



THE MULTIVERSE: A PHILOSOPHICAL INTRODUCTION

by Jeremy Butterfield

ISBN 978-1-77385-699-5

THIS BOOK IS AN OPEN ACCESS E-BOOK. It is an electronic version of a book that can be purchased in physical form through any bookseller or on-line retailer, or from our distributors. Please support this open access publication by requesting that your university purchase a print copy of this book, or by purchasing a copy yourself. If you have any questions, please contact us at ucpress@ucalgary.ca

Cover Art: The artwork on the cover of this book is not open access and falls under traditional copyright provisions; it cannot be reproduced in any way without written permission of the artists and their agents. The cover can be displayed as a complete cover image for the purposes of publicizing this work, but the artwork cannot be extracted from the context of the cover of this specific work without breaching the artist's copyright.

COPYRIGHT NOTICE: This open-access work is published under a Creative Commons licence. This means that you are free to copy, distribute, display or perform the work as long as you clearly attribute the work to its authors and publisher, that you do not use this work for any commercial gain in any form, and that you in no way alter, transform, or build on the work outside of its use in normal academic scholarship without our express permission. If you want to reuse or distribute the work, you must inform its new audience of the licence terms of this work. For more information, see details of the Creative Commons licence at: <http://creativecommons.org/licenses/by-nc-nd/4.0/>

UNDER THE CREATIVE COMMONS LICENCE YOU MAY:

- read and store this document free of charge;
- distribute it for personal use free of charge;
- print sections of the work for personal use;
- read or perform parts of the work in a context where no financial transactions take place.

UNDER THE CREATIVE COMMONS LICENCE YOU MAY NOT:

- gain financially from the work in any way;
- sell the work or seek monies in relation to the distribution of the work;
- use the work in any commercial activity of any kind;
- profit a third party indirectly via use or distribution of the work;
- distribute in or through a commercial body (with the exception of academic usage within educational institutions such as schools and universities);
- reproduce, distribute, or store the cover image outside of its function as a cover of this work;
- alter or build on the work outside of normal academic scholarship.



Acknowledgement: We acknowledge the wording around open access used by Australian publisher, **re.press**, and thank them for giving us permission to adapt their wording to our policy <http://www.re-press.org>

All the Worlds from the Primordial Bubbles

This chapter discusses our third multiverse. It was proposed by some cosmologists from the early 1980s onwards, on the basis of their theories about the Big Bang origin of the cosmos. Like the Everettian multiverse in the last chapter, it is agreed even by its proponents to be very speculative and hard to confirm. This common feature is unsurprising. As we discussed in chapter 1, there might well be, within a branch of physics, theoretical (or more generally, conceptual) reasons for a proposal that is hard to confirm—and whose assessment thus calls on conceptual, even philosophical, arguments.

But there is also a dissimilarity with the Everettian multiverse. For that multiverse arose as a solution to a problem about the general structure of our supremely successful quantum theory. Recall that although we concentrated for simplicity on elementary quantum theory, e.g. the quantum replacement of classical point particles, all our more advanced quantum theories, including our well-established quantum field theories of electrons, quarks, etc., equally face the measurement problem (sections 3 and 4 of chapter 4). And the reasons in favour of the Everettian solution to that problem apply equally to them (as do, broadly speaking, the reasons against the Everettian solution).

On the other hand, the cosmological multiverse arises by combining two speculations that are very specific. One is a mechanism called *inflation*, which is speculated to have operated for a tiny fraction of a second in the very early history of the universe, i.e. very soon after the Big Bang. The idea is that for this tiny fraction of a second, the cosmos expanded vastly, and also at an accelerating rate. Cosmologists' original rationale, in the

early 1980s, for proposing inflation was not that it solved a problem about the general structure of a physical theory, quantum or otherwise. Rather, it answered pressing “Why?” questions. It promised to explain some facts that were otherwise puzzling, even mysterious. These questions were about the value of a physical quantity “having to be just-so,” i.e., having to take a value that was constrained to many decimal places, if our cosmological theory was to adequately describe empirical observations. This just-so-ness is usually called *fine-tuning*. (I shall focus on two such questions, called the *flatness problem* and the *horizon problem*.)

These “Why?” questions, i.e., “Why is the value fine-tuned?,” prompt more general philosophical questions: “What counts as an *explanation*?” and “How could science *confirm* a multiverse proposal?” These will be this chapter’s main philosophical questions, which we pursue from section 5 onward.

