What's so special about initial conditions?

Understanding the past hypothesis in directionless time

[To appear in Y. Ben-Menahem (ed), *Rethinking Laws of Nature*, Springer.][*]

Matt Farr

University of Cambridge

you@mattfarr.co.uk • mattfarr.co.uk

November 16, 2021

Abstract

It is often said that the world is explained by laws of nature together with initial conditions. But does that mean initial conditions don't require further explanation? And does the explanatory role played by initial conditions entail or require that time has a preferred direction? This chapter looks at the use of the 'initialness defence' in physics, the idea that initial conditions are intrinsically special in that they don't require further explanation, unlike the state of the world at other times. Such defences commonly assume a primitive directionality of time to distinguish between initial and final conditions. Using the case study of the time-asymmetry of thermodynamics and the so-called 'past hypothesis' — the hypothesis that the early universe was in a state of very low entropy —, I outline and support a deflationary account of the initialness defence that does not presuppose a basic directionality of time, and argue that there is a relevant explanatory asymmetry between initial conditions and the state of systems at other times only if certain causal conditions are satisfied. Hence, the initialness defence is available to those who reject a fundamental direction of time.

[*] Comments welcome (especially if they concern unfortunate typos or project-undermining objections).

Contents

1 Introduction

Eugene Wigner once remarked that '[t]he sharp distinction between Initial Conditions and Laws of Nature was initiated by Isaac Newton and I consider this to be one of his most important, if not *the* most important, accomplishment', adding that prior to Newton, 'there was no sharp separation between the two concepts' and that after Newton this sharp separation 'was taken for granted and rarely even mentioned' (Wigner, 1995, p. 334). The terminology of 'initial' conditions is so widespread within physics as to make its usage seemingly unremarkable. But the very word 'initial' connotes the idea of time as directed, deriving from the Latin for 'the beginning'; something could not be 'initial' without being chronologically first in some sequence of events, and something could not be chronologically first without a direction of time relative to which it is first. However, there are many motivations within fundamental physics for holding time to lack a direction,[1] with the appearance of a direction of time being an emergent feature of higher-level physics. So what is the relationship between initial conditions and the direction of time, and how do philosophical views that reject a directionality of time make sense of the role of initial conditions in physics?

This chapter looks at a particular kind of argument I'll call the 'initialness defence'. The general idea is that initial conditions, as opposed to the state of the system at other times, are explanatorily special in virtue of temporally preceding all other temporal stages of the relevant system, and not being the result of evolution from earlier states. In ordinary experiments, one can consider the initial conditions — the state in which the relevant system is prepared in the experimental set-up — to be a set of information that one does not seek to explain in terms of the theory being tested. Conversely, the final conditions of the experiment, namely the measured state at the end of the experiment, is to be explained in terms of the relevant theory together with the initial conditions. This explanatory asymmetry between initial conditions and final conditions is present in cosmology, such as the cosmological explanation of the second law of thermodynamics, the details of which I'll turn to in a moment. The idea in the cosmological case is that one can design a cosmological model with some assumed initial conditions, such as the extremely low-entropy early universe, and appeal to this to explain some particular observed regularity, in our case the steady increase of thermodynamical entropy towards the future. In such cases, the initial conditions are ordinarily treated as explanatorily special.

Initialness defence. The earliest state of some system is explanatorily special because it is the only state that did not evolve from some earlier state of that system.

Initialness defences commonly assume a primitive directionality of time in order

[1] See Farr (2020).

justify distinguishing between initial and final conditions in this way. But what exactly is the relationship between the initialness defence and the direction of time? Do initialness defences presuppose a specific metaphysical view of time direction? And can metaphysical views that hold time to be *adirectional* — to lack a preferred direction — employ initialness defences? In what follows, I assess these issues using the case study of the time asymmetry of thermodynamics.

The standard statistical mechanical account of why macroscopic systems obey the second law of thermodynamics relies on the assumption that in the past, entropy was far lower than it is now. Contemporary cosmology posits an extremely low entropy and thus low probability initial state of the universe, which following Albert (2000) is commonly termed the *past hypothesis*. There has been much disagreement as to whether the extremely low-probability initial state posited by the past hypothesis stands in need of explanation (e.g. Price (2004) argues it does; Callender (2004b) argues it does not), and what would constitute an explanation of such a thing (e.g. Callender (2004a) questions whether there could in principle be such an explanation). One thought implicit in a number of considerations of this issue is a version of the initialness defence, that the past hypothesis is explanatorily special due to being a hypothesis about the *earliest* temporal part of the universe, and so not causally dependent upon earlier states. Prima facie, realists about time direction — those who take the directionality of time to be some basic fact about the world, such as Maudlin (2007) — can make use of the initialness defence, but antirealists about time direction — those who hold either that time has no direction, or that the directionality of time is an emergent feature of the macroscopic world, such as Price (1996) — cannot, since on their picture of the world there is nothing intrinsically special about either temporal 'end' of the universe. I argue this dichotomy is false. Instead, antirealists about time direction, or 'C-theorists' as I will call them, can and should adopt a restricted version of the initialness defence.

The chapter is structured as follows. Section 2 introduces the background to the debate, introducing the problem of time asymmetry in statistical mechanics and the role played by the past hypothesis. Section 3 introduces the B- and C-theories of time, and the debate between realist and antirealist accounts of the direction of time. In section 4 I assess Maudlin's B-theoretic version of the initialness defence in the context of the past hypothesis, and Huw Price's C-theoretic rejection of it. Section 5 presents a C-theoretic 'deflationary' version of the initialness defence, motivated by a position outlined by Hans Reichenbach, and argues the deflated initialness defence preserves the ontological simplicity of the C-theory whilst marking out the past hypothesis as explanatorily special. Section 6 is the conclusion.

