Philosophical Issues in Early Universe Cosmology

Adam Koberinski* †and Chris Smeenk*‡

October 25, 2023

Abstract

This article highlights some interesting foundational and philosophical issues made salient in early universe cosmology. One major focus is on issues that arise at the boundaries of distinct theories or frameworks, when trying to merge them for describing the early universe. These include issues at the boundary of gravity and statistical physics, as well as gravity and quantum field theory. These foundational issues arise in trying to unify distinct domains of physics. Another major theme of early universe cosmology is the methodological goal of finding dynamical explanations for striking features in the universe. This review surveys many examples of such a methodology—including the cosmic arrow of time, posits of a Past Hypothesis for the initial state of the universe, inflation, baryogenesis, and emergence of spacetime. There is much philosophical debate about the prospects for success of such a methodology; these are surveyed below.

1 Introduction

For the last half century cosmology has turned into an increasingly important testing ground for fundamental physics. In the late 1960s, Zel'dovich (and others) recognized the appeal of pursuing high energy physics through observations of what he called the "poor man's accelerator," the early universe. As $t \to 0$ in the hot big bang models, the interaction energy steadily increases, leading to an abundance of exotic particles. The mix of high-temperatures and densities in a dynamical spacetime provides promise of unifying quantum field theory, gravity, and statistical physics in theories of the early universe. But cheaper is not always better: there is only one run of the experiment to study, and backgrounds cannot be controlled, nor conditions varied, as with other accelerators. For many aspects of fundamental physics, the early universe now provides the sole avenue for investigation rather than a complement to earthbound experiments. Cosmology needs novel physics to describe the late universe as well,

---

[*] Department of Philosophy and Rotman Institute of Philosophy, University of Western Ontario, London, ON N6A 5BY, Canada

[†] akoberin@uwo.ca

[‡] csmeenk2@uwo.ca

in particular to account for the formation of structures such as galaxies, and the observed accelerating expansion. The name of cosmology's standard model, ΛCDM, acknowledges two novel types of matter-energy that contribute the overwhelming majority of mass-energy in the universe, cold dark matter (CDM) and dark energy in the form of a cosmological constant (Λ).[1] In both regimes, cosmology poses profound puzzles for fundamental physics and imposes significant constraints on the space of viable solutions.

This new role raises a challenge: how effectively can fundamental physics be developed and tested via cosmology? Over the past century, cosmologists have constructed an increasingly detailed account of the history of the universe drawing on all the resources of modern physics. Success has encouraged an ambitious research agenda, not just to describe the universe but to explain why it has the features it does. Pursuit of such ambitious aims has prompted strikingly open-ended debates, reflecting the poor fit between an understanding of scientific inquiry based on experimental physics and the distinctive nature of cosmology. Along with being limited to passive observations, cosmology often involves a strikingly different kind of theorizing than in other areas of physics — regarding, for example, the origins of the universe, or its global properties. At the boundaries of known physics, early universe cosmology makes salient foundational and philosophical issues from all areas of fundamental physics. Philosophers have contributed by clarifying the implicit principles guiding inquiry in cosmology, the conceptual obstacles facing theories with such an unusual subject matter, and intrinsic limitations to what can be tested. This review focuses on foundational debates in early universe cosmology. One unifying methodological thread for early universe cosmology is the drive to explain any non-generic features of our universe via dynamical effects, rather than including them in initial conditions. This leads to important foundational debates regarding the "initial state" of the universe.

In pursuing the more ambitious agenda, cosmologists have often treated a "theory of origins" as a supplement to dynamical descriptions of subsequent evolution, needed to improve on explanations of the following type: the universe is as it is, because it was as it was. Yet what form should a "theory of origins" take? In other areas of physics, initial and boundary conditions typically reflect the impact of the environment, or a decision regarding how to delimit the target system, and are not the primary focus of investigation. In early universe cosmology views regarding how to treat the initial state set the agenda for further theorizing. The discussion below considers three prominent views, highlighting that the choice among them reflects assumptions regarding the aims of cosmology with implications for a variety of topics in foundations of physics. §2 discusses the proposal that a low entropy initial state must be posited to account for time's arrow. Further debates concern whether this initial state can be explained, and what form that might take. The target of this explanation may be mistaken, however, if quantum gravity resolves the singularities predicted

---

[1] In what follows, the term 'Standard Model' refers exclusively to the standard model of particle physics, and 'ΛCDM' refers to what is sometimes also called the standard model (of cosmology).

by classical general relativity (GR), as discussed in §3. In place of an "initial state," there may instead be a well-defined continuation through a "bounce," or spacetime may "emerge" from a non-spatiotemporal basis. In §4, attention is turned to the dominant line of work in early universe cosmology, based on extensions of the Standard Model that predict phase transitions (among other effects). In direct contrast to the first view, researchers in this line often aim to eliminate "special" initial conditions entirely. Observed features of the universe should be explained as a consequence of dynamical evolution rather than appealing to features of the initial state, and early universe cosmology promises to provide insights regarding otherwise inaccessible dynamical regimes.

In all of these domains, physicists are pushing the boundaries of applicability of known physics. At these boundaries lies the promise for unification, but also deep foundational issues in reconciling known physics. These boundaries are where philosophy and physics meet, as new perspectives and clarity on foundations of current theory are essential for moving forward. Though there is an enormous literature (both empirical and theoretical) on cosmology, the philosophy of cosmology is a relatively new field.[2] The focus here is on providing an orientation to the philosophical concepts most illuminating to issues in early universe cosmology. This review draws on philosophy of science, cosmology, particle physics, and statistical physics; the issues in each field are not separable, and all play a role in shaping current debates. With regret this review does not cover the late time effects of dark matter and dark energy, which raise equally intriguing philosophical questions of a quite different character.[3]

This article assumes basic familiarity with the ΛCDM model; see Weinberg (2008a) for a self-contained textbook treatment, or Longair (2006) and Peebles (2020) for more historical accounts of the construction of ΛCDM. These sources provide an orientation to the extensive scientific literature on these topics, whereas the primary aim below in choosing citations has been to guide readers to useful sources in philosophy and foundations of physics.

2 Time's Arrow and the Past Hypothesis

Consider describing the history of some physical system in terms of a trajectory $\Sigma(t)$ – a sequence of instantaneous states indexed by a parameter t, where the state specifies values for all the relevant dynamical variables. In familiar simple cases, a point in an appropriate phase space represents the state, and a trajectory corresponds to a curve through phase space. For many familiar dynamical laws, if a trajectory is possible according to the laws, so is the same

---

[2] But see Chamcham et al. (2017) for a collection of papers in the philosophy of cosmology, as well as overviews in Ellis (2014) and Azhar and Butterfield (2017).

[3] See Sus (2014), Swart, Bertone, and Dongen (2017), Sanders (2010), and Vanderburgh (2003) for historical and philosophical discussions about dark matter in ΛCDM, and Rugh and Zinkernagel (2002), Ellis (2008), Bianchi and Rovelli (2010), Koberinski and Smeenk (2023), and Koberinski, Falck, and Smeenk (2023) for philosophical discussion about dark energy and the cosmological constant problem. Additionally, there is a 2021 special issue of Studies in History and Philosophy of Science dedicated to Dark Matter and Modified Gravity.

sequence of states run in reverse. This is true if the dynamics is time-reversal invariant, as is the case for Newtonian gravity.[4] This idea can be extended to describe a history of the entire cosmos, given an appropriate conception of the instantaneous state. Since the dynamical laws relevant in cosmology are time-reversal invariant, a movie representing the history of the universe could run in either direction. Yet the actual universe exhibits striking temporal asymmetries: it expands and cools, initial clumps of matter form larger structures, particle species "freeze out" from equilibrating reactions with a characteristic density (discussed later), and so on, and the reverse of these processes is never observed. As with the ubiquitous asymmetries of everyday experience, the fact that these processes are governed by time-reversal invariant dynamics raises foundational questions:[5] what accounts for these asymmetries, also known as "arrows of time"? And how are the various observed arrows related to one another?

Since Boltzmann, physicists have considered posits regarding the initial state, or the global structure of the cosmos, as a source of potential answers. One prominent line of thought takes all of these asymmetries to be consequences of the "Past Hypothesis" (Albert 2000). The universe exhibits an arrow of time, on this view, because it currently has an entropy much lower than the maximum entropy. This is surprising because "typical" dynamical evolution tends to drive a system towards equilibrium. The present entropy is low because the entropy at even earlier times was lower, and there has not been sufficient time to reach equilibrium. Iterating this argument leads to the Past Hypothesis: the universe began in an initial state of extremely low entropy.[6]

Cosmologists have made similar posits for different reasons. Based on reflections regarding structure formation, Peebles (2020, p. 212) advocates a general principle: "our universe has been evolving from order to chaos." "Order" corresponds to a homogeneous and isotropic universe, with small departures that can be treated as growing modes. Note that there is a striking contrast between "order," heuristically identified with low entropy, in the case of gravity and other

---

[4] More precisely, for time-reversal invariant laws, for any trajectory that is a solution of the dynamics D, the time-reverse is also a solution. The "reverse" trajectory, $\Sigma(t) \to R[\Sigma(-t)]$, is obtained by taking instantaneous states in the opposite order, and acting with an operator R that "reverses" each of these states (e.g., by reversing the velocities). Whether GR is time reversal invariant is subtle. Locally there is a transformation that corresponds to time reversal. But in general there is not a natural global operation of time reversal that can be applied to the state of the universe (e.g., on a 3-dimensional spatial surface Σ), in order to specify what is meant by the "reversed" trajectory, and then to determine whether it is also a solution. For some classes of solutions with further structure, however, such as the FLRW models used in the ΛCDM model, there is a natural way to define time reversal and establish invariance.

[5] For an overview of related philosophical debates, see Price (1996) and North (2011).

[6] The Boltzmannian version of the second law of thermodynamics holds that if at $t = t_0$ the Boltzmann entropy $S_B(t_0)$ of the system is low compared to the equilibrium value, then at $t' = t_0 + \Delta t$, it is highly probable that $S_B(t') > S_B(t_0)$. Yet this holds regardless of the sign of Δt for time-reversal invariant dynamics. Hence entropy increases in *both directions* from t_0, apparently implying disastrous retrodictions to earlier times – e.g., the coffee cup cools in both directions from t_0, and the Past Hypothesis must apply to the *initial* state. The initial state can be taken to be specified at the boundary of the domain of applicability of current theories; see the discussion below.

domains of physics. We would usually associate a thermalized, uniform state, such as that observed via the CMB, with *high* entropy. The reverse is true once gravity (or any long-range, unscreened attractive force) is included: high entropy states correspond to maximally clumpy states. Taking GR into account, Penrose (1979) argued that black holes maximize the entropy for gravitational degrees of freedom, and estimated the maximum entropy based on black hole thermodynamics. Peebles's evolution from order to chaos then exemplifies thermodynamic asymmetry, going from an extremely low entropy initial state to the presently observed "chaos," with higher—but still far from maximum—entropy.