The other speculation, the second ingredient of the cosmological multiverse, is *string theory* (which also began to be developed in the early 1980s). It is a quantum theory that unifies general relativity’s successful account of gravitation with quantum field theory’s successful account of nature’s other fundamental forces. Section 4 will give a few more details about it. But the main point for us is that it combines with inflation so as to give a multiverse. For this, the key contribution from string theory is that (i) it has many *vacuum states* (also called *ground states*, or for short, *vacua*); (ii) these states differ from each other in the values of what we usually call *constants of nature*; and (iii) each vacuum state has an associated set of states that share with the vacuum state its values for these constants. (String theorists realized this only in about 2000; they originally hoped that the theory would have a unique vacuum state.)

Obviously, a full understanding of (i), (ii), and (iii) requires the demanding technicalities of string theory. But—fortunately, for a philosophical book—we can get by with a sketch. Sections 3 and 4 will give more details. But for the moment, the following three points will suffice.

- (i.) Beware that “vacuum” does *not* mean “nothing.” Nor does it mean “no physical system,” i.e., sheer absence of the physical system that the theory is intended to describe. Instead, it means “state of lowest energy.” Hence, the synonym “ground state,” with states of higher energy being called “excited” or “above the ground state.” By and large,

physical systems tend to lose energy and thereby to evolve over time toward their vacuum states. Many physical systems have a unique vacuum state, but by no means all do—a glass has many vacuum states.

- (ii.) “Constant of nature” means a value of a physical quantity that, so far as we have measured it, does *not* vary across the entire observable universe. (So, this is unlike chapter 4’s examples of a system’s values of position, momentum, etc., which obviously vary.) Here are three examples: the amount of electric charge on an electron, the ratio of the strengths of electromagnetic and gravitational forces, and the speed of light. And there are many others, such as the masses of elementary particles, like electrons and quarks, and the strengths of the forces between them. And as we shall see, the quantities whose fine-tuning gives the flatness and horizon problems are also examples. But string theory is, so to speak, liberal or unopinionated, in the sense that its different vacuum states have different values for these “constants.”
- (iii.) For each vacuum state, its associated set of states is a set of higher energy, excited, states that share the vacuum state’s values for these “constants.” (But I shall often drop the double “scare” quotes that are meant to signal that the value can vary across the string theory’s state space, and so I will just write the word “constant.” Another common term avoiding the connotations of “constant” is “parameter” or “cosmological parameter.”) We can think of this set as a *tower* of states, standing above the vacuum.

So to sum up (i) to (iii), the striking fact about the many vacua of string theory, and their associated towers of excited states, is that the values of the constants listed in (ii) vary from one vacuum and its tower to another.

One natural (perhaps the most natural) first response to this fact is that string theory should simply restrict itself to the tower of states whose values match the values that we actually measure. That is, it should explore the theoretical features and experimental consequences of states in that tower,

one of which would then be (according to string theory) the actual state of the cosmos. Thus, one might say that it is no demerit of string theory that it could describe various non-actual values of the constants, but also no merit. After all, the same is true of our other theories. For example, classical electromagnetic theory could describe electromagnetism with various non-actual values of the charge on an electron.

But it turns out that combining string theory with inflation, i.e., the very early, very brief but also very rapid expansion of the cosmos, yields a mechanism that makes for a multiverse. Roughly speaking, the expansion makes a quantum state in one tower evolve over time so as to have a component in (an amplitude for) other towers. Besides, this happens in a runaway fashion, called *eternal inflation*, so that states in very many towers, and so very many combinations of possible values of the “constants” are equally allowed.

It is evident in the light of chapter 4 that here we face questions about the interpretation of the quantum state. Thus, suppose one takes a non-zero (or at least large enough) amplitude in the quantum state to correspond to something real. (Note that this does *not* commit one to being an Everettian in any of the senses discussed in chapter 4. For they all maintain that the Schrödinger equation is, in Bell’s phrase, “always right,” whereas this supposition is entirely compatible with (i) believing that wave-function collapse is a real dynamical process, and with (ii) understanding “correspond to something real” in some different way than the straightforward Everettian way.) And suppose also that one accepts the cosmologist’s quantum state, with amplitudes for states in many towers, as correct. Then one must conclude that each of many combinations of values of the “constants” corresponds to something real.