2 Explaining the arrow of time

2.1 The time asymmetry of thermodynamics

Imagine an ordinary irreversible process like the smashing of a wine glass: I take a wine glass out of the cupboard, drop it on the floor, and it shatters into tiny pieces. Such a process is (hopefully) not especially familiar but it's certainly far more familiar than the temporal opposite process; the process run from future to past involves a bunch of shards of glass spontaneously starting to jiggle on the kitchen floor due to an injection of impetus from heat on the floor causing the shards to jump together, forming a smooth wine glass, before bouncing up into my hand. Whereas the original process certainly sounds a plausible candidate for what can happen in a kitchen, the second process does not. This asymmetry between two processes here reflects a greater asymmetry of things in time described by the second law of thermodynamics. One of the most familiar statements of the second law is Rudolf Clausius' (1864, p. 44) 'Die Entropie der Welt strebt einem Maximum zu,' or 'the entropy of the world tends towards a maximum.' This introduces the notion of entropy, which is taken to be a measure of the energy of a system from which work cannot be extracted. More familiarly, we can state the second law as follows:

Second Law of Thermodynamics. The entropy of an isolated subsystem of the world does not decrease.[2]

Or, denoting entropy by the letter S,

$$\frac{dS}{dt} \geq 0 \tag{1}$$

2.2 Boltzmannian statistical mechanics

The latter half of the 19th century saw the development of statistical mechanics in the work of (among others) Maxwell, Boltzmann and Gibbs, with the central idea being that thermodynamic regularities can be accounted for in terms of the statistics of the constituents of sufficiently large mechanical systems, or in Gibbs' case the statistics of ensembles of mechanical systems. The discussion of this chapter is based on the Boltzmannian approach to statistical mechanics.[3] To understand the Boltzmannian statistical mechanical account, we need a bit of mathematics. The state of a classical Hamiltonian system of n particles is given by the specification of the positions and momenta of all the particles in the system. To simplify, we can use the concept of a *phase space*, denoted by the symbol Γ. A phase space is an abstract mathematical space in which each point represents a complete kinematically possible microstate of the system in question—in this case the position

[2]There exist many different statements of the second law (see Uffink (2001)), and many different definitions of entropy (see Maroney (2009)). As Price (2002) notes, the precise version of the second law, and the definition of 'entropy' are relatively inessential to the issue of the relation between time direction and thermodynamics. The example of the wine glass is sufficient to illustrate the kind of time asymmetry with which this chapter is concerned.

[3]There are two main approaches to statistical mechanics, owing to the work of Boltzmann and Gibbs respectively. I shall work within the Boltzmannian framework. Gibbsian statistical mechanics (Gibbs, 1902) is distinct from Boltzmannian statistical mechanics in a number of ways. The Gibbsian approach is characterised by considering an *ensemble* of systems, each with slightly different microscopic states. As Frigg (2009) notes, both approaches are widely used within the literature but generally for different purposes: the Gibbsian approach is used in practice (i.e. experiments); the Boltzmannian approach is used in consideration of foundational issues (e.g. philosophical issues concerning the reduction of thermodynamics to statistical mechanics). See Wallace (2020) and Robertson (2020, 2021) for recent defences of the Gibbsian approach.

and momentum values of each particle.[4] For a system of n particles, the phase space has $6n$ dimensions, with a point in the $6n$ phase space picks out a unique set of position and momentum values (in three-dimensional physical space) for each particle in the system. Given this, the evolution of a many-particle classical system can be illustrated by a curve in Γ, and thus the dynamics describes curves in Γ. A point in Γ picks out a *microstate* of the system in question. In thermodynamics, a state of a system is given in terms of macroscopic (observable) parameters, such as the pressure, volume and temperature of a gas. One of the key observations in the development of statistical mechanics is the identification of the thermodynamic *macrostate* of a system with a *region* of Γ. In this sense, the macrostate of a system supervenes on its microstate—there can be no change in the system's macrostate without a change in its microstate—but not *vice versa*.

Central to the Boltzmannian account is the application of the 'standard' (Lebesgue) measure to the 'microcanonical ensemble'—that each microstate on a hypersurface of fixed energy on the phase space is assigned equal probability.[5] Given this, regions of phase space can be assigned measurable *volumes*. Boltzmann famously provided the following statistical mechanical definition of the entropy, S, of a system as proportional to natural logarithm of the volume of its phase space corresponding to its macrostate, where, given the standard measure, W corresponds to the *amount* of microstates corresponding to the system's

[4]In order for a point in phase space to play this role, the phase space must have the right dimensionality. The state of a one-particle system is provided by two quantities: the position and the momentum of the particle. If the particle is in a three-dimensional *physical* space, then its position and momentum each have coefficients in each spatial dimension: the position is provided by the triple (q_x, q_y, q_z) corresponding to its location in each spatial dimension (x, y and z); and the momentum is provided the triple (p_x, p_y, p_z). Thus, for a one-particle system, the phase space has 6 dimensions—three for position; three for momentum—such that each point in this space denotes a unique set of position and momentum values, with the entire space covering every possible state of the one-particle system.

[5]This assumption is one of the central problems concerning Boltzmannian statistical mechanics—see Sklar (2006) for a survey. The justifiability of Boltzmannian microcanonical probability distribution has historically concerned Liouville's Theorem (which holds that regions of phase space of fixed energy are invariant under Hamiltonian evolution) and the Ergodic Hypothesis. See Frigg (2008) for a discussion of the relation between the ergodic hypothesis and Boltzmannian statistical mechanics.

macrostate, and k is Boltzmann's constant ($1.38062 \times 10^{-23} JK^{-1}$).

$$S = k \ln W \tag{2}$$

This tells us that the larger the volume of phase space corresponding to some macrostate of a system X, the greater the entropy of X. It follows from this that for any gas, given no further constraints on its macrostate than the Boltzmannian statistical considerations, at any time the gas is most likely to be in the highest-entropy available macrostate and hence in equilibrium.

To complete the necessary details of the Boltzmannian account, we need to make an assumption about the dynamics of points in phase space—*micro*states—namely that they are quasi-ergodic.[6] The ergodic hypothesis holds that the infinite time averages of systems are such that the proportion of time spent in each macrostate corresponds to the volume of the macrostate given by the standard measure. What this means is that for any point in Γ (except for a set of measure zero), its evolution over infinite time is such that the time spent in each of its available macrostate is directly proportional to its volume.[7] Given the assumption that systems are quasi-ergodic, it follows that systems in small regions of phase space (low entropy) tend to evolve towards larger regions of phase space (higher entropy), and spend the majority of time in the largest region of phase space, which corresponds to the equilibrium state—maximal entropy. It follows that given a non-maximal entropy macrostate, it is overwhelmingly probable that its microstate is such that its evolution leads towards macrostates of higher entropy.