Peebles presents the principle as a reasonable generalization, motivated by the phenomenological success of structure formation models that satisfy it: treating the early universe as highly ordered in this sense has been a crucial part of successful theorizing in cosmology over the last half century. This success provides some level of empirical justification for the assumption that the universe began in a highly ordered initial state, which would plausibly have low entropy. Before turning to further questions about how satisfying this justification is, a pause to acknowledge several foundational challenges that need to be met in order to even formulate the assumption should be made.

Begin by considering the problem Peebles had in mind, the growth of small density contrasts due to gravity. This case will serve to illustrate two subtle points about such asymmetric processes. Starting from a nearly homogeneous distribution of matter with a slight density contrast, described with Newtonian gravity, it would be expected that the density contrast increases — because the rate of gravitational collapse is higher in regions with higher density. This expectation is supported by gravitational clumping observed in self-gravitating systems at various scales. Since Newtonian gravity is invariant under time reversal, however, it is also possible for an initially clumpy matter distribution to expand and smooth out. (The general solution includes both growing and decaying modes.)

To resolve this apparent conflict, it must first be acknowledged that the description of gravitational clumping reflects the choice of a macroscopic variable — namely, fluctuations away from average density. This amounts to coarse-graining the space of states of the system, lumping together a large number of microscopic states (full specifications of the distribution of matter and velocities) that have the same macroscopic observables. It is only with respect to some chosen coarse-graining that the definition of the Boltzmann entropy S_B, as proportional to the volume of phase space corresponding to the macro-state, can be attempted. Expectations and observations of gravitational clumping regard macroscopic variables, whereas the reversibility of the laws ensures the existence of a trajectory defined at the level of the microscopic states. Second, treatments of gravitational clumping yield time-asymmetric equations in terms of these macroscopic variables. Some initial states allowed by Newtonian gravity are incompatible with these equations: for example, there are solutions describing a pure "decaying mode," expanding and smoothing out an initially clumpy state. Such initial states require tightly correlated initial velocities, in order to "unwind" the initial clumps. The plausible step of ruling out such correlations is needed to support the time asymmetry in the macroscopic equations, absent in the microscopic dynamics.[7]

Despite increased complexity and several quantitative differences introduced by GR, this analysis of the source of time asymmetry carries over to perturbations in an expanding universe. Similar points apply to other temporal asymmetries described in terms of time-reversal invariant microscopic dynamics: the asymmetry arises at the level of coarse-grained variables, and only holds if initial or boundary conditions are imposed. The restriction on the initial state is essential to account for the asymmetry, but it can seem question-begging: apparently time-asymmetric evolution is explained by fiat, by ruling out states that would lead to the opposite behaviour.

These general points about the nature of temporal asymmetry are generally regarded as uncontroversial, but there are active debates about two further issues. First, how do the cosmological arrow of time, and the Past Hypothesis, relate to other observed asymmetries? Albert (2000) treats the Past Hypothesis as something like a transcendental condition for the possibility of investigating the past, which furthermore underwrites all the other arrows — even down to the asymmetries exhibited by quasi-isolated systems like cups of coffee.[8] Making the case in favor of this position requires linking the global constraint on the initial state to the features of the various other systems that exhibit temporal asymmetries, sufficient to explain, inter alia, why all quasi-isolated systems exhibit the same temporal orientation. Earman (2006) rejects the view of Albert (and other advocates of a similar view) as providing little more than a "just-so story" linking the Past Hypothesis to other asymmetries. But there are good prospects for an alternative account that provides such a linkage, based on a much more detailed understanding of how cosmological evolution relates to subsystems. Rovelli (2019), drawing on Wallace (2020), provides such an account, arguing that other temporal asymmetries track back to the fact that the rate of expansion of the universe starts far from its equilibrium value. Spelling out this connection requires considering how local systems fall out of equilibrium in an expanding universe, and are then trapped in metastable states of low entropy. For example, roughly three minutes "after" the big bang, deuterium would rapidly react with free neutrons to form various helium nuclei, as this is an entropically favorable reaction. Shortly after this, however, deuterium and other hydrogen isotopes were stuck in their relatively low-entropy state due to the rapid expansion of the universe; temperature and density of free neutrons rapidly decreased, leaving roughly the relative abundances of hydrogen and helium seen today. The role of the initial state in these explanations is to

---

[7] This is analogous to Boltzmann's *Stoßzahlansatz*: his derivation of a macroscopic transport equation for a dilute gas required that initially the velocities of the gas's molecules were randomly distributed and uncorrelated. See Uffink (2006) for a comprehensive overview of Boltzmann's contributions and the foundations of statistical physics more broadly.

[8] See Wallace (2011) for a critical assessment and alternative account of how to understand the nature of the initial state, which is drawn on here. Note further that the Past Hypothesis has a specific role in Albert's project; as Myrvold (2021, §8.7) emphasizes, the Past Hypothesis is needed to correct faulty retrodictions based on over-extending a statistical postulate appropriate only for systems at equilibrium.

provide the out of equilibrium, “special” expansion rate that limits equilibration time for subsystems.

The second issue concerns the propriety of invoking thermodynamic concepts such as entropy in formulating the Past Hypothesis and establishing its connections with other temporal asymmetries. The account of asymmetry sketched above can arguably be given entirely in terms of statistical mechanics without relying on thermodynamic concepts. This is a virtue given the difficulty of defining a sensible notion of gravitational entropy. Astrophysicists have developed sophisticated treatments of the statistical mechanics of self-gravitating systems, extending well beyond the simple case of linear perturbations to a homogeneous background. But it is unclear how to apply thermodynamics to such systems, due to distinctive features of gravity. For example, adding energy to a self-gravitating system reduces the temperature (negative heat capacity), and more fundamentally the gravitational interaction fails to satisfy the conditions necessary to establish the existence of the thermodynamic limit. These problems arise in attempts to apply thermodynamics, and to define quantities such as S_B, to self-gravitating subsystems of the universe, such as globular clusters.

Further obstacles arise with respect to the “entropy of the whole universe.” Thermodynamics depends on distinctions that only seem applicable to subsystems of the universe, in particular the contrast between work (energy that can be utilized as a resource) and heat (Myrvold 2016). Hence it is unclear how to introduce thermodynamic concepts that apply globally. A distinct set of issues arises due to features of GR: it lacks the “time translation invariance” usually assumed in statistical physics, and there is no generally accepted definition of the entropy of the gravitational field. On a more optimistic note, the success of black hole thermodynamics suggests that a suitable notion of entropy exists (e.g., Wald 1994), although insights from quantum gravity will be needed to identify the fundamental degrees of freedom underlying these concepts. Penrose’s argument noted above relies on these results to provide heuristic estimates of the entropy of the early universe. Thermodynamic concepts may apply reliably in some regimes even if no general definition can be found, based for example on the entropy associated with surface areas of black holes or more generally with causal horizons.

The above suggests that much of the analysis of temporal asymmetry can be based on statistical mechanics without the need to employ problematic thermodynamic concepts. This approach is defended in broader terms by Robertson (2019), who argues that the statistical mechanics of self-gravitating systems does not depend on the emergence of thermodynamics. Callender (2010) and Callender (2011) also provide a clear overview of the relevant foundational issues and reconsiders the nature of thermodynamics and its relation to other physical theories.

Finally, waiving these concerns about the precise formulation of the Past Hypothesis, there are debates regarding its status. Despite widespread consensus that the Past Hypothesis, or at least a constraint on boundary conditions with a similar character, plays a role in accounting for temporal asymmetries, there is little agreement regarding the propriety of the further demand to explain or justify the Past Hypothesis itself. Those who see these demands as pressing usually start by emphasizing just how incredibly special the initial state must be: according to Penrose, the initial state has a volume of 1 part in $10^{10^{123}}$ of the entire phase space volume.[9] Penrose often includes an illustration, of a bearded Creator choosing this state within the phase space with exquisite precision, intended to highlight the incredibly low probability that the initial state could be chosen "at random." The illustration implicitly acknowledges the kind of speculative hypothesis that would be needed to associate probability with phase space volume in this case.

There are essentially four options regarding the status of the Past Hypothesis (see also Albrecht 2004; Callender 2004; Price 2004):

1. *Empiricist Denial*: Physical explanations all have explanatory stopping points, brute facts that are not a fruitful target for further work. This general point is often bolstered by criticisms of positive arguments meant to distinguish the initial state, such as the one based on "improbability" (Callender 2004).

2. *Law-like Constraint*: The constraint on the initial state is a law-like feature of the universe. The view comes in two versions. On the first, it qualifies as a "law" based on carefully attending to what features a "law of nature" should have. In the Lewisian best systems account, for example, it plausibly features as one of the basic assumptions in an axiomatic system that optimizes strength and simplicity, the defining feature of a law (Albert 2000; Loewer 2012). On the second, the Past Hypothesis is expected to follow as a consequence of quantum gravity, as in Penrose (1979)'s Weyl Curvature Hypothesis.

3. *Multiverse plus Selection*: This view postulates that on extremely large scales, the features of the universe differ dramatically from our observed region, and vary from one region to another. If we select regions of this vast multiverse that resemble our observations, it is claimed that they will with high probability have evolved from a low entropy initial state. Proponents of this view need to counter the obvious rejoinder that it is much cheaper, in terms of entropy, to recover our observations with a fluctuation somewhat smaller than the entire observed universe. Such a view is discussed, though not endorsed, in Albrecht (2004); see W. Kinney (2022) for a recent defense.

4. *Dynamics*: On this view, the need to postulate a special initial state provides a target for theorizing. A more successful theory of the origins of the universe should demonstrate that an initial "chaos," or "generic" initial state, would lead to something compatible with observations due to subsequent dynamical evolution. This is largely the mainstream methodological

---

[9] This estimate is based on the heuristic mentioned above, namely calculating total entropy or phase space volume based on black hole thermodynamics; see, e.g., Penrose (2004, §27.13).

strategy in early universe cosmology, and finds different proponents working on inflation (including eternal inflation) and quantum gravity (Guth 2007; Linde 2015).

As discussed in the following sections, the fourth option has been the dominant approach, at least in sociological terms, in early universe cosmology.

3 Singularity Resolution and Quantum Gravity

3.1 Singularities in General Relativity

ΛCDM describes the universe, at large scales, using the simple Friedmann-Lemaître-Robertson-Walker (FLRW) models. These models describe geometrically uniform expanding universes, composed of a sequence of spacelike surfaces of constant curvature $\Sigma(t)$ indexed by the cosmic time t. The dynamics of GR reduces to two ordinary differential equations for a quantity called the scale factor $R(t)$, governing how these models evolve based on the types of matter and energy present. The scale factor represents the change in spatial distance between nearby observers who flow with the cosmic expansion; the cosmic time t corresponds to the proper time measured by these observers. Ordinary matter leads to decreasing expansion (with $\ddot{R} < 0$), reflecting the fact that gravity is a force of attraction.