This is the cosmological multiverse, whose different universes are described by states in different towers, differing in their constants of nature. These universes are often called “bubbles” (or “domains”)—hence this chapter’s title.

So much by way of introducing inflation and string theory, and how they combine to give a cosmological multiverse. From a philosophical perspective, it is remarkable how they *both* raise the philosophical theme of explanation. For I said above that inflation was originally motivated by “Why?” questions about fine-tuning. And we have just seen that string theory is a framework with constants of nature—including the quantities whose fine-tuning gives the flatness and horizon problems—varying across

the multiverse. This confluence obviously prompts an ambitious project to somehow *explain* the values of all these constants—i.e., the values that we in this universe actually measure—by invoking some appropriate features of how the values vary across the multiverse. Hence, this chapter takes explanation and confirmation as its main philosophical topics.

With this background in place, I can now explain the plan of the chapter. There will be four stages. The first stage (just section 1) clarifies the relation between the Everettian and cosmological multiverses. This will develop the discussion above, about the cosmological multiverse involving not just inflation but also string theory and its many vacuum states. But again, we will be able to proceed with hardly a mention of the advanced physics involved.

Then, in the second stage, I introduce modern cosmology and its multiverse. Section 2 will summarize what cosmology had established by about 1980. Section 3 will do two jobs. First, I report the puzzling facts about fine-tuning that remained unexplained (i.e., the flatness and horizon problems), and how inflation could explain them. Second, I sketch how inflation led to a multiverse, though without details of string theory. Then section 4 supplies some details about string theory.

The third and fourth stages (section 5 onward) are philosophical: they address explanation and confirmation.

In the third stage (sections 5 to 7), section 5 first reviews some of the philosophical literature about explanation. Then it formulates two overall strategies we could adopt to explain the values we actually measure. (The distinction between the strategies will not depend on details of the physics.) The first strategy is the obvious one: to argue that the value we measure is in some precise sense *generic*, or *typical*, of the values across the multiverse, and is thereby to be expected. But there is a second, less obvious, strategy. It invokes what are called *selection effects* or (a better-known term) *the anthropic principle* to argue that the value we measure is *likely to be observed*, even if it is not generic or typical across the whole multiverse, simply because most “regions” of the multiverse have no observers. I will label these two strategies, “strategy (Gen)” and “strategy (Obs),” with the labels standing for “generic” and “observation,” respectively.

The rest of the third stage (the next two sections) develops the strategy (Gen), that is, of explaining a fact by showing that it is generic or typical. Section 6 describes the strategy’s successes and its positive features. Section 7 describes its difficulties, including in cosmology.

The fourth stage (sections 8 and 9) is about the strategy (Obs): explaining a fact by showing it is probable (or probable enough) that it be observed, even if, observation apart, it is improbable (or not probable enough). Section 8 explains the strategy with examples from outside cosmology, indeed, from outside physics. It also introduces the term *selection effects*. Section 9 discusses the strategy within cosmology. It introduces the term *the anthropic principle*. And, as an example, it summarizes the anthropic explanation of the value of a “constant” (also called “parameter”), which I have not mentioned so far, namely, the *cosmological constant*.

Finally, section 10 concludes. Here, I recommend a framework for confirming a theory of the multiverse that incorporates ideas from *both* strategies: the idea of being generic and the idea of being probable (or probable enough) to be observed.

1. Comparing the Everettian and Cosmological Multiverses

The mere phrase “the cosmological multiverse” suggests there should be connections with the Everettian multiverse proposal. For as we saw, the Everettian proposes a quantum state of the cosmos, written with the capital Greek letter “Psi,” i.e., written as Ψ . (Recall section 6 of chapter 4’s decision to use this phrase, not the more usual “quantum state of the universe.”) So, the job of this section is to describe these connections. I shall, of course, have to set aside many details of advanced physics, especially string theory, but this omission will, I think, be justified by its not affecting our philosophical questions about explanation and confirmation. We can spell out the connections in four comments, covered in 1.1 to 1.4.

1.1 Contrasts Between the Multiverses

When one first meets the phrase “the cosmological multiverse,” and bears in mind that modern cosmologists, of course, accept quantum theory, one naturally expects that (i) some cosmologists will endorse the Everett interpretation, and even (ii) the cosmological multiverse will turn out to be an elaboration of the Everettian one.