As such, we can see quite simply that the reason non-equilibrium systems tend towards higher entropy is that there are more *ways* for a system to have high entropy. Given this, as Wallace (2011, p. 1) puts it, the overwhelming tendency of

[6]Frigg and Werndl (2011) argue that a weaker property that that of ergodicity—'epsilon ergodicity'—is sufficient for Boltzmannian statistical mechanics and avoids the problems of the KAM and Marcus-Meyer theorems.

[7]The status of the ergodic hypothesis has been the subject of a particularly massive literature that does not concern us in the present chapter. See Sklar (1993, ch. 5) for a review.

systems to approach equilibrium is 'basically just a consequence of the geometry of phase space.' For a system of low entropy left to evolve in phase space, it will invariably tend towards higher entropy simply because high entropy macrostates take up much larger regions of phase space than do low entropy macrostates. This gives us two kinds of 'normal' or 'typical' states for systems; ones that, assuming nothing else, systems can be expected to be in without requiring further explanation:

Normal Macrostate. A normal macrostate for a system is one that corresponds to maximal entropy (its 'equilibrium state').

Normal Microstate. A normal microstate for a system in non-equilibrium is one that evolves towards equilibrium.[8]

2.3 What is the past hypothesis?

The problem with this kind of statistical mechanical reasoning is that it works too well. Towards the future, where we expect entropy to increase, the reasoning works fine. But if we try to use the same reasoning towards the past then we should also think that it's far more likely for any system in a non-maximum-entropy state to have evolved from a higher-entropy state and not from a lower-entropy state, contrary to our beliefs and apparent observations. This is a major problem insofar as the statistical method only gives us a time asymmetry of entropy if we explicitly assume that in the past entropy was far lower than it is now. Without such an assumption, the statistical method implies that both the future and past are higher entropy than the present, so we end up with a sceptical problem about the past (implying that our apparent records and memories of the past are false) and are unable to ground the time asymmetry of the Second Law.

Boltzmann suggested that such a low-entropy assumption about the past is a crucial aspect of the statistical explanation of the second law:

[8] A normal microstate for a system *in equilibrium* is one that stays in equilibrium (in both temporal directions).

> The second law of thermodynamics can be proved from the mechanical theory if one assumes that the present state of the universe [...] started to evolve from an improbable state and is still in a relatively improbable state. (Boltzmann, 1967)

The idea suggested by Boltzmann in this passage is that we need to put in an assumption of a low-entropy past in by hand, and that once we do that then the statistical method gives back the tendancy for entropy to increase monotonically towards the future.[9] The same conclusion is reached by Richard Feynman in his *Lectures on Gravitation*:

> The success of [scientific inferences] indicates that the world did not come from a fluctuation [...] Therefore I think it is necessary to add to the physical laws the hypothesis that in the past the universe was more ordered [...] than it is today—I think this is the additional statement that is needed to make sense, and to make an understanding of [thermodynamic] irreversibility. (Feynman et al., 1971)

This idea that we need to make an assumption of a low-entropy past to complement the Boltzmannian statistical method is elaborated on by Albert (2000), where he terms the idea the 'past hypothesis', which has become the common terminology in the literature. The idea is that the past hypothesis complements the statistical mechanical account of the entropy gradient by stipulating a low entropy constraint on a past macrostate of the world. Thus, for some non-maximal entropy system, such as a half-melted ice cube on a warm plate, the statistical mechanical *pre*diction is unconstrained—we predict that the ice cube will be more melted in 5 minutes—, but the *retro*diction is constrained insofar as it is conditional upon the stipulated low-entropy past macrostate. Call this macrostate (following Callender 2004a) the *Past State*. We can define the past hypothesis as

[9] However, this remark is not elaborated on in much detail by Boltzmann, and he elsewhere entertained the hypothesis that the low entropy past is simply the result of a large fluctuation from thermal equilibrium made probable by the universe being significantly older than standard cosmological models assume, and for the most part in thermal equilibrium.

follows:

Past Hypothesis. The hypothesis that at some point in the distant past, the world occupied an extremely low-entropy macrostate—the Past State.

On Albert's (2000) account, the past hypothesis corrects the faulty retrodictions of statistical mechanics, namely the overwhelming likelihood that the past was higher entropy, implying that all apparent records and memories of a lower-entropy past are misleading. The big problem with the past hypothesis is the sheer improbability of the Past State. Penrose (1989) calculates the probability of the universe having been in a state of sufficiently low entropy to be 1 in $10^{10^{123}}$.[10] According to the past-hypothesis-based statistical mechanical explanation of the entropy gradient, the assumption that such a low-probability state occurred is a key part of the explanation of any mundane thermodynamic regularity, such as the inevitably underwhelming final sips of my now-cold flat white.

Boltzmann himself thought the past hypothesis 'is a reasonable assumption to make since it enables us to explain the facts of experience, and one should not expect to be able to deduce it from anything more fundamental' (Boltzmann, 1967), emphasising the lack of a need to offer a special explanation of the past hypothesis itself. Various authors have offered a range of explanatory options, from Boltzmann's no-explanation-needed option, to Albert's related suggestion that the past hypothesis possesses a Kantian-type epistemic status in that it is a necessary *precondition* for knowledge of the past, to Callender's (2004a) suggestion that the past hypothesis has the status of a non-dynamical law of nature. But we are here going to assess only the role of the initialness defence in the explanation of the past hypothesis, and to do that we'll focus on Maudlin's initialness defence, Price's rejection of such an approach, and my own preferred 'deflated' version of the initialness defence that is motivated by remarks made by Hans Reichenbach (1956).

[10] Though it should be used that Penrose's calculation is in the context of black hole thermodynamics, using Bekenstein–Hawking entropy rather than Boltzmann entropy.

