
- (2002a). Boltzmann's time bomb. *British Journal for the Philosophy of Science*, 53, 83-119.
- (2002b). Burbury's last case: the mystery of the entropic arrow. In C. Callender (Ed). *Time, reality and experience* (pp. 19-56). Cambridge: Cambridge University Press.
- Sklar, L. (1993). *Physics and chance: philosophical issues in the foundations of statistical mechanics*. Cambridge: Cambridge University Press.