The first of these expectations (i) is indeed true. Among cosmologists interested in interpretative questions about quantum theory—and, of course, the cosmological multiverse raises such questions—the Everettian interpretation is popular.

But, on the other hand, the second expectation (ii) is true only in a liberal, i.e., logically weak, sense of “elaboration.” Yes, a cosmologist may well accept that there is a quantum state of the cosmos and may also argue that it describes the cosmological multiverse they advocate. But the universes (or worlds) in this multiverse are very different from the Everettian universes (or worlds) defined by the well-established, continual, ubiquitous, and rapid process of decoherence applying to macrosystems, which we discussed in section 7 of chapter 4. For (as I said in this chapter’s preamble), the cosmologists’ many universes are produced by a specific mechanism, *inflation*, that is speculated to have operated for a tiny fraction of a second very soon after the Big Bang. Agreed, in many such universes, there will indeed be (later, long after the universe starts) macrosystems such as dust particles, or even pointers and cats: macroscopic objects that can decohere. And for these, the Everettian can then claim that the last chapter’s account, with its Ψ applies.

So, the upshot is that if we accept the cosmological multiverse based on the idea of inflation, then the quantum state of the cosmos—in the sense we envisaged in the last chapter, namely, Ψ with components, i.e., summands, describing various decoherent worlds (cf. section 8 of chapter 4) —is at best a description of a *single* universe within the cosmological multiverse.

Besides, I say “at best” not just because inflation occurs vastly earlier than any macroscopic objects exist. Also, as I said in the preamble, the cosmological multiverse is based on string theory (combined with inflation). So, the cosmologist’s “quantum state of the cosmos” will be a state in string theory, not a state in a well-confirmed quantum theory. In particular, it is not a state in our well-confirmed theory of electrons, quarks, etc., which is nowadays called “the standard model.” This distinction matters because there is a “large gap” between string theory states and those in established quantum field theories. Indeed, this is a notorious fact about string theory. Namely, it is very hard to get out of string theory any empirical predictions or even the recovery of specific theoretical postulates of confirmed quantum field theories. So, the upshot announced in the last paragraph assumes in effect that this large gap has been bridged. It assumes that we can get by deduction (or by some approximation to deduction) from the official “stringy” state of the cosmos to the states of our confirmed quantum field theories.

So, the cosmologist we envisaged two paragraphs above—who accepts that there is a quantum state of the cosmos, and says it describes the cosmological multiverse they advocate—will mean by “quantum state

of the cosmos” something yet more dizzying than the last chapter’s (the Everettian’s) Ψ . There are two reasons for this.

First, it means a string theory state, and one needs to bridge the gap to quantum field theory. Second, as announced in this chapter’s preamble, it means a state that encompasses various alternative values of (what we call!) the constants of nature. Here, my “encompasses” is deliberately vague. For the way that the theory (this state) describes eternal inflation’s runaway process of the different universes (with their different constants of nature) coming to exist is complicated and controversial (not least because of the challenge of the first reason).

But for our purposes, it will suffice to boldly ignore the gap between string theory and quantum field theory, and to construe “encompasses” as “has amplitudes for.” That is, it will suffice to take our cosmologist’s “quantum state of the cosmos” to be a sum of “Everettian-chapter-4 states,” i.e., a sum (superposition) of “quantum *states* of the cosmos,” using this last phrase in the sense of the last chapter’s Everettian. So, we take it to be a sum of different Everettian-chapter-4 states Ψ_i , where the suffix “i” ($i = 1, 2, \dots$) is a label on the summands, i.e., the different universes, which in general disagree with each other about the values of (what we call) constants of nature.

1.2 Decoherence Without Macrosystems

Here, I should give a clarification. For the sake of a clearer exposition, I have in 1.1 above taken decoherence as a process that occurs in macrosystems due to their interaction with an environment, e.g., dust particles immersed in air, just as I did in chapter 4. As just explained, this gives a crisp contrast between chapter 4’s ideas and this chapter’s new idea, the mechanism of inflation. But cosmologists *also* apply the idea of decoherence, in various detailed ways, to the cosmos as a whole, including at very early times.