So, we are in a position now to present the past-hypothesis-based statistical mechanical explanation of the entropy gradient with which the rest of the chapter is concerned: the universe is, at one temporal end, in a macrostate of very low entropy (the Past State), and there is a monotonically-increasing entropy gradient towards the other temporal end (entailed by applying the Boltzmannian method to the Past State and projecting in one temporal direction) until a maximal-entropy state at the other temporal end, which we may term the *Future State*. On this model, the past hypothesis plays two main problem-solving roles. Firstly, it allows for there to be an entropy gradient at all (without the existence of a low-entropy state, there'd be no room for entropy to increase). Secondly, the past hypothesis corrects the retrodictions of statistical mechanics; we need the past hypothesis to validate our beliefs about a low-entropy past.

Even given the past hypothesis, there is still a major problem. Towards the past we have processes which are anti-thermodynamical in that they are entropy-decreasing and hence highly improbable. Maudlin (2007) argues that the assumption of a fundamental direction of time allows us to make an initialness defence here: the apparently improbable behaviour towards the past is really just probable behaviour towards the future, with the lowest-entropy state (the Past State) being the temporally earliest, ensuring all 'real' evolution is entropy-increasing and so thermodynamically normal. On the contrary, Price (2002, 2004), an antirealist about time direction, holds that this improbable behaviour towards the past is reason to think that the past hypothesis requires further explanation. I shall now argue that this sets up a false dichotomy between realists and antirealists about time direction. Instead, I argue that antirealists about time direction, or 'C-theorists' as I will call them, can and should adopt a restricted version of the initialness defence. In order to set this out, I'll first introduce the C-theory of time and what it means to reject that time has a direction.

3 Directionless time: Understanding the C-theory

In order to set up the opposing camps regarding the initialness defence, I'll introduce two alternative theories of time: the B-theory; and the C-theory. We can start with simple definitions of the two theories before adding some details to how they differ.

B-theory of time. The world is directed in time from earlier to later.

C-theory of time. The world is not directed in time.

The B- and C-theories derive from McTaggart's (1908) B-series and C-series. McTaggart presented three different understandings of how things could be ordered in time. The A-series orders things in terms of whether they're past present or future, such as the extinction of the dinosaurs being past, the displaying of dinosaur fossils in the Natural History Museum present, and the genetically-modified-dinosaur apocalypse future. The A-series captures how things appear to change in time: for example, your reading of this sentence is, for me (at the time of writing), future, but for you it is present, and as of the time of your reading the next sentence it is past. In this sense events change their A-series properties as time passes. The B-series preserves the directionality of the A-series, ordering things from earlier to later — e.g. my writing of this sentence is earlier than your reading it. But the B-series doesn't change in the way the A-series does; if x is earlier than y, this is a fact that does not change over time in any sense, so there is no need to update or redraw the B-series over time.

McTaggart also introduced the much-less-talked-about C-series, which he took to provide an undirected ordering of things in time, noting that the C-series, 'while it determines the order, does not determine the direction' of a set of events. The implied ordering relation of the C-series in McTaggart (1908) is *temporal betweenness*: a C-series tells us what events are between which other events in time, with this betweenness ordering being insensitive to whether the ordering goes in one time direction rather than the other.

> [T]he *C* series, while it determines the *order*, does not determine the *direction*. If the *C* series runs M, N, O, P, then the *B* series from earlier to later cannot run M, O, N, P, or M, P, O, N, or in any way but two. But it can run either M, N, O, P (so that M is earliest and P latest) or else P, O, N, M (so that P is earliest and M latest). And there is nothing [...] in the *C* series [...] to determine which it will be. (McTaggart, 1908, p. 462; my emphasis.)

In the C-series, the event of you reading this sentence is located between your reading of the previous sentence and the next sentence, but this 'temporal betweenness' ordering has no directionality built into it. Building time out of an undirected relation like betweenness gives a picture of time as ordered but undirected, with one direction no more important than the other. A consequence of this is that there is no such thing as the temporal reverse of a C-series of events, since whichever direction one chooses to describe events, the same betweenness relations hold (even 'backwards' in time your reading of this sentence is between your reading of the previous and next sentences).

Unlike the A-theory, the B- and C-theories contain no special 'present moment' of time — the time we consider present can be considered to be indexical, like the term 'here', in that what is 'present' is simply what is simultaneous with us at the particular relevant moment of time that we are thinking or speaking. As such, there is no observer-independent flow or passage of time on the B- and C-theories — they are both versions of the so-called 'block universe', wherein events we think of as being in the past or future are as real as events we think of as present. The difference between the B- and C-theories is simply that the B-theory contains a basic directionality of time — it is a fact that the world runs from earlier to later and not later to earlier, that the Big Bang happened at the beginning and not the end of time. On the C-theory, the universe does not contain sufficient temporal structure to distinguish between an earlier-to-later universe and a later-to-earlier universe, and as such there is no such fact about the Big Bang being at the beginning rather than the end of time.

	Dynamic	Directed	Ordered
A-theory	✓	✓	✓
B-theory	✗	✓	✓
C-theory	✗	✗	✓

Table 1: A-theories, B-theories and C-theories understood in terms of a descending hierarchy of structure attributed to time, from Farr (2020).

Though the terminology might be new, various people have defended versions of the C-theory. Thomas Gold, when setting out a temporally adirectional position motivated by the time symmetry of classical mechanics, noted:

> [T]he description of our universe in the opposite sense of time [...] sounds very strange but it has no conflict with any laws of physics. Our strange description is not describing another universe, or how it might be but isn't, but it is describing the very same thing. (Gold, 1966, p. 327)

A decade earlier, Hans Reichenbach, in a book that almost certainly influenced Gold's position, held:

> Positive and negative time supply equivalent descriptions, and it would be meaningless to ask which of the two descriptions is true. Reichenbach (1956, pp. 31–32)

The basic idea behind both claims is that whether we describe things from past to future, or from future to past, we are simply giving two alternative descriptions of the same possible world; there's no such thing as *this world but run backwards in time*; there's no basic underlying direction of time. If you try to imagine this universe running backwards in time, the Reichenbach–Gold view holds that you're just imagining the same universe but from a different and unfamiliar temporal perspective. Crucially, all the same temporal facts (i.e. the temporal betweenness facts) are the same across these two alternative time-directed descriptions.