One particularly striking feature of these models follows directly from the dynamics: for an expanding model filled with ordinary matter and radiation, there is an upper limit to how far the cosmic time can be extended back from the present moment. Physical quantities such as matter density and curvature invariants blow up as $t \to 0$; in other words, there is a singularity at a finite time to the past in these models.[10] This does not imply that there is actually a “first instant”: although the density increases without bound as t is rolled back, the “region of infinite density” cannot be a part of a classical relativistic spacetime. Trying to identify a “first instant” close to this region makes as much sense as choosing the real number closest to zero. Hence the discussion of the “initial state” in the previous section either reflects an arbitrary choice, or follows from some other considerations – such as an estimate of the limits of applicability of classical GR.

Physicists usually take the existence of singularities to signal the breakdown of GR. In the early days of relativistic cosmology, singularities could be plausibly dismissed as artifacts of unphysical idealizations, such as the symmetries of the FLRW models. Several theorems from the 1960s, however established that singularities arise generically under much weaker, and physically plausible, conditions (see Earman 1995 for a philosophical overview). On this approach, a spacetime is “singular” if it is geodesically incomplete — that is, it includes curves that have (roughly speaking) finite length, and it cannot be embedded

---

[10] Spacetime curvature in GR is represented by the Riemann curvature tensor; curvature invariants are scalar quantities constructed from the Riemann curvature tensor that can be used to characterize spacetime geometry.

in a larger spacetime. (The second condition rules out the possibility that the incompleteness reflects a mistake in taking a "small" spacetime, such that the missing points of incomplete geodesics can be added back in by extending it.) While this definition proved to be extremely fruitful for theorem proving, there have been extended debates regarding what singularities (in this sense) reveal about the limits of GR and possible successor theories. One source of difficulty is the divergence between what can be proven mathematically and definitions that have clear physical implications. In particular, in the simple case of the FLRW models, along every timelike geodesic extending towards the past, for matter with physically reasonable equations of state, the density and curvature invariants increase without bound within finite time. The mathematical structures needed to provide a "local" analysis of singularities in a wider class of cases have, however, various pathological features. After reviewing various attempts along these lines, Earman (1995) argues that the idea of a "localizable" singular region should be dropped, and replaced with the idea that a given spacetime as a whole qualifies as "singular" based on global features.

Based on this assessment, Earman challenges the conventional wisdom in physics that singularities clearly signal that GR is incomplete, or breaks down. In many singular spacetimes, there is no way of making precise what it would mean to "approach a singular region." This problem in part stems from considering the complexity of causal structures and anomalies that can be appear in GR (see also Geroch, Can-bin, and Wald 1982; Curiel 1999). Furthermore, other features of relativistic spacetimes may indicate the limits of GR. For example, in some cases a reasonable "initial state" (on an analog of a surface $\Sigma(t)$ in an FLRW model) does not determine, through the dynamics, the physical state throughout the entire spacetime.[11] Such a failure of deterministic evolution is often taken to indicate that a theory breaks down, but such failures need not stem from singularities.

If the clearest characterization of singularities is as a global feature of a spacetime—defined in terms of singular spacetimes, rather than singularities located within a specific region—how can quantum gravity help resolve singularities? If a theory of quantum gravity is expected to resolve all singularities or causal anomalies in GR, this would seem to require a more radical change to our understanding of global properties of spacetimes. The more naive approach of delimiting regions of spacetime "near" singularities where GR is expected to break down is at best limited in scope, and at worst must be entirely rethought. The former option is typically taken in response to the challenge of creating a local definition for singularities. When asked to explain *why* one should expect quantum gravity to resolve singularities, the response really only applies to restricted classes of curvature singularities: in the presence of hot, dense matter, gravitational fields become so strong that one often finds divergence of curvature invariants. But this is precisely the regime where one would expect GR to break down and quantum gravity to take over. Given such a restricted

---

[11] These spacetimes have a Cauchy horizon, a null surface which marks the boundary of the region whose physical state is determined via the laws given the initial state.

focus, it would then be possible to assess the implications of candidate theories of quantum gravity for specific types of singular behavior.

3.2 Emergence of Spacetime

Several different approaches to quantum gravity suggest that spacetime "disappears" – roughly speaking, that the fundamental degrees of freedom the theory describes cannot be characterized as spatio-temporal quantities. There is also a widespread expectation that the singular behaviour evident in classical GR may be resolved in this new description. For example, quantum observables that correspond to important classical quantities, such as the scale factor or spatial volume, may remain bounded rather than diverging. This would be particularly satisfying if the theory also clarified how the fundamental degrees of freedom relate to the classical picture, and explained where and why the classical picture breaks down. Yet one of the main open technical challenges facing several approaches to quantum gravity is precisely the problem of how to recover classical GR as an appropriate limit. Furthermore, as emphasized above, whether singularities have been successfully "resolved" depends on what type of singular behaviour is targeted.

Any merger of quantum theory and gravity serves to evade many of the singularity theorems, in one relatively straightforward sense, because matter described by quantum fields does not obey the standard energy conditions that are assumed in their proofs. At the semiclassical level, then, features of the quantum fields may lead the collapsing matter to "bounce," and prevent the formation of singularities. However, this type of evasion is not the same as having a fully quantum treatment of spacetime that resolves physically significant singularities, such as those in black holes or at the big bang, by giving an alternative physical account that is entirely non-singular. In a full theory of quantum gravity, many physicists expect — or even demand — that the restricted class of localizable curvature singularities will be resolved by some underlying quantum description that is non-spatiotemporal. Classical spacetime as described in general relativity would be recovered as an appropriate limit in some physical domains, but not in the quantum gravity regime "near" a singularity.

Recovering spacetime from a theory which lacks fundamental spacetime degrees of freedom raises several interesting philosophical questions alongside the technical challenges associated with this limit. First, how is it even possible to gather evidence for a theory regarding fundamentally non-spatiotemporal quantities (Huggett and Wüthrich 2013a)? Measurements are always localized in some region of spacetime, which suggests that fundamentally non-spatiotemporal quantities cannot be measured. Huggett and Wüthrich reject this claim, and emphasize that most approaches to quantum gravity have principled ways of bridging the gap between the proposed fundamental degrees of freedom and familiar spatio-temporal concepts. Recovery of a spacetime structure and the presence of locally defined quatities ("local beables") are hurdles that can be overcome, in at least some domains. If singularities are resolved in a theory of quantum gravity through the introduction of fundamentally non-spatiotemporal degrees of freedom, it might not be possible to bridge the gap "near" a singularity in spatiotemporal terms. A classical spacetime description would have a status similar to, for example, a classical description of an electron orbiting a nucleus; there is a breakdown of the description as the electron can fall directly into the nucleus. In that sense, singularities in GR should be regarded as indicating the breakdown of classical spacetime concepts.

Second, philosophers have tried to make sense of spacetime emerging from something that is fundamentally non-spatiotemporal, when the very concepts of causation, change, and transition seem to depend on at least some notion of temporal evolution (see Huggett and Wüthrich's (2021) upcoming book on spacetime emergence in quantum gravity). This leads to more detailed analyses of how one can obtain a meaningful form of emergence, with classical gravity emerging at a higher level from a fundamental theory that dispenses with classical spacetime. How this account of emergence works in detail differs among candidate theories of quantum gravity.

It will be useful to give a relatively theory-neutral classification scheme to demarcate different senses of spacetime emergence, following Oriti (2018). Oriti distinguishes three different levels at which classical general relativistic spacetime can be said to emerge from a theory of quantum gravity. At the lowest level (Level 0), the gravitational degrees of freedom are quantized, and classical spacetime "emerges" as the classical limit of these quantum degrees of freedom. Classical spacetime would emerge in much the same sense as classical trajectories emerge from a quantum description of particles. Oriti's next step, Level 1, involves new non-geometric degrees of freedom, upon which the concepts of classical spacetime supervene. This differs from the lowest level in that there is not a clear correspondence between the classical degrees of freedom and the more fundamental description. Such a relation is characterized in metaphysics as a supervenience relation: a directed relation of ontological dependence. The classical spacetime structure supervenes on the quantum level, but the quantum level is not itself (fully) geometric. One example of such a relation would be spacetime supervening on spin networks from loop quantum gravity (Bianchi, Rovelli, and Vidotto 2010). Changes in the spacetime structure depend on changes to the spin networks, and a change in one level happens if and only if a change at the other level also happens. Level 2 introduces the possibility that the underlying degrees of freedom have both geometric and non-geometric phases. In the geometric phase spacetime concepts can apply, but in a non-geometric phase there is no way to recover classical spacetime ideas. In that phase, it would be simply inappropriate to describe the system using any spacetime concepts at all. Oriti characterizes Level 3 as involving one final step, namely acknowledging the possibility of a phase transition from a non-geometric to a geometric phase.

There is another dimension along which we can distinguish two different types of emergence (Crowther 2020). First, there is *synchronic* emergence: a relationship that holds between a classical continuum spacetime and more fundamental, non-spatiotemporal "atoms," in a description of one and the same system at the same "time". But, second, there is also the possibility of *di-*

achronic emergence if a theory of quantum gravity suggests some form of transition from a non-spatiotemporal phase to a phase that can be described with spacetime concepts (as in Oriti's level 3). Obviously it is difficult to characterize a "transition" between distinct phases without appealing to a concept of time. Approaches like loop quantum cosmology solve this problem by introducing a scalar clock field into the model, so that time is relational and parameterized by the field (Ashtekar and Singh 2011).

If spacetime emergence from a theory of quantum gravity is more radical—Level 2 or Level 3 emergence in Oriti's hierarchy—then perhaps a more radical global redefinition of spacetimes and of (at least some) singularities should be expected.[12] Such a theory of quantum gravity may imply that only certain subsets of GR spacetimes can emerge from the quantum theory, and that these plus the quantum matter conditions preclude the existence of singularities. This would be a more radical form of emergence, as it would force a revision of ideas about what sorts of spacetimes are physically possible.