One way this is done is to take all the material in the cosmos to be the system (i.e., the analogue of the dust particle), and space-time to be its environment (i.e., the analogue of the air). Here, “material” will include not just matter, e.g., atoms, electrons, and quarks, but also radiation, like electromagnetic radiation. And to give content to the idea of this material *interacting* with space-time, one needs states of space-time that respond to the states of matter, so that one describes space-time with general relativity, according to which it is indeed responsive in this way. (But as mentioned in section 4 of chapter 8, the system–environment split is sometimes made in

a less “obvious” way, e.g., taking long-wavelength modes to be the system, which is then decohered by short-wavelength modes.)

In this kind of way, even without macrosystems such as dust particles, decoherence can occur. So, cosmologists will talk of all the cosmos’ material degrees of freedom being in an improper mixture (cf. the end of section 7 of chapter 4) of states that are definite for some appropriate quantities. So, in short, the contrast between chapter 4’s Everettian multiverse and the cosmological multiverse is not as crisp as I first suggested. That is, the contrast is not as crisp as decoherence for macrosystems, without variation in constants of nature vs. string theory and inflation, with varying constants.

Nevertheless, in the rest of this chapter, we can safely take this contrast to be correct. That is, we can think of the cosmologist’s “quantum state of the cosmos” as a sum of different Everettian-chapter-4 states Ψ , and these states in general disagree about the values of constants of nature, such as the charge of the electron. This picture of the relation between the two multiverse proposals will set us in good stead for our philosophical questions about explanation and confirmation.

Of course, these are not just speculative, but also imprecise and dizzying ideas. Just as I remarked (in section 6 of chapter 4) that no Everettian has the faintest idea how to write down in detail the Everettian quantum state of the cosmos Ψ , no quantum cosmologist can now write down in detail *their* quantum state of the cosmos. For as this chapter will report, we do not know—and we may never know—the underlying physics of inflation. So, a sketch definition of “universe” for the cosmological multiverse, on analogy with the Everettian’s sketch definition of “world” (section 8 of chapter 4), is far beyond current knowledge.

1.3 The Measurement Problem on a Cosmic Scale

But agreed, you can’t keep a good idea down. I should also report that some quantum cosmologists, sympathetic to the Everett interpretation, have proposed mathematical formulas for the quantum state of the cosmos at very early times. Of course, the details vary from one author or research programme to another. Thus, some of these formulas are independent of whether there was a very brief period of inflation, while some incorporate such a period. Some are string-theoretic, some are not. Some are independent of decoherence at very early times, e.g., of the material degrees of freedom, while some incorporate it. One such formula, proposed by Hartle and Hawking in 1983 (and independently of all three topics above: inflation,

string theory, and decoherence), is called *the no-boundary proposal* or *no boundary condition*. It has been much studied and developed since then, and indeed, it has been related to inflation, string theory, and decoherence.

This reflects the more general fact that most quantum cosmologists recognize the relevance of interpretative questions and related methods and ideas, like decoherence, to their scientific work. This relevance is shown by the point above: that cosmologists' models often incorporate decoherence at very early times, long before there were macrosystems, like dust particles.

But for this book's theme of the multiverse, the most vivid example of this relevance will be the point I made in the preamble (and will develop in section 3 below): that if (i) one takes a non-zero (or at least large enough) amplitude in the quantum state to correspond to something real, and (ii) one accepts the cosmologist's quantum state, with amplitudes for states in many towers with differing constants, as correct, then one must conclude that each of many combinations of values of the constants are real.

As noted in the preamble, (i) does not require being Everettian. In particular, it does not require the Schrödinger equation to be, in Bell's phrase, "always right." Nor does combining (i) with (ii). For there might be a model of wave-function collapse (a revision of the Schrödinger equation of the kind discussed in section 5 of chapter 4), which describes the formation of a bubble in the expanding inflationary cosmos; and which can also describe the formation of many bubbles—all, thanks to (i), equally real—and thus, a multiverse. But Everettian or not, the relevance of interpretative questions to the cosmological multiverse is vivid. One might well say that taking all the universes ("bubbles") in the cosmological multiverse to be equally real involves assuming some solution, on a cosmic scale, of the quantum measurement problem.

1.4 The Cosmic Microwave Background as an Illustration

Nor is this vivid example the only place where modern theoretical cosmology meets the measurement problem on a cosmic scale. Even without going back in time as far as the putative period of inflation, and without postulating a multiverse, modern cosmology describes early states of the universe, e.g., a minute, a year, or ten thousand years after the Big Bang, in terms of quantum theory—and so the measurement problem arises.