Though both these quotes are from within the discussion of the time symmetry of classical mechanics, Reichenbach (1956, pp. 128–9) also discusses this C-theoretic attitude within the context of time-asymmetric laws like the second law of thermodynamics, noting that 'it has no meaning to say that entropy "really" goes up or that its time direction is "really" positive', but rather that we conventionally prefer to talk about the universe in the direction in which entropy increases; if we were to describe it in the direction which entropy decreases (from future the past) we wouldn't be getting anything wrong about time, but would simply going against the standard convention and adopting an unfamiliar temporal perspective.

4 Initialness and the past hypothesis

With the B- and C-theories of time in hand we can think about how they differ when it comes to explaining the low-entropy past.

4.1 Maudlin's initialness defence

Maudlin (2007) is a prominent defender of the version of the B-theory.[11] Though an advocate of the block universe, he takes there to be a primitive directionality of time (which he calls 'passage', though devoid of the metaphysics of the A-theory) in virtue of which later states of the world 'are produced by' earlier states of the world, with no such feature being assigned to the opposite direction of time. On this picture, it is a fact that we live in an entropy-increasing world and not an entropy-decreasing world, something Maudlin takes not to be available to the antirealist about time direction (the C-theorist). Crucially, Maudlin takes this B-theoretic fact to play a key role in explaining the thermodynamic asymmetry in a very specific way.

[11] Indeed, Maudlin (2007, p. 126, fn. 11) describes his own position as a 'B-series theory' as opposed to a 'C-series theory', fitting with the B/C-theory distinction I've set out.

Maudlin distinguishes between two different ways in which entropy-increasing processes (such as the spreading out of gas in a room, the melting of an ice cube, and the smashing of wineglasses) are atypical and hence needy of explanation. Firstly they are *macroscopically atypical* in that their initial states are low-entropy. But there is also a second kind of atypicality we can call *dynamical atypicality*: the microstates of entropy-increasing systems are atypical insofar as backwards evolution (towards the temporal past) leads towards lower entropy. Imagine that midway through the process of an ice cube melting on a warm plate you freeze-frame the system. Given the assumption that the ice cube was less melted in the past, it follows that if you imagine this process running in reverse, the microstate of the system would have to be vastly improbable insofar as the arrangement of positions and momenta of the molecules would lead the system to a state of macroscopically lower entropy, despite there being vastly more microstates compatible with the present macrostate that would lead towards *higher* entropy in the past. In other words, any apparently ordinary thermodynamical process is dynamically atypical if viewed from future to past. Of this kind of dynamical atypicality, Maudlin remarks:

> [The microstates of such systems are] atypical in a way that can only be characterized in terms of how the state will 'evolve' though time [in that...] temporal evolution in one direction from it will lead, over a very long period of time, to monotonically lower entropy. (Maudlin, 2007, pp. 132–3)

Maudlin's solution is to hold that past-to-future evolutions of systems are real and future-to-past evolutions are not, and so there is no sense in which systems in the world are actually dynamically atypical. In other words, once the direction of time (from past to future) is stipulated as a fundamental fact about the world, this entails that macroscopic systems only evolve towards more probable and not less probable states over time. As such, the problem of dynamical atypicality, Maudlin argues, is explained away on the B-theory.

> [Dynamical atypicality] is completely accounted for by how it was generated or produced[; it is] the product of an evolution from [an] initial state[; and] [t]his sort of explanation requires that there be a fact about which states produce which[, which] is provided by a direction of time. (Maudlin, 2007, pp. 133–4)

This is not to say that the macroscopic atypicality of the past hypothesis is explained away, but rather that the fact that systems behave anti-thermodynamically towards the past is explained away. On Maudlin's B-theoretic account, the Past State is not the 'result' of temporal/causal evolution; it is this that makes Maudlin's argument a version of the initialness defence. It is thus implied by Maudlin that dynamical atypicality is a genuine problem for the C-theorist, since by rejecting a fundamental direction of time the C-theorist is committed to the idea that future-to-past evolutions are as real as past-to-future evolutions, and so in this way the B-theorist has recourse to an initialness defence that the C-theorist lacks.

4.2 Price's rejection of the initialness defence

So is Maudlin right to hold that the C-theorist is troubled by dynamical atypicality in a way the B-theorist is not? I'll argue that he is not, and that the C-theorist does indeed have access to the right kind of initialness defence here. But first it is important to show that at least some defenders of the C-theory appear to agree with Maudlin's contention. Consider, for example, the following passage from Price:

> Imagine that in recent years physics had discovered that the matter in the universe is collapsing towards a big crunch [...] 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 matter will [...] spread itself out with great uniformity. (Price, 2004, p. 230)

But, Price continues,

> [T]his state of affairs is exactly what physics has discovered! I've merely taken advantage [...] of the fact that [...] there is no objective sense in which what we call the future is *really* the 'positive' direction of time. [...] Relabelled [from future to past], 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 way. (Price, 2004, pp. 230–1; emphasis in original)

Though Price presents this in the form of an intuition pump, we can go further and construct it in the form of a deductive argument as follows:

P1. Forwards and backwards time are equivalent [central thesis of C-theory]

P2. An evolution from a low-entropy past state is equivalent to an evolution towards a low-entropy 'future'[12] state [implied entailment of P1]

P3. An evolution towards a low-entropy 'future' state would demand a special explanation [intuition]

C. Therefore, an evolution from a low-entropy past state demands a special explanation [from P2 & P3]

It is not clear whether Price explicitly endorses this exact argument, but these passages function as an intuition pump for P3 based on the assumption of P2, the key premises of the argument. It is P2 that I think can and should be rejected.

[12] Here, a 'future' state is simply the Past State described in negative time.

5 Deflating the initialness defence

5.1 Reunderstanding the C-theory

Rejecting P2 involves putting forward a different understanding of the C-theory. Such a version is motivated by remarks made by Reichenbach (1956), which I elsewhere (Farr, 2020, 2021) term *time direction conventionalism*, according to which there are no time-directed facts, but it is more convenient to describe the world from past to future. Central to the Reichenbachian account is the idea that although past-to-future and future-to-past descriptions of the world are equivalent, it does not follow that it is as true to say that the Past State is a product of entropy-decreasing behaviour as it is to say that the present state is the product of entropy-increasing behaviour. As such, this offers a middle-ground option that preserves intuitions behind both Maudlin's and Price's approaches by holding that there is an objective causal direction of processes in time, but that it is as true to describe things from future to past as from past to future.