The potential for quantum gravity to result in an emergent spacetime highlights some of the important methodological challenges faced in the pursuit of new physical theories. Philosophers talk of reduction in science in slightly different terms than physicists. While it is common in physics to speak of, for example, GR *reducing* to Newtonian gravity in the weak field limit, philosophers often talk of an Newtonian gravity *reducing* to GR, since the latter is the more fundamental theory. Reduction in the philosopher's sense is a directed relationship from a higher-level approximate theory to a lower-level, more fundamental theory. Emergence is a directed relationship going the other way; phenomena *emerge* in a higher-level theory out of more complex relationships in the fundamental theory. Thus, spacetime would emerge from a theory of quantum gravity, while GR—and perhaps particle physics—would reduce to a theory of quantum gravity. For some philosophers, emergence is only possible when there is no reduction (Batterman 2011), though it has become more common in the philosophy of physics to articulate forms of emergence that are compatible with reduction (Butterfield 2011). However, recent work in the philosophy of quantum gravity has tried to disentangle emergence and reduction. Crowther (2015; 2018) has argued that emergence of spacetime is not a useful methodological heuristic for constructing a theory of quantum gravity. Rather, the focus should be on the intertheoretic limiting relations that classical GR to be recovered as an approximation to quantum gravity in the appropriate limits. Only once a strong candidate theory of quantum gravity exists can it be examined to see if and how spacetime emerges from it. Methodologically, reduction seems like the better choice for guiding the construction of a theory of quantum gravity. Once that is in place, the theory itself can be examined to help clarify the ways (if any) in which classical spacetime is emergent from some non-geometric quantum structure.

Singularity resolution also has implications for the initial state of the uni-

---

[12] For more philosophical work on the different notions of emergence of spacetime see (Huggett and Wüthrich 2013a; Huggett and Wüthrich 2013b) and the articles contained in the latter special issue.

verse. Assuming that at least curvature singularities get resolved by a theory of quantum gravity, this has direct implications for the ΛCDM model of our universe. For example, if the "initial" state (in the spatiotemporal regime) arises from a pre-existing non-spatiotemporal state, this might eliminate the need for something like the past hypothesis to set the initial conditions by fiat. Ideally, the theory of quantum gravity will constrain the initial state to be in some sense determined by the underlying microphysics. It may be the outcome of a transition from a non-spatiotemporal phase. Huggett and Wüthrich (2018) discuss various ways to describe a "transition" from a non-spatiotemporal state to a spatiotemporal one in quantum gravity. They argue that, in cases where the degree of emergence is not too drastic (i.e., not the geometrogenesis of Oriti's Level 3), classical cosmic time can be extended beyond its proper domain, making it possible to speak loosely of the "duration" of the non-spatiotemporal phase and its transition to classical spacetime. The notions of "transition" and "emergence" may break down entirely in more radical theories of quantum gravity. Evolution might then be parameterized with respect to some internal degrees of freedom, whose identification with a cosmic clock may only hold in some phases.

Exactly what to make of a cosmology without a big bang depends on how quantum gravity resolves the singularity. In loop quantum cosmology, for example, the region of cosmic evolution where the cube of the scale factor is smaller than the minimal volume is replaced by a short quantum transition. In this region, the concepts of space and time break down. Continuing through to the "other side" of the quantum region reveals a symmetric contracting universe phase. In loop quantum cosmology, therefore, cosmic time can be thoguht of as extending back indefinitely, such that there is no initial state. Alternatively, since the entropic arrow of time points "backwards" on the other side of the quantum transition, a cosmic bounce may be considered to actually consist of a single quantum creation event, out of which two universes emerge with opposite arrows of time. In either case, the proposed quantum gravity theory provides more insight into the state of the universe at minimal volume, and avoids the initial big bang singularity. Though it may not make sense any longer to speak of the initial state, there is a natural choice for a time at which to impose the past hypothesis: the cosmic time at which the volume of the universe is minimized, while classical concepts of space and time still apply.

The idea of spacetime emergence in string theory is less concrete, but there are suggestions to indicate that, in the aspirational final form of string theory, space and time emerge from the underlying stringy structure (Vistarini 2019). Further, from the point of view of string theory, effective field theory methods are expected to break down due to nonlocalities present in string interactions, as well as the increased relevance of higher dimensions and new symmetries like dualities. The search for singularity resolution in cosmology ties the big bang singularity together with divergences in effective field theory as evidence pointing to new physics, including the emergence of spacetime from fundamental string dynamics (Brandenberger 2021; Koberinski and Smeenk 2023). However, this is currently more of a promissory note, since exact, non-perturbative formulations of string theory are still not well understood. Brandenberger's (2008) attempt to use string theory to eliminate an initial big bang singularity, for example, involves an approximate description with a classical spacetime background. In this string gas cosmology, there is a phase transition from a constant temperature phase to an expanding phase, but all of these remain within the realm of classical spacetime concepts. As with loop quantum cosmology, however, the choice of initial state might naturally be taken to be the time at which the phase transition occurs and expansion begins, where cosmic time becomes relevant.

Philosophical accounts of change, emergence, and time are set to play a central role in quantum gravity in the early universe. It is likely that quantum gravity completely revolutionizes the notion of an initial state, either by a breakdown of the idea of "initial" state, or by setting an unambiguous starting point where temporal concepts apply. In either case, the initial conditions assumed in something like a past hypothesis might themselves be subject to a generalized form of dynamical explanation supplied by quantum gravity.

4 Beyond Standard Model Effects

The discussion above has considered "initial state" of the universe, and the impact that a theory of quantum gravity might have on how it is treated. These are areas where unification of thermodynamics, quantum theory, and gravity in the early universe would be expected. However, evolution from the quantum gravity epoch until the stage when known physics can be used to describe the early universe spans a large dynamic range. According to the current ΛCDM model, known physics from the Standard Model applies at about 10^{-12} seconds after the big bang with the electroweak phase transition, well after the expected limit of the quantum gravity epoch at 10^{-43} seconds, or one Planck time, after the big bang. The field of early universe cosmology has explored a variety of proposals based on beyond Standard Model physics that have a distinctive impact during this period. Here, two of the most widely studied scenarios, inflationary cosmology and baryogenesis, will be discussed. According to inflation, the universe underwent a transient period of rapid (exponential) expansion after the quantum gravity epoch (anywhere from $10^{-42} - -10^{-35}$ seconds, depending on the model). Following inflation, baryogenesis occurs (the expected time is model-dependent), generating a large excess of matter over antimatter. After the electroweak transition, more familiar physics takes over, including a quark epoch (dominated by the dynamics of a quark-gluon plasma), neutrino decoupling, nucleosynthesis, and recombination.[13]

The very early universe provides an appealing target for theorists for two quite different reasons. First, from the point of view of particle physics, rewinding back toward the initial state of the universe leads to higher temperatures and therefore higher energies for interactions. For (relatively) late times, when the expected interactions can be described by the Standard Model, the early

---

[13] See Dodelson and Schmidt (2020) for a textbook treatment of the thermal history of the universe, with more detailed descriptions of these and later phases.

universe provides an opportunity to test novel effects; for earlier times the interaction energies exceed the regime where the Standard Model applies. Furthermore, symmetries are expected to be restored at higher energies, leading to successive unification of the fundamental forces at early times. This connects with the idea of the universe evolving from order to chaos (Peebles 2020). Proposed extensions of the Standard Model would be expected to have consequences for the early universe, and it is plausible to hope that these would include distinctive and robust predictions. Second, physicists started to take the account of the early universe provided by cosmologists much more seriously by the late 70s, due to the successes of the big bang model and the prospects of further precision measurements of the cosmic microwave background radiation (CMBR). Cosmologists provided an increasingly clear account of the early state of the universe, raising the prospects for compelling observational constraints. But this work also raised further explanatory questions: in particular, why did the universe have precisely the features revealed by cosmological observations?

The dominant approach to early universe theorizing has been to explain as much as possible in terms of the dynamical evolution of the universe, such that the initial state of the universe can be treated as "simple" or "generic". Any features that are in some sense special stand out as targets for explanation. This is in contrast with expectations of the initial state from thermodynamic considerations discussed above. Despite this tension, the goal of finding dynamical explanations for special features of the universe is undeniably appealing. Both inflation and baryogenesis exemplify this appeal. Inflation explains several large-scale properties of the early universe — such as its flatness and uniformity — as the outcomes of a phase of dynamical evolution, and accounts of baryogenesis similarly propose dynamical mechanisms to account for another seemingly special feature, the observed asymmetry of matter and antimatter.

Both proposals face a similar methodological challenge: that cosmological evidence may not be sufficient to test or confirm the details of the proposed evolutionary stage. For later stages of the universe's evolution, such as big bang nucleosynthesis, the relevant nuclear physics can be grounded directly in a wide variety of experimental tests. There is then no possibility that, for example, the predicted pattern of light element abundances (as a function of cosmological parameters such as matter density) can be adjusted to match observations by changing nuclear physics. By contrast, both inflation and baryogenesis require beyond Standard Model physics that is weakly constrained by non-cosmological experiments, and the cosmological evidence — in the case of baryogenesis, essentially a single parameter — is not sufficient to compensate. Currently work in both areas has produced a number of different toy models, based on new physics, that give different descriptions of early universe dynamics. Toy models that are compatible with observational constraints, while an important achievement, provide at best "how possibly" explanations of the relevant features of the early universe. The challenge is to justify treating one of these toy models as providing an accurate account of how the universe *actually* evolved. In other words, how can the worry that a toy model should be regarded as a just-so story, in which the novel physics has been adjusted to match (sparse) observations, be countered? This kind of skeptical challenge has arguably been met effectively in other areas of physics (e.g. Smith 2014), based on the ability of a physical theory to support a line of inquiry that leads to the identification of new features that can be independently checked. Explaining striking features of the early universe as consequences of accidental, contingent properties of the initial state cuts off further inquiry, but so does postulating a mechanism tuned to reproduce a single effect. Ideally, then, proposals for inflation or baryogenesis would be at least partially grounded in known physics, or would be part of a theory that makes several distinct predictions to be tested outside of the domain of cosmology. Unfortunately, as discussed next, the most appealing candidate models of this sort have been all but ruled out by lack of empirical evidence. The discussion starts with the case of baryogenesis.

4.1 Baryogenesis

The problem of baryogenesis is relatively easy to state: why is there so much matter—and so little antimatter—in the observable universe, when the underlying laws of known physics do not favour baryon production over antibaryon production? Baryon number is conserved in all known Standard Model particle interactions, and yet there is an enormous abundance of baryons over antibaryons in the universe. There are three ways to explain the observed baryon abundance: (1) the initial conditions of the universe were such that the baryon asymmetry was built in; (2) the current Standard Model of particle physics has the capacity to explain the asymmetry, in the form of a phase transition in the early universe; or (3) new physics beyond the Standard Model will give rise to baryon nonconserving interactions, possibly also in the form of phase transitions. As the name baryogenesis indicates, option (1) is rarely considered as a serious option. Further, as discussed above, most physicists prefer dynamical explanations of any striking features of the universe; retreating to initial conditions is seen by many to be a failure of explanation, despite the thermodynamic concern that the initial state will have to be special in being a low entropy state (§ 2). This leaves (2) and (3) as options, both of which have been seriously pursued over the last 35 years.