Indeed, it arises in connection with something so basic and vivid to us as the existence of stars, planets, and galaxies. For as I will explain in section 2, there is a weak electromagnetic signal throughout space called the

“cosmic microwave background” (CMB) radiation that dates from about 380,000 years after the Big Bang—and which we can directly observe. We observe this CMB radiation to be very smooth and uniform: it looks the same in all directions. But it has tiny “wrinkles,” which are really variations in the quantum amplitudes for various densities of mass in regions of space. This means that a peak among these wrinkles is a “seed” of a clump of matter becoming gradually localized, under the gravitational attraction of its component parts, in one region of space rather than another. Such a clump, once localized, can grow, pulling yet more matter in by its gravity, and eventually produce, for example, a galaxy.

But note that here, the word “seed” is a metaphor that hides the problematic issue of the collapse of the wave function. For whereas a real seed is an actually existing object that grows into an actually existing plant, this peak of quantum amplitude is only a higher (square root of a) probability for the event of clumping to happen here, rather than there. (Unfortunately, the metaphor is entrenched in textbooks as well as popular expositions. Only the better textbooks admit that this transition, from peaks and troughs of quantum amplitude to a classical, slightly uneven, distribution of mass density across space, is problematic, since it is a cosmic version of the “collapse of the wave function,” which all agree is problematic.)

1.5 The Need for More Precision

So much by way of sketching connections between the Everettian and cosmological multiverses. Or to put it more generally and precisely, so much for the connections between (i) the interpretation of quantum theory, especially its measurement problem, and (ii) quantum-to-classical transitions in the very early cosmos (“primordial bubbles”) and in the not-so-early cosmos (“wrinkles” in the CMB).

Obviously, these connections are important—indeed, fundamental. But this chapter will not go into further details about them for two reasons: one negative and one positive. The negative reason is that most cosmologists, even quantum cosmologists, believe that these cosmological aspects of the measurement problem (or more neutrally: of the interpretation of quantum states) are not yet precisely enough formulated to be addressed as a problem within physics. In short, we do not know enough, and the time is not yet ripe. The positive reason is that (as we shall see) even if we restrict ourselves to a “classical outlook,” there is so much to explore, in both the physics and

the philosophy of modern cosmology. At least, there is certainly enough for this chapter.

2. A Golden Age of Cosmology

We live in a golden age of cosmology. It began in the twentieth century, especially in its second half. I will describe it in three stages, covering the first half of the twentieth century in 2.1, and the subsequent history in 2.2 and 2.3.

2.1 *The First Half-Century*

Already in the first half of the century, there were four momentous developments in cosmology: two observational and two theoretical.

First, we discovered that nebulae appearing in telescopes as cloudy smudges rather than point-like stars were actually other galaxies, each containing vast numbers of stars like our own Milky Way. So, the cosmos turned out to be vastly larger than had been envisaged. Second, we discovered that the cosmos is expanding. More precisely, any two galaxies were found to be receding from each other, i.e., the distance between them was increasing. (The speed of recession is approximately proportional to the distance between them.) But this is not an expansion of matter into a pre-existing empty space, like an explosion of a firework or a bomb. Rather, the space itself is expanding. Agreed, that is impossible to visualize. We are bound to think of an ambient or embedding space relative to which the expansion occurs; and it was a struggle for physicists to accept this idea.

In this acceptance, the third development was crucial, namely, Einstein's discovery of general relativity and its application to the whole cosmos. As mentioned in chapter 2, general relativity is a relativistic theory of gravitation. According to this theory, gravitational influence propagates across space at the speed of light, not instantaneously, as in Newton's theory. And like Einstein's earlier theory of special relativity, it unifies space and time into a single entity, space-time, which, being four-dimensional, is again unvisualizable. (Physicists' acceptance of these unvisualizable ideas, of expanding space or space-time, was helped by the rise of pure mathematics, reviewed at the end of chapter 2 (the increasing formalization of mathematics included liberating geometry from visual intuition)).

Einstein himself, immediately after formulating general relativity, applied the theory to the cosmos as a whole. In terms of our terminology for

systems and their state spaces (cf. section 3 of chapter 3), he took the cosmos as his system, and he described it as a space-time, extending not just arbitrarily far in all directions in space, but also throughout the past