Whilst discussing past-to-future and future-to-past descriptions of the entropy gradient, Reichenbach (1956, p. 154) remarks:

> The two languages L_1 [a language in which the positive direction of time is that of increasing entropy] and L_2 [a language in which the positive direction of time is that of decreasing entropy] represent *equivalent descriptions*; one is as true as the other.

But he adds:

> If someone argues that it is a matter of convention to select the direction of growing entropy [as opposed to decreasing entropy] as the direction of time, [their] conception cannot be called false. But [they] must not commit the error often connected with other forms of conventionalism: the error of overlooking the empirical content associated with the use of this convention. (Reichenbach, 1956, p. 154)

Table 2: Four options for the C-theorist

	Causal	Teleological
Entropy-increasing	1	2
Entropy-decreasing	3	4

In order to understand what he has in mind here, it is important to first clarify how Reichenbach views the relationship between the direction of cause–effect and direction of increasing entropy. Reichenbach (1956, p. 153–6) makes a subtle point that has rarely if at all been discussed in the secondary literature, namely that there are two separate conventions that guide our talk about the increase of entropy over time. Firstly there is the *time direction convention* – the tendency to describe the world in the direction of increasing entropy (and so past-to-future) and not in the direction of decreasing entropy. Reichenbach says there's nothing factually wrong about adopting the opposite convention, but it simply goes against our standard way of talking about the world. There are pragmatic grounds for adopting the standard convention; for instance Reichenbach (1956, p. 154) suggests that it matches the 'time direction of psychological experience' (though for criticism of this claim, see Farr (2022)).

Reichenbach also speaks of a second convention I'll call the *causal convention*, namely the convention that happenings in the world are explained by the temporal past rather than temporal future. The alternative convention would be to prefer a teleological mode of explanation, what Reichenbach refers to as a 'principle of finality', according to which things are explained in terms of the temporal future. What is key is how the time-direction convention and causal convention go together. See Table 2 for the four options these provide. On Reichenbach's conventionalist account, Options 1 and 4 are equivalent descriptions, and Options 2 and 3 are equivalent descriptions. But neither 1 or 4 are equivalent to 2 or 3.

Reichenbach takes the former of these equivalence classes (1 and 4) to adequately describe the world: the world is *either* a causal entropy-increasing world

(Option 1), or a teleological entropy-decreasing world (Option 4), noting that when taking the direction of time to be a convention, one must not overlook 'the empirical content associated with the use of this convention' (ibid). The empirical fact Reichenbach points out is what he calls the 'parallelism of entropy increase' for systems: for quasi-isolated subsystems of the universe, as well as the universe itself, the directions of increasing entropy are parallel in that they increase in entropy *in the same direction* and not in different directions.[13] This allows us to talk of a general positive time direction being defined, which Reichenbach takes to be an 'empirical hypothes[is] which [is] convincingly verified' (Reichenbach, 1956, p. 154). This does not pick out the direction of entropy-increase or entropy-decrease on its own, since the parallelism itself is not directed. But Reichenbach ties this parallelism to the notion of causality, noting that 'the convention of defining positive time through growing entropy is inseparable from accepting causality as the general method of explanation' (ibid.), and further:

> Once the direction is assumed in the usual sense [i.e. positive time is the direction of increasing entropy], it is not a matter of personal preference, not a mode of consideration, whether we should describe the world in terms of causes or of ends: it is a physical law that causality, and not finality, governs the universe. (Reichenbach, 1956, p. 154)

This can be understood in terms of the asymmetric structure of causal networks, in which correlations between variables that are not the direct cause or effect of each other are explainable in terms of their causal past and not their causal future. Reichenbach discusses this in the context of his principle of common cause (Reichenbach, 1956, ch. 19), remarking that while some causal forks are closed to the future, such as when two independent events have a common effect,

[13] Reichenbach's principle of parallelism of entropy increase: 'In the vast majority of branch systems, the directions toward higher entropy are parallel to one another and to that of the main system' (Reichenbach, 1956, p. 136). Equivalently, the directions towards lower entropy are also parallel.

'a common effect cannot be regarded as an explanation' (Reichenbach, 1956, p. 163) of the correlation. As such, the method for establishing causal direction is tied to the entropy gradient (e.g. causes are lower in the entropy gradient than their effects). This asymmetry is taken by Reichenbach to be due to the second law of thermodynamics in that the existence of a sufficiently low-entropy state makes it possible for there to be the kind of past-directed common cause structure for correlated events, meaning that if the Past State is stipulated as the lowest entropy state, it cannot be considered to be causally influenced by any other temporal states of the world.

This understanding of the Past State gives a deflationary account of the direction of causation, entailing that there is a direction of causation *only* in the presence of the right kind of probabilistic asymmetries, such as irreversible processes, time-asymmetric screening-off conditions, and other such things common to the causal modelling literature. This position is taken by Reichenbach also in the case of time-symmetrical systems; for example in the case of classical particle mechanics Reichenbach (1956, ch. II) holds that the relevant causal notion is one of 'causal betweenness' — the physical laws are sensitive to which events are causally between others, but so long as there is no basic lawlike temporal asymmetry displayed by systems there is no sense in which some events *causally precede* others; for sufficiently time-symmetric systems, causal-direction talk breaks down. For our purposes, an important consequence of this is that entropy gradients, rather than being *explained* by further causal facts, are what allows for the existence of causal explanations of events. This runs contrary to Maudlin's realism about causal direction, which holds that 'even if the world were always in thermal equilibrium [...] later states would arise out of earlier ones' (Maudlin, 2007, p. 177), where by 'arise' Maudlin is referring to causal production, meaning that for Maudlin there would be causal direction facts even in the absence of the kinds of probabilistic asymmetries central to standard causal discovery algorithms like Reichenbach's common cause principle and modern descendents of it, such as the causal Markov condition (see Hausman and Woodward (1999)).