There are agreed upon criteria that any plausible baryogenesis model must satisfy, known as the Sakharov criteria, first posed by Sakharov (1967). They are three conditions that must jointly be in place for an excess production of baryons over antibaryons in the early universe. First, the mechanism must include an interaction that violates baryon number conservation. Without a baryon nonconserving interaction, no dynamical explanation of baryogenesis is possible. Second, the baryon nonconserving interaction must go out of equilibrium at some time. Assuming that the model has CPT invariance, a state in thermal equilibrium will have a balance of baryon generating and antibaryon generating interactions, so no global excess can be produced. Finally, there must be a mechanism in place in which C and CP symmetries are both violated. This is necessary to ensure that baryon production is biased over antibaryon production. Since Sakharov, physicists pursuing solutions to baryogenesis have focused on showing explicitly how a given model satisfies the above criteria. Since baryogenesis must occur after inflation[14], the underlying assumption that the physics of baryogenesis is described in the language of quantum field theory is a reasonable one. However, the insistence on such a framework should be noted, as the solution criteria may rule out radically different alternative approaches.

Trends in popular explanatory strategies have shifted over time; early on, electroweak baryogenesis and grand unified theory (GUT) baryogenesis were the favoured approaches. Empirical constraints have made these accounts less plausible, and now new toy models are created with increasingly novel and independent mechanisms. The problems with toy models here are clear: since an explanation of baryogenesis only leads to the prediction of a single quantity—the baryon to photon ratio—toy models must be fit to a larger physical theory in a natural way. Without some coherence with other domains of physics, toy models can be tuned to provide the right ratio, so mechanisms built to explain baryogenesis must provide further testable consequences. Given the modern understanding of the Standard Model as an effective field theory, global symmetries like baryon number are expected to be generically broken at high energies, and as such are treated as merely accidental symmetries. While the effective field theory (EFT) perspective therefore makes baryon nonconservation less surprising, it does not provide any candidate mechanisms for generating the observed distribution of baryons, only an expectation that some high-energy effects should break accidental symmetries. The EFT framework can *accommodate* these effects, but it does not provide an explanation as to what high-energy physics leads to the asymmetry. This was the appeal of solutions from within the Standard Model or GUTs; since these theories make several other testable predictions, worries of tuning the models to get the desired result could be avoided.

The 1980s saw an increased interest in electroweak baryogenesis, which proved an active area of research into the late 1990s and is still a small area of research today (Kolb and Wolfram 1980; Bergerhoff and Wetterich 1998; Riotto and Trodden 1999; Cline 2006; Garbrecht 2020). The prospect of explaining the universe's excess baryon number from within the Standard Model is tempting, since the Standard Model is highly confirmed by numerous experiments. It would be appealing to use only known, established physics to explain the asymmetry (option (2) above). Baryogenesis would thus be strongly linked to well-established physics, and the explanation would serve to further extend the empirical domain of the Standard Model to just before the assumed electroweak epoch. The earlier strategies start with the idea that a global U(1) anomaly—a global symmetry present in the classical form of the Standard Model Lagrangian but broken by quantum corrections—allows for high energy baryon nonconserving interactions. The presence of this anomaly is a nonperturbative effect dependent on the details of the electroweak Lagrangian, and a solution involving the anomaly is called a sphaleron anomaly. The baryon nonconservation would therefore be within Standard Model physics, but only visible as

---

[14] As discussed below, inflation tends to drive particle density down due to rapid expansion; the observed particle content of the universe is then created in a process called reheating, in which the inflaton energy decays and sources energy in the other quantum fields.

a nonperturbative effect. Electroweak baryogenesis requires that the early universe electroweak phase transition was a first order phase transition, which is a discontinuous process. As pockets of the early universe break the electroweak symmetry, a transition wall would form separating regions of broken symmetry from those with the symmetry retained. The transition wall breaks the parity symmetry, and provides a barrier to equilibration of particles and antiparticles across the barrier. If the symmetry breaking is smooth, then there is no barrier to equilibrium, and no boundary for sphaleron transitions. The dynamics under a first-order phase transition lead to a buildup of quarks and leptons on the symmetry broken side, while antiquarks and antileptons build up on the symmetric side. Then sphalerons—occurring abundantly on the symmetric side, but scarce once symmetry is broken—wipe out the antiquarks and antileptons, leading to an abundance of matter over antimatter at the end of the phase transition.

While this would be an elegant solution to baryogenesis from the domain of known physics, it turns out that the observed Higgs mass is too large for the electroweak transition to be first order. With a Higgs mass greater than about 70 GeV, the electroweak gauge interactions near the critical temperature are too strong to allow for a first order transition. Lower bounds on the Higgs mass cast doubt on this approach beginning in the early 2000s. By 2012, measurements of the Higgs mass ruled out this form of Standard Model electroweak baryogenesis, indicating that new physics is required to account for baryogenesis. Some physicists have continued to pursue electroweak baryogenesis in the context of extensions of the Standard Model (e.g., supersymmetry, composite Higgs, etc., cf. Morrissey and Ramsey-Musolf (2012)), while others have pursued baryogenesis as an effect of entirely new physics.

GUTs have been studied since shortly after the construction of the Standard Model. They initially held promise to unify the strong and electroweak forces, explain baryogenesis in the early universe, and, once inflation was proposed, provided a framework in which to fit the inflaton. Baryon nonconserving interactions are a generic feature of large classes of possible GUTs, making them appealing for explaining several features of the early universe. The details vary depending on the model considered, but in broad strokes GUTs explain baryogenesis in the same way that the Standard Model explains nucleosynthesis (Riotto 1998; Riotto and Trodden 1999). In the very early universe, temperatures would be high enough for the non-gravitational forces to restore grand unified symmetries, and exotic high-energy reactions would be common. Each model requires an overproduction of baryons compared to antibaryons, assumed to be tied to the violation of baryon number conservation by GUT interactions. As the universe cools, the grand unified symmetry breaks, and exotic baryon production halts. Baryon production must occur after or near the end of the inflationary epoch, placing bounds on the energy scale of the phase transition. Since GUTs constitute new physics, a successful explanation of baryogenesis must be accompanied by further testable predictions, confirmation of which would establish the particular GUT as a likely successor to the Standard Model.

Like electroweak baryogenesis, GUT baryogenesis would be an explanation situated within a full theory. GUTs are similar to the Standard Model, in that they predict a multitude of fundamental and composite particles, as well as interaction strengths and cross sections. The GUT would provide a comprehensive framework on which to draw in order to add subleading physical effects to increase the precision of predictions, leading to a theory-guided research program. Of course, unlike electroweak baryogenesis, the GUT explanation would remain only a candidate explanation until further aspects of the GUT were confirmed, making this solution fit under option (3). Unfortunately for the prospects of this account, GUTs are also heavily empirically disfavoured. The major issue barring the acceptance and further pursuit of GUTs today is the lack of observed proton decay. GUTs generically predict proton decay, a process that has not been observed despite immense experimental effort. The lower bounds on proton half-life are in the area of 10^{34} years (Bajc et al. 2016), sufficient to rule out many GUTs and tightly constrain others.

Baryogenesis remains an open problem in cosmology. Though many toy models have been constructed, there is no widely accepted paradigm for a solution. The Sakharov criteria must be met, but there are numerous ways to do so, involving different physics at radically different energy scales. Some, like electroweak baryogenesis, could be supported by measurements at LHC energy scales. Others, like GUT baryogenesis, require indirect evidence due to the extremely high energy at which these effects become significant. Toy models designed to explain baryogenesis should be met with skepticism; since there is really only one free parameter in need of explanation, the explanation must be independently motivated by fitting with other aspects of beyond Standard Model physics. Though electroweak and GUT baryogenesis are empirically disfavoured, they exemplify what a fruitful methodology to for dynamical explanations of aspects of the early universe: explanations should come from one part of a fuller theory that unites several independent observations. There are several candidate models that might explain baryogenesis, though it is difficult to unify these into a simple parameterized form, since the explanations come from a wide range of potential theories beyond the Standard Model. Along with inflation and the cosmological constant problem, baryogenesis is one of a few topics that is known to require new physics beyond the Standard Model and GR, and is therefore an important litmus test for new candidate theories seeking to unify physics in the early universe.

4.2 Inflation

The ΛCDM model describes the universe using the extremely simple FLRW models. The spatial surfaces $\Sigma(t)$ (labeled by cosmic time, t) in these models are uniform in two different senses: from any given point, there is no geometrically preferred direction, and any two spatial locations within a given Σ are equivalent. With the discovery of the CMBR, cosmologists were shocked to discover that these uniform models apply, even to the early universe, with remarkably high accuracy. Perturbations have to be added to the models, in order to provide seeds that grow into larger structures through gravitational enhancement. Initial observations of the CMBR showed that it has a strikingly uniform temperature, but increasingly precise measurements eventually revealed temperature anisotropies on the order of 1 part in 10^5. These measurements constrain the properties of the initial spectrum of perturbations (see, e.g. Durrer 2020), such as their overall amplitude. Inflationary cosmology accounts for the uniformity and small temperature anistropies of the CMBR as the consequence of a phase of evolution driven by a scalar field (or fields) in the very early universe. The proposal is called "inflation" because it involves a phase of exponential growth (with the scale factor $R(t) \propto e^{\xi t}$) in place of the more sedate radiation-dominated evolution of the FLRW models (with $R(t) \propto t^{1/2}$).

Inflation has been a widely accepted part of cosmological theory for several decades. Yet questions similar to those posed regarding baryogenesis have been the focus of ongoing and persistent debates. In particular, why should the claim that a period of inflationary expansion occurred be believed? In other words, to what extent does inflation's compatibility with observations justify this claim? And, furthermore, what are the implications of accepting inflation — specifically, how does the physical source of the inflationary phase fit into a more complete account of high energy physics?

Guth (1981) provided an influential motivation for pursuing inflationary cosmology, based on its ability to resolve the horizon and flatness problems.[15] Horizons in cosmology delimit the range of possible causal interactions. In the FLRW models, there is a finite horizon distance d_h, and regions beyond this distance have not been in causal contact. A straightforward calculation shows that CMB photons coming from regions of the sky separated by more than about 1° are separated by a distance greater than d_h. How, then, could these regions have come to have the same physical properties, such as the same temperature to one part in 10^5? The flatness problem stems from a feature of the FLRW dynamics, the dynamical instability of the "flat" model (the FLRW model with critical density, whose spatial sections Σ have flat geometries). Any model that is not exactly flat evolves rapidly away from the flat model; the distance from the flat model grows as a function of cosmic time. As a result, the early universe had to be *extremely* close to the flat model to be compatible with the observed flatness at late times.