The convention of taking the direction of entropy-increase to be the direction of positive time is tied to the convention of causality over finality: once one is fixed, the other follows empirically. We can regard these as on a par — neither alone is strictly an empirical matter, but once one is chosen, the other is not open to choice.

> [T]he convention of defining positive time through growing entropy is inseparable from accepting causality as the general method of explanation. Those who prefer to give an explanation in terms of finality would be compelled to use the opposite time direction and to regard time as going from high-entropy to low-entropy. (Reichenbach, 1956, p. 154)

What's *not* a convention is that *either* the world is an entropy-increasing causal world *or* an entropy-decreasing teleological world.

Compare this to Price's intuition pump and the argument we constructed from it. Premise 2 holds that an entropy-increasing causal world is equivalent to an entropy-decreasing causal world, aligning with Price's suggestion that, when described from future to past, we see a world full of coincidences in which local causes force systems to decrease in entropy over time. But the C-theorist is not compelled to accept Premise 2 and can instead adopt the Reichenbachian option. The key difference between the Pricean and Reichenbachian approaches here is whether or not the Past State can reasonably be considered the result of a series of coincidences due to an improbably fine-tuned present state of the world. On the picture Price describes, the description of the world evolving thermodynamically from the Past State is equivalent to the description of the world evolving anti-thermodynamically towards the Past State. But on the Reichenbachian account, we can pull apart the time-direction convention and the causal convention and clarify that it is not true to say that Past State is the causal product of anti-thermodynamic evolution from a higher-entropy state, and so we get rid of the problematic unexplained coincidences apparent in Price's future-to-past description.

On the Reichenbachian account, when we switch from a past-to-future to a future-to-past description of the world, we switch a causal explanation of the entropy gradient for a teleological explanation of the entropy gradient, and as such, in both directions we have an explanation of the entropy gradient in terms of the past hypothesis and not vice versa. In this way we find a best-of-both-worlds option between Price and Maudlin's accounts: (1) we avoid Price's conclusion that an entropy-increasing evolution *from* the Past State requires a special explanation; and (2) we capture the idea from Maudlin's account that Past State is not the 'result' of temporal/causal evolution without committing to an underlying directionality of time. This shows that we can have a C-theory of time — one that holds that past-to-future and future-to-past descriptions of the world are equivalent — that nonetheless supports the idea there is preferred direction of causal explanation (lower-entropy states can be used to explain higher-entropy states but not vice versa).

5.2 Initialness, deflated

This takes us to the central conclusion: there is a 'deflated' version of initialness defence available to the C-theorist. On the one hand, B-theorists like Maudlin take the past-to-future direction as metaphysically privileged, and so the low-entropy past not as needy of explanation as a low-entropy future would be. And on the other, C-theorists like Price give up the initialness defence on the grounds that the description of the world as evolving away from a low entropy past is equivalent to the alternative description of the world as evolving towards a low entropy future. But the best-of-both worlds option given by conventionalism holds instead that when described from future to past, the story of the past hypothesis needs to be understood in teleological rather than in causal terms, so that entropy-lowering processes are explained in terms of a fact about their temporal future. On this view, the past hypothesis is still explanatorily special insofar as other temporal states of the world are explained by way of reference to it and there is no equivalent way in which it is explained in terms of other states of the world.

On the Reichenbachian approach, we get back a C-theoretic, deflationary version of Maudlinian Production; if we are to use the language of causal production to explain the entropy gradient, then the Past State is a producer but not a product of other temporal states of the world. This is something briefly suggested Reichenbach (1956):

> Once the time direction is assumed in the usual sense, it is not a matter of personal preference [...] whether we should describe the world in terms of causes or ends: it is a physical law that causality, and not finality, governs the universe. [...] For the time direction of growing entropy [...] the interaction point is the beginning, not the end, of the evolution of the branch system. The statistical relationships [...] account for our conception that the past *produces* the future, and not vice versa. (Reichenbach, 1956, p. 155; my emphasis)

Recall that for Maudlin, a 'primitive',[14] 'irreducible' and 'fundamental'[15] directionality of time plays a crucial role in this initialness defence, explaining the dynamical atypicality of microstates at all times apart from the initial time. Maudlin's argument is an inference to the best explanation in favour of the B-theory. It follows from Maudlin's reasoning that only by assuming such a direction of time (from past to future) can we explain away dynamical atypicality. He adds, in defence of the B-theory:

> [W]e cannot run this trick in reverse. [...W]e cannot specify an independent, generic constraint on the *final* state that will yield (granting the final macrostate is typical) ever decreasing entropy in one direction. (Maudlin, 2007, p. 133; emphasis added)

But this is simply not the relevant contrast class. The reason we can explain the Future State[16] in terms of the Past State and not vice versa is not because the Past

[14] Maudlin (2007, 172).

[15] Both Maudlin (2007, p. 3 & p. 107).

[16] Recall that the 'Future state' is the hypothetical maximal entropy state in our distant future.

State is earlier than the Future State. It is instead because, given the Boltzmannian statistical mechanical account, we can explain the Future State by conditionalising on the Past State, and we cannot explain the Past State by conditionalising on the Future State. The explanans here is simply that the Future State is a *normal* macrostate, and the Past State is not. Whether the Past State is 'really' earlier or later than the Future State does not enter into this explanation. If the Past State is earlier than the Future State, then statistical mechanics dictates that the 'forwards' direction of time provides the natural (causal) explanation. However, if we flip to using the future-to-past direction, it is now the 'backwards' direction of time that provides the natural (teleological) explanation.

As such, it is not at all clear why a primitive or fundamental directionality of time needs to enter the explanation here. Consider two different models of the entropy gradient using a primitive directionality of time: one in which entropy increases over time; and one in which entropy decreases over time. The two different models correspond to different B-series. The statistical mechanical explanation of the Future State in terms of the Past State, and the lack of an explanation of the Past State in terms of the Future State, is itself independent of the issue as to which state is 'earlier than' the other. The statistical mechanical explanation *can only go one way — from* past *to* future, and this is respected by Reichenbachian time-direction conventionalism, which holds that regardless of whether we choose an entropy-increasing or entropy-decreasing description, the entropy gradient is explained in terms of the existence of the Past State and not vice versa. The important thing is that *regardless* of whether we use a past-to-future or future-to-past description, the explanation goes *from* the Past State (i.e. the *low entropy* state) *to* the Future State (i.e. the *high entropy* state), and as such there is no clear need for a further background directionality of time.