Both of these features seem deeply puzzling on the supposition that the universe began in an initially "chaotic" state, with each causally independent region having a physical state "randomly chosen" from some space of possibilities. Horizons prevent causal interactions from smoothing out the physical differences amongst these distinct patches, and randomly choosing a state sufficiently close to flatness at early times seems, intuitively, quite improbable. These problems arguably illustrate a deep conflict between this proposed understanding of the initial state and the use of the FLRW models. Prior to Guth, cosmologists re-

---

[15] Guth's characterization of the advantages of inflation, and several aspects of his proposed model, were original (Guth 1997), but the general idea of a false-vacuum dominated expansion phase had been explored by several others, in particular Starobinsky, prior to Guth (Smeenk 2005). McCoy (2015) also argues that the horizon and flatness problems are too poorly-defined to be said to be solved by inflation (see also Earman and Mosterin 1999; Smeenk 2014; Holman 2018).

garded uniformity and flatness as enigmas (Dicke and Peebles 1979), but Guth's proposed solution convinced the community to see them instead as the crucial explanatory targets for early universe theorizing while also proposing a dynamical explanation. A stage of inflationary expansion alters the horizon structure of the early universe, stretching the horizon distance d_h; for a sufficiently long inflationary phase, the entire observed universe falls within this distance. The inflationary phase also reverses the instability of the FLRW models responsible for the flatness problem: the departure from the flat model rapidly *decreases* during the inflationary phase. Inflation made it possible to reconcile the idea that the initial state is in some sense "generic" with the applicability of the FLRW models.

Guth's model was proposed in the context of a GUT, where the inflationary stage was thought to be driven by one of several new particles that would appear at the unification scale. Inflation would then exemplify a strong form of unification between particle physics and cosmology: the physical source of the inflationary stage that resolves cosmological puzzles would also be a specific field included in a full theory of beyond Standard Model physics. But this initial promise has failed to come to fruition (see also Zinkernagel 2002). As noted above, the failure to experimentally observe proton decay casts doubt on GUTs. This is hardly an objection to inflation as long as it can be incorporated in other proposals for beyond Standard Model physics. The obstacle to doing so arises from a further appealing feature of inflation: namely, it provides a dynamical mechanism to generate the perturbations that leave an imprint as temperature anistropies in the CMBR. The amplitude of the observed anisotropies constrains the features of the field driving inflation. This ruled out the proposed identification of the field driving inflation with specific candidate fields found in GUTs, and it appears to have quite distinctive features compared to fields in other extensions of the Standard Model. Cosmologists now routinely discuss toy models of an "inflaton" field (the field, or fields, driving the inflationary stage), and have explored a variety of ways in which the inflaton can be embedded within high energy physics. Over the last 40 years cosmologists have explored a vast space of hundreds of possible models (Martin, Ringeval, and Vennin 2014). With a few exceptions, in these models the inflaton field is treated as an independent component of beyond Standard Model physics. Inflation thus became a paradigm largely independent from any underlying theory.

Ongoing precision cosmological measurements, primarily of the CMB, can be used to evaluate what regions of the vast space of possible models are most plausible, and to constrain parameters for specific models. The fact that inflation has remained compatible with measurements is an important success, and the observations are discriminating enough to rule out models of structure formation in the early universe based on topological defects. Inflation is widely accepted based on this success in providing an account of the origin of density perturbations. But this empirical success has not led to the resolution of persistent foundational debates regarding inflation, of which three are briefly described here.[16]

The first concerns the status of the problems originally used to motivate inflation. The view regarding the initial state that generates these problems contrasts sharply with the discussion of the past hypothesis in Section 2 above. Inflation apparently eliminates the need for special initial conditions — pre-established harmony among different causally disconnected regions, near exact flatness, and coherent large-scale perturbations. On the one hand, thermodynamic arguments lead to the expectation that the initial state of the universe must have been one of extremely low entropy, and therefore also highly ordered. On the other hand, inflation promised to eliminate the need to explain striking features of the universe by recourse to initial conditions. The goal of such an explanatory project is to provide a mechanism that carries generic states to homogeneous, flat states consistent with current observations. One problem with such an explanatory project is that a state space and accompanying measure with respect to which states are "generic" or "special" must be specified. Furthermore, the same thermodynamic arguments that lead to the postulation of a special initial state to account for temporal asymmetry apply to inflation. Penrose, among others, has argued (e.g., Penrose 2004, §28.5) that the pre-inflationary state of the inflaton field has to be extremely "special" to ensure that an inflationary period takes place. See Albrecht (2004) for further discussion, and a defense of the consistency of inflation with a thermodynamic arrow of time.

Moreover, the extent to which inflation solves the fine-tuning problems depends on the status of the inflationary model itself. The simplest models of inflation are determined by specifying a function called the effective potential that appears in the Lagrangian of the inflaton field. If the effective potential is treated as a free function, then it is possible to generate (nearly) any specified cosmic expansion history. It is plausible to restrict consideration to models with effective potentials that are not "fine-tuned," to rule out matching the observations by design. However, as philosophers have noted, defining fine-tuning and assessing its further implications are both quite contentious. Diagnosing fine-tuning requires an underlying notion of a measure over values of numerical constants (or the space of possible effective potentials), along with some understanding of the source of "randomness" in the system (see White 2000; Colyvan, Garfield, and Priest 2005; Norton 2010). This becomes even more challenging when it comes to adjudicating whether inflation has simply shifted one type of fine-tuning (in the features of the big bang's initial state) into another type (regarding the inflaton potential and the initial state of the inflaton field), as critics have claimed (Earman and Mosterin 1999; Ijjas, Steinhardt, and Loeb 2013).

These criticisms only establish, at most, the flaws of one way of justifying inflation. The initial case for novel physical theories is often replaced by a

---

[16] One thing that is not discussed here is the problem of the quantum to classical transition of fluctuations in the inflationary accounts of structure formation. Some link this to the much broader issue of the measurement problem in quantum theory (Sudarsky 2014), though this framing is controversial.

more compelling justification as scientists develop a better understanding of the the theory and its consequences. This leads to the second foundational debate regarding inflation, which concerns the strength of an empirical case in favour of inflation and how this depends on an account linking inflation to high-energy particle physics. In the early days of inflation, it was plausible to hope that further work would show how inflation fits into a specific extension of the Standard Model. This would resolve questions about the physical source of the inflationary phase — by identifying the inflaton field with a scalar field in the particle physics model — and enable a wide variety of empirical constraints from different domains, including particle physics and cosmology, to be brought to bear. But, as with the models of baryogenesis described above, current inflationary modeling is much less constrained, and there are several different views regarding the status of inflation in response. One possibility is to regard the identification of the "inflaton" as an open physical problem, with the hope that progress in cosmology and particle physics will fulfill the initial promise of a unified account of early universe cosmology. An alternative view regards inflation as merely an effective field theory (EFT) (Cheung et al. 2008; Weinberg 2008b).

There are two senses in which inflation could be treated as an EFT. First, inflation could be an EFT in the minimal sense of an effective parameterization of some unknown high-energy physics. In this sense, the inflaton Lagrangian is a phenomenological Lagrangian, whose terms include all configurations of the scalar field and its derivatives, and whose coupling coefficients are parameterizations of unknown higher-energy physics. This captures the idea that some underlying physics can be appropriately modelled with enough free parameters provided by an EFT expansion of a scalar field on curved spacetime. Second, and usually meant in addition to the first, inflation could be an EFT in that it is properly decoupled from the details of higher-energy physics. This ensures a more robust sense in which the inflaton Lagrangian accurately captures the relevant physics in a circumscribed domain, since decoupling ensures the stability of the EFT and its insensitivity to unknown high-energy physics. Using the standard machinery of renormalization group analysis (Burgess 2020), the inflaton would decouple from higher-energy physics if all terms in the Lagrangian are either marginal or irrelevant at comparatively low energies.

The first, minimal sense provides a useful method of unifying many distinct models of single-field inflation into a single framework. It provides a relatively model-independent formalism for comparing different inflation proposals with cosmological observations. But if inflation is to be a stable part of early universe theory, it must also be an EFT in the latter sense as well. However, a term in the inflaton potential proportional to the square of the field (usually interpreted as a mass term) is relevant, and so couples sensitively to higher-energy physics. If inflation is more than a phenomenological parameterization of the early stages of the universe, then further arguments are needed to show that inflation appropriately decouples from higher-energy physics.

One view of inflation seems to undermine the sense in which inflation decouples from high-energy physics. Some have argued that inflation is a sort of

"Planck microscope" (Martin and Brandenberger 2001; Brandenberger 2014). Since tiny quantum fluctuations are squeezed and stretched during exponential expansion—until they freeze out on crossing the Hubble radius—sufficiently long inflation would lead to Planck scale fluctuations playing a role in structure formation. This sort of mode crossing undermines the idea that inflation can be treated as effectively decoupled from high-energy physics. An ad hoc EFT treatment of inflation can be maintained only by introducing a principle like the Trans-Planckian censorship conjecture (Bedroya et al. 2020). Either some hard momentum cutoff must be imposed on the evolution of fluctuations, or the inflationary epoch must be short enough such that Planck scale fluctuations don't grow large enough to influence structure formation. In either case, the restrictions are ad hoc from the point of view of current physics, and both violate Lorentz-invariance by introducing a momentum scale cutoff. If the EFT framework is abandoned, however, the Planck microscope nature of inflation provides a useful probe into physics at the Planck scale. As Brandenberger (2014) illustrates, some models of quantum cosmology predict deviations from standard inflation in the power spectra of polarization modes, for example. Thus, if inflation is viewed as a Planck microscope, then precision observations of the CMB might illuminate Planck scale physics.

A second, contrasting interpretation of inflation is that it provides a Planck scale eraser, effectively providing a mechanism for decoupling pre-inflationary physics from later stages of cosmological evolution (Guth 2000; Albrecht 2004). Rather than focusing on perturbations, the Planck eraser view emphasizes the tendency of inflation to take "generic" spacetime geometries to flat, homogeneous spacetimes with matter fields unexcited. Since the inflaton dominates evolution until reheating, many of the details of pre-inflationary physics are assumed to "wash out" during inflation. In order to quantify the degree to which inflation is an attractor to flat homogeneous spacetimes, we need an idea of the sample space of pre-inflationary spacetimes, as well as a measure over such a space. See, for example, Azhar and Kaiser (2018), who only require minimal assumptions about the pre-inflationary spacetime background. The plausibility of these assumptions, however, depends on the details of physics prior to inflation. Since the Planck eraser view implies that very little about pre-inflationary physics can be discovered, the prospect of such an assessment seems dim.

The contrast between inflation as a Planck microscope versus a Planck eraser is not a true dilemma; both points of view could be correct, for example, by applying to different degrees of freedom. It could be the case that, though most of the details of pre-inflationary physics are wiped out during inflation, the growth of Planck-scale fluctuations provide one window to quantum gravity effects. Then these details could provide indirect evidence for something like an initial state, where quantum gravity reduces to an appropriate classical spacetime picture.

A further question regarding inflation concerns whether it can be incorporated as part of a full theory of quantum gravity. There have been active debates regarding whether inflation is incompatible with string theory. Swampland conjectures have led some string theorists to conclude that observationally favoured models of inflation are incompatible with string theory (Vafa 2005; W. H. Kinney, Vagnozzi, and Visinelli 2019). If true, then there are two attitudes to take in response. First, if inflation is strongly confirmed observationally, then so much the worse for string theory. That string theory and inflation are incompatible might be accepted, and therefore alternatives to string theory within which inflation is incorporated sought. Alternatively, the swampland conjectures can be viewed as more general constraints on possible UV completions of effective field theories. In this case, the problem might lie with the formulation of inflation, and alternatives may be sought to reproduce its successful predictions for density perturbations.

As a standalone paradigm, inflation provides a successful phenomenological account of structure formation, but arguably falls short of a full physical theory. Without convincing further justification, inflation is an ill fit for the standard decoupling assumptions of EFT, and therefore inflation should be embedded in a full theory of high-energy physics. As a purely phenomenological account of structure formation, inflation falls victim to problems of transient underdetermination. The situation here is much like that of baryogenesis; though inflation shows the promise of predicting more than a single parameter value, there are a vast number of toy models that currently do not connect to a full high-energy theory. Parameters in a toy model can be tuned to fit the known data for inflation. Unless a model of inflation includes a wide range of otherwise unrelated physical predictions, none can be established as fact. Instead of using new data to test predictions from a small set of models, the data is often used to constrain the space of possible inflationary models and to fix the parameters appearing in specific models. This approach, of course, assumes that the space of inflationary models is already well-established. However, there are alternative early universe scenarios that do not involve a period of inflation.

Brandenberger (2014) has pointed out that any model which decouples the horizon from the Hubble radius, has cosmological scale perturbations that originate inside the Hubble radius at early times, squeezes fluctuations on super-Hubble scales, and has a mechanism for producing a (near) scale-invariant spectrum of curvature fluctuations at super-Hubble scales will entirely replicate the successes of inflation—including the horizon and flatness problems.[17] These are the major selling points of inflation. Until inflation finds a home as part of a theory of the early universe, competing explanations pose real threats to establishing that an inflationary phase occurred (Smeenk 2019) and these alternatives should be taken seriously as candidate explanations.

One example of a competitor to inflation is the class of bouncing universe cosmologies (Battefeld and Peter 2015). Several distinct bouncing models meet all of the requirements stated above, provided there is sufficient control over conditions at the bounce to allow continuation "through" the big bang. Inflation is widely accepted as the "better" account of structure formation, but until inflation is fit with a broader theoretical context it is challenging to decide

---

[17] The Hubble radius, $\frac{c}{H}$ where H is the Hubble constant, roughly measures the spatial distance over which light travels during the time in which the scale factor increases by e (see Dodelson and Schmidt 2020).

between inflation and a cosmic bounce as an explanation for the available data. From a philosophical standpoint, it is to be hoped that open questions posed by inflation and its competitors would lead to further opportunities for empirical testing, allowing for evidence to decide between the competitors. This should be seen as an opportunity to fill out our account of the early universe, and to learn more about the connection between inflation and higher-energy physics. At this point, the step from establishing that inflation and its rivals are compatible with the known data to conducting further probing tests has yet to be made. This latter step is necessary to eliminate toy models that are simply tuned to known data, and to move from research programs that are merely attractive and worthy of pursuit to those favoured by the evidence (Wolf and Duerr 2023). This is a common feature of all beyond Standard Model effects; when reconstructing the evolution of the universe at earlier and earlier times, evidential support is likely to get scarcer and scarcer. Tuning toy models to fit the limited data is dangerous; it is preferable to seek robust physical effects in a coherent theory of high-energy physics.

Finally, the third foundational problem, which has generated the most heated controversy, is examined: the implications of "eternal inflation." Above inflation has been described as a supplement to the standard big bang model: a brief burst of exponential expansion that sets the stage for subsequent evolution. Yet many cosmologists have argued that inflation implies a very different picture of the universe on (extremely) large scales, well outside the observable universe. On this view, most of the universe is still undergoing exponential expansion, even as regions like the observable universe transition out of the expansionary phase (Guth 2000; Aguirre 2007). Inflation then only leads to uniformity within each of these "bubbles." At much larger scales, the universe has widely varied physical conditions: "most" of the universe would consist of regions undergoing inflation, and it is usually supposed that low-energy physics varies considerably across the ensemble of bubbles. The low-energy physics results from a series of phase transition that occurs within the bubbles, and these are expected to unfold differently in distinct regions. This account results from extrapolating physics to describe the earliest phases of inflation. (Although it has to be admitted that it is rather loosely based, and in particular the source of variation in the laws in the different regions is usually not specified.) Based on these calculations, many cosmologists claim that "inflation is generically eternal"—if an inflationary stage occurs, it will inevitably generate the global features just described. This remains a source of debate, however, given that many critics regard these calculations as extending well beyond their domain of applicability.

Granting that inflation inevitably leads to eternal inflation poses important questions regarding testability (see, e.g., the essays in Carr 2007). Obviously most of the implications of eternal inflation will be forever unknown, given that only the observable universe can be accessed — which comprises, if the view is correct, just part of one bubble in the vast multiverse. Clearly, direct tests of eternal inflation based on its novel implications for the extremely large scale structure of the universe are impossible. Critics have argued that eternal inflation fails to be scientific because it fails to satisfy a criterion of "falsifiability" as a result. This argument seems overstated: scientists routinely accept the implications of scientific theories for things that cannot be directly observed based on their success in accounting for what can be observed. Recent work in philosophy has considered more liberal accounts of scientific confirmation to account for the fact that scientists appear to have high credence in ideas that cannot be supported in this way, but instead are justified based on features of the research program that generated them (Dawid 2013). Yet there is a further challenge posed by eternal inflation that is directly related to the line of argument above. In particular, Ijjas, Steinhardt, and Loeb (2013) and Ijjas, Steinhardt, and Loeb (2014) argue that eternal inflation undermines what was originally so appealing about inflation. Eternal inflation appears to predict that anything can happen, somewhere in the multiverse, and the theoretical understanding needed to assign probabilities to different possible outcomes is lacking (Smeenk 2014; Norton 2018). Guth, Kaiser, and Nomura (2014), for example, responds to this line of criticism. In the opinion of the authors, it points to the need for a more careful evaluation of the empirical case in favour of inflation among those who grant that inflation is generically eternal.

References

Aguirre, Anthony (2007). “Eternal Inflation, past and future”. In: *arXiv preprint arXiv:0712.0571.*

Albert, David Z. (2000). *Time and Chance.* Cambridge: Harvard University Press.

Albrecht, Andreas (2004). “Cosmic inflation and the arrow of time”. In: *Science and ultimate reality: Quantum theory, cosmology and complexity.* Cambridge University Press, pp. 363–401.

Ashtekar, Abhay and Parampreet Singh (2011). “Loop quantum cosmology: a status report”. In: *Classical and Quantum Gravity* 28.21, p. 213001.

Azhar, Feraz and Jeremy Butterfield (2017). “Scientific realism and primordial cosmology”. In: *The Routledge Handbook of Scientific Realism.* Routledge, pp. 304–320.

Azhar, Feraz and David I Kaiser (2018). “Flows into inflation: An effective field theory approach”. In: *Physical Review D* 98.6, p. 063515.

Bajc, Borut et al. (2016). “Threshold corrections to dimension-six proton decay operators in non-minimal SUSY SU (5) GUTs”. In: *Nuclear Physics B* 910, pp. 1–22.

Battefeld, Diana and Patrick Peter (2015). “A critical review of classical bouncing cosmologies”. In: *Physics Reports* 571, pp. 1–66.

Batterman, Robert W (2011). “Emergence, singularities, and symmetry breaking”. In: *Foundations of physics* 41, pp. 1031–1050.

Bedroya, Alek et al. (2020). “Trans-Planckian censorship and inflationary cosmology”. In: *Physical Review D* 101.10, p. 103502.

Bergerhoff, Bastian and Christof Wetterich (1998). “Electroweak phase transition in the early universe?” In: *Current Topics in Astrofundamental Physics: Primordial Cosmology.* Springer, pp. 211–240.

Bianchi, Eugenio and Carlo Rovelli (2010). “Why all these prejudices against a constant?” In: *arXiv preprint arXiv:1002.3966.*

Bianchi, Eugenio, Carlo Rovelli, and Francesca Vidotto (2010). “Towards spinfoam cosmology”. In: *Phys. Rev. D* 82 (8), p. 084035. DOI: `10.1103/PhysRevD.82.084035`.

Brandenberger, Robert H. (2008). *String Gas Cosmology.* DOI: `10.48550/ARXIV.0808.0746`. URL: `https://arxiv.org/abs/0808.0746`.

Brandenberger, Robert H. (2014). “Do we have a theory of early universe cosmology?” In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 46, pp. 109–121.

Brandenberger, Robert H. (2021). “Limitations of an Effective Field Theory Treatment of Early Universe Cosmology”. In: *arXiv preprint arXiv:2108.12743.*

Burgess, Cliff Peter (2020). *Introduction to effective field theory.* Cambridge: Cambridge University Press.

Butterfield, Jeremy (2011). “Emergence, reduction and supervenience: A varied landscape”. In: *Foundations of physics* 41, pp. 920–959.

Callender, Craig (2004). "Measures, explanations and the past: Should 'special' initial conditions be explained?" In: *British Journal for the Philosophy of Science* 55.2.

Callender, Craig (2010). "The past hypothesis meets gravity". In: *Time, chance, and reduction: Philosophical aspects of statistical mechanics*, pp. 34–58.

Callender, Craig (2011). "The past histories of molecules". In: *Probabilities in physics*, pp. 83–113.

Carr, Bernard (2007). *Universe or multiverse?* Cambridge: Cambridge University Press.

Chamcham, Khalil et al. (2017). *The Philosophy of Cosmology*. Cambridge University Press.

Cheung, Clifford et al. (2008). "The effective field theory of inflation". In: *Journal of High Energy Physics* 2008.03, p. 014.

Cline, James M. (2006). *Baryogenesis*. arXiv: `hep-ph/0609145 [hep-ph]`.

Colyvan, Mark, Jay L Garfield, and Graham Priest (2005). "Problems with the argument from fine tuning". In: *Synthese* 145.3, pp. 325–338.

Crowther, Karen (2015). "Decoupling emergence and reduction in physics". In: *European Journal for Philosophy of Science* 5.3, pp. 419–445.

Crowther, Karen (2018). "Inter-theory relations in quantum gravity: Correspondence, reduction, and emergence". In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 63, pp. 74–85.

Crowther, Karen (2020). "As below, so before:'synchronic' and 'diachronic' conceptions of spacetime emergence". In: *Synthese*, pp. 1–29.

Curiel, Erik (1999). "The analysis of singular spacetimes". In: *Philosophy of Science* 66, S119–S145.

Dawid, Richard (2013). *String theory and the scientific method*. Cambridge University Press.

Dicke, RH and Phillip James Edwin Peebles (1979). "The big bang cosmology-enigmas and nostrums." In: *General Relativity: An Einstein centenary survey*. Ed. by Stephen Hawking and Werner Israel. Cambridge: Cambridge University Press, pp. 504–517.