6 Summing up

We started with a puzzle about the nature of initial conditions: in what sense do laws of nature require there to be genuinely 'initial' conditions, and so for time to have some basic or fundamental preferred direction? By focusing on the problem of the low-entropy early universe, we considered a potential role for a fundamental directionality of time — that the dynamical atypicality of the world when viewed from future to past can be explained away only on the assumption that really time runs from past to future. I have argued that dynamical atypicality is not as serious a problem as it appears, and that the C-theorist of time can perfectly well make sense of the Past State as explanatorily special without there being a fundamental direction of time. On the version of the C-theory defended, the low-entropy early universe is not intrinsically special through being 'temporally first' or anything to that effect. Rather, due to the relationship between entropy and causality, it is the existence of a low-entropy state that allows there to be a causal direction of systems and hence a clear direction of explanation. As such, initialness defences are available to those that reject that time has a direction.

References

Albert, D. Z. (2000). *Time and Chance.* Massachusetts: Harvard University Press.

Boltzmann, L. (1967). On Zermelo's Paper "On the mechanical explanation of irreversible processes". In S. Brush (Ed.), *Kinetic Theory, Vol. 2 – Irreversible Processes.* Pergamon.

Callender, C. (2004a). Measures, explanations and the past: Should 'special' initial conditions be explained? *British Journal for Philosophy of Science 55*(2), 195–217.

Callender, C. (2004b). There is no puzzle about the low entropy past. In C. Hitchcock (Ed.), *Contemporary Debates in Philosophy of Science*, pp. 240–256. Oxford: Blackwell Publishing.

Clausius, R. (1864). *Abhandlungungen über die mechanische Wärmetheorie*, Volume 1. Braunschweig: F. Vieweg.

Farr, M. (2020). C-theories of time: On the adirectionality of time. *Philosophy Compass 15*(12), e12714.

Farr, M. (2021). Conventionalism about time direction. *Unpublished manuscript (coming soon...)*.

Farr, M. (2022). Perceiving direction in directionless time. In K. M. Jaszczolt (Ed.), *Understanding Human Time*. Oxford University Press.

Feynman, R. P., F. B. Morinigo, and W. G. Wagner (1971). *Feynman Lectures on Gravitation*, Volume 13. California Institute of Technology Pasadena, CA.

Frigg, R. (2008). A field guide to recent work on the foundations of statistical mechanics. In D. Rickles (Ed.), *The Ashgate Companion to Contemporary Philosophy of Physics*, Chapter 3, pp. 99–196. London: Ashgate.

Frigg, R. (2009). What is statistical mechanics? In C. Galles, P. Lorenzano, E. Ortiz, and H. J. Rheinberger (Eds.), *History and Philosophy of Science and Technology*, Volume 4 of *Encyclopedia of Life Support Systems*. Isle of Man: Eolss.

Frigg, R. and C. Werndl (2011). Explaining thermodynamic-like behavior in terms of epsilon-ergodicity. *Philosophy of Science 78*(4), pp. 628–652.

Gibbs, J. W. (1902). *Elementary Principles in Statistical Mechanics*. Woodbridge: Ox Bow Press.

Gold, T. (1966). Cosmic processes and the nature of time. In R. G. Colodny (Ed.), *Mind and Cosmos*, pp. 311–329. Pittsburgh: University of Pittsburgh Press.

Hausman, D. and J. Woodward (1999). Independence, invariance and the causal markov condition. *The British Journal for the Philosophy of Science 50*(4), 521–583.

Maroney, O. (2009). Information processing and thermodynamic entropy. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2009 ed.).

Maudlin, T. (2007). *The Metaphysics Within Physics.* Oxford: Oxford University Press.

McTaggart, J. M. E. (1908). The unreality of time. *Mind 17*(68), 457–474.

Penrose, R. (1989). *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics.* Oxford: Oxford University Press Press.

Price, H. (1996). *Time's Arrow and Archimedes' Point: New directions for the physics of time.* Oxford: Oxford University Press.

Price, H. (2002). Boltzmann's time bomb. *The British Journal for the Philosophy of Science 53*(1), 83–119.

Price, H. (2004). On the origins of the arrow of time: Why there is still a puzzle about the low-entropy past. In C. Hitchcock (Ed.), *Contemporary Debates in Philosophy of Science*, pp. 219–239. Oxford: Blackwell Publishing.

Reichenbach, H. (1956). *The Direction of Time.* Berkeley: University of California Press.

Robertson, K. (2020). Asymmetry, abstraction, and autonomy: Justifying coarse-graining in statistical mechanics. *The British Journal for the Philosophy of Science 71*(2), 547–579.

Robertson, K. (2021). In search of the holy grail: How to reduce the second law of thermodynamics. *The British Journal for the Philosophy of Science.*

Sklar, L. (1993). *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics.* Cambridge: Cambridge Univ Press.

Sklar, L. (2006). Why does the standard measure work in statistical mechanics? In V. F. Hendricks, K. F. Jörgensen, J. Lötzen, and S. A. Pedersen (Eds.), *Interactions*, Volume 251 of *Boston Studies in the Philosophy and History of Science*, pp. 307–320. Springer Netherlands. 10.1007/978-1-4020-5195-1-10.

Uffink, J. (2001). Bluff your way in the second law of thermodynamics. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics 32*(3), 305 – 394.

Wallace, D. (2011). The logic of the past hypothesis. *PhilSci Archive http://philsci-archive.pitt.edu/8894/.*

Wallace, D. (2020). The necessity of Gibbsian statistical mechanics. In V. Allori (Ed.), *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*, Chapter 15, pp. 583–616. World Scientific.

Wigner, E. P. (1995). Events, laws of nature, and invariance principles. In *Philosophical Reflections and Syntheses*, pp. 321–333. Springer.