Dodelson, Scott and Fabian Schmidt (2020). *Modern cosmology*. Academic press.

Durrer, Ruth (2020). *The cosmic microwave background*. Cambridge: Cambridge University Press.

Earman, John (1995). *Bangs, crunches, whimpers, and shrieks: Singularities and acausalities in relativistic spacetimes*. Oxford University Press.

Earman, John (2006). "The 'Past Hypothesis': Not even False". In: *Studies in History and Philosophy of Modern Physics* 37.3, pp. 399–430.

Earman, John and Jesus Mosterin (1999). "A critical look at inflationary cosmology". In: *Philosophy of Science* 66.1, pp. 1–49.

Ellis, George Francis Rayner (2008). "Cosmology: patchy solutions". In: *Nature* 452.7184, pp. 158–161.

Ellis, George Francis Rayner (2014). "On the philosophy of cosmology". In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 46, pp. 5–23.

Garbrecht, Björn (2020). “Why is there more matter than antimatter? Calculational methods for leptogenesis and electroweak baryogenesis”. In: *Progress in Particle and Nuclear Physics* 110, p. 103727.

Geroch, Robert, Liang Can-bin, and Robert M Wald (1982). “Singular boundaries of space–times”. In: *Journal of Mathematical Physics* 23.3, pp. 432–435.

Guth, Alan H (1981). “Inflationary universe: A possible solution to the horizon and flatness problems”. In: *Physical Review D* 23.2, p. 347.

Guth, Alan H (1997). *The inflationary universe: the quest for a new theory of cosmic origins*. London: Jonathan Cape.

Guth, Alan H (2000). “Inflation and eternal inflation”. In: *Physics Reports* 333, pp. 555–574.

Guth, Alan H (2007). “Eternal inflation and its implications”. In: *Journal of Physics A: Mathematical and Theoretical* 40.25, p. 6811.

Guth, Alan H, David I Kaiser, and Yasunori Nomura (2014). “Inflationary paradigm after Planck 2013”. In: *Physics Letters B* 733, pp. 112–119.

Holman, Marc (2018). “How problematic is the near-Euclidean spatial geometry of the large-scale universe?” In: *Foundations of Physics* 48.11, pp. 1617–1647.

Huggett, Nick and Christian Wüthrich (2013a). “Emergent spacetime and empirical (in) coherence”. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 44.3, pp. 276–285.

Huggett, Nick and Christian Wüthrich (2013b). “The emergence of spacetime in quantum theories of gravity”. In: *Studies in the History and Philosophy of Modern Physics* 44.3, pp. 273–275.

Huggett, Nick and Christian Wüthrich (2018). “The (a) temporal emergence of spacetime”. In: *Philosophy of Science* 85.5, pp. 1190–1203.

Huggett, Nick and Christian Wüthrich (2021). “Out of Nowhere: Introduction: The emergence of spacetime”. In: *arXiv preprint arXiv:2101.06955*.

Ijjas, Anna, Paul J Steinhardt, and Abraham Loeb (2013). “Inflationary paradigm in trouble after Planck2013”. In: *Physics Letters B* 723.4-5, pp. 261–266.

Ijjas, Anna, Paul J Steinhardt, and Abraham Loeb (2014). “Inflationary schism”. In: *Physics Letters B* 736, pp. 142–146.

Kinney, Will (2022). *An Infinity of Worlds: Cosmic Inflation and the Beginning of the Universe*. Cambridge, MA: MIT Press.

Kinney, William H, Sunny Vagnozzi, and Luca Visinelli (2019). “The zoo plot meets the swampland: mutual (in) consistency of single-field inflation, string conjectures, and cosmological data”. In: *Classical and quantum gravity* 36.11, p. 117001.

Koberinski, Adam, Bridget Falck, and Chris Smeenk (2023). “Contemporary Philosophical Perspectives on the Cosmological Constant”. In: *Universe* 9.3, p. 134.

Koberinski, Adam and Chris Smeenk (2023). “Λ and the limits of effective field theory”. In: *Philosophy of Science* 90.2, pp. 454–474.

Kolb, Edward W and Stephen Wolfram (1980). “Baryon number generation in the early universe”. In: *Nuclear Physics B* 172, pp. 224–284.

Linde, Andrei (2015). "Inflationary cosmology after Planck". In: *Post-Planck Cosmology: Lecture Notes of the Les Houches Summer School 2013.* Ed. by Cédric Deffayet et al. Vol. 100. Oxford University Press Oxford, England, UK.

Loewer, Barry (2012). "Two accounts of laws and time". In: *Philosophical studies* 160, pp. 115–137.

Longair, Malcolm S. (2006). *The cosmic century: a history of astrophysics and cosmology.* Cambridge University Press.

Martin, Jerome and Robert H. Brandenberger (2001). "Trans-Planckian problem of inflationary cosmology". In: *Physical Review D* 63.12, p. 123501.

Martin, Jerome, Christophe Ringeval, and Vincent Vennin (2014). "Encyclopædia inflationaris". In: *Physics of the Dark Universe* 5, pp. 75–235.

McCoy, Casey D (2015). "Does inflation solve the hot big bang model s fine-tuning problems?" In: *Studies in History and Philosophy of Modern Physics* 51, pp. 23–36.

Morrissey, David E and Michael J Ramsey-Musolf (2012). "Electroweak baryogenesis". In: *New Journal of Physics* 14.12, p. 125003.

Myrvold, Wayne C. (2016). "Probabilities in Statistical Mechanics". In: *The Oxford Handbook of Probability and Philosophy.* Ed. by Christopher Hitchcock and Alan Hájek. Oxford: Oxford University Press, pp. 573–600.

Myrvold, Wayne C. (2021). *Beyond chance and credence: A theory of hybrid probabilities.* Oxford University Press, USA.

North, Jill (2011). "Time in thermodynamics". In: *The oxford handbook of philosophy of time*, pp. 312–350.

Norton, John D (2010). "Cosmic confusions: Not supporting versus supporting not". In: *Philosophy of Science* 77.4, pp. 501–523.

Norton, John D (2018). "Eternal inflation: when probabilities fail". In: *Synthese*, pp. 1–23.

Oriti, Daniele (2018). *Levels of spacetime emergence in quantum gravity.* arXiv: `1807.04875 [physics.hist-ph]`.

Peebles, Phillip James Edwin (2020). *Cosmology's Century: An inside history of our modern understanding of the universe.* Princeton: Princeton University Press.

Penrose, Roger (1979). "Singularities and time-asymmetry". In: *General Relativity: An Einstein centenary survey.* Ed. by Stephen Hawking and Werner Israel. Cambridge: Cambridge University Press, pp. 581–638.

Penrose, Roger (2004). *The road to reality: A complete guide to the laws of the universe.* Jonathan Cape.

Price, Huw (1996). *Time's arrow & Archimedes' point: new directions for the physics of time.* Oxford University Press, USA.

Price, Huw (2004). "On the origins of the arrow of time: Why there is still a puzzle about the low-entropy past". In: *Contemporary debates in philosophy of science*, pp. 219–239.

Riotto, Antonio (1998). "Theories of baryogenesis". In: *Proceedings, Summer School in High-energy physics and cosmology: Trieste, Italy*, pp. 326–436.

Riotto, Antonio and Mark Trodden (1999). “Recent progress in baryogenesis”. In: *Annual Review of Nuclear and Particle Science* 49.1, pp. 35–75.

Robertson, Katie (2019). “Stars and steam engines: To what extent do thermodynamics and statistical mechanics apply to self-gravitating systems?” In: *Synthese* 196.5, pp. 1783–1808.

Rovelli, Carlo (2019). “Where was past low-entropy?” In: *Entropy* 21.5, p. 466.

Rugh, Svend E. and Henrik Zinkernagel (2002). “The quantum vacuum and the cosmological constant problem”. In: *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics* 33.4, pp. 663–705.

Sakharov, Andrej Dmitrievich (1967). “Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe”. In: *Journal of Experimental and Theoretical Physics Letters* 5, pp. 24–27.

Sanders, Robert H (2010). *The dark matter problem: a historical perspective.* Cambridge University Press.

Smeenk, Chris (2005). “False vacuum: Early universe cosmology and the development of inflation”. In: *The universe of general relativity.* Ed. by J. Eisenstaedt and A. J. Kox. Boston: Springer, pp. 223–257.

Smeenk, Chris (2014). “Predictability crisis in early universe cosmology”. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 46, pp. 122–133.

Smeenk, Chris (2019). “Gaining access to the early universe”. In: *Why Trust a Theory.* Ed. by Radin Dardashti, Richard Dawid, and Karim Thébault. Cambridge University Press, pp. 315–335.

Smith, George E (2014). “Closing the loop”. In: *Newton and empiricism.* Ed. by Zvi Biener and Eric Schliesser. Oxford University Press, USA, pp. 262–352.

Sudarsky, Daniel (2014). “The inflationary origin of the seeds of cosmic structure: Quantum theory and the need for novel physics”. In: *General Relativity, Cosmology and Astrophysics: Perspectives 100 years after Einstein’s stay in Prague*. Springer, pp. 349–367.

Sus, Adán (2014). “Dark matter, the Equivalence Principle and modified gravity”. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 45, pp. 66–71.

Swart, Jaco de, Gianfranco Bertone, and Jeroen van Dongen (2017). “How dark matter came to matter”. In: *arXiv preprint arXiv:1703.00013.*

Uffink, Jos (2006). “Compendium of the foundations of classical statistical physics”. In:

Vafa, Cumrun (2005). *The String Landscape and the Swampland.* arXiv: `hep-th/0509212 [hep-th]`.

Vanderburgh, William L (2003). “The dark matter double bind: Astrophysical aspects of the evidential warrant for general relativity”. In: *Philosophy of science* 70.4, pp. 812–832.

Vistarini, Tiziana (2019). *The emergence of spacetime in string theory.* Routledge.

Wald, Robert M (1994). *Quantum field theory in curved spacetime and black hole thermodynamics.* University of Chicago press.

Wallace, David (2011). “The logic of the past hypothesis”. In:

Wallace, David (2020). “Gravity, entropy, and cosmology: In search of clarity”. In: *The British Journal for the Philosophy of Science*.

Weinberg, Steven (2008a). *Cosmology*. Oxford University Press.

Weinberg, Steven (2008b). “Effective field theory for inflation”. In: *Physical Review D* 77.12, p. 123541.

White, Roger (2000). “Fine-tuning and multiple universes”. In: *Noûs* 34.2, pp. 260–276.

Wolf, William J and Patrick M Duerr (2023). “The Virtues of Pursuit-Worthy Speculation: The Promises of Cosmic Inflation”. In: *arXiv preprint arXiv:2309.16266*.

Zinkernagel, Henrik (2002). “Cosmology, particles, and the unity of science”. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 33.3, pp. 493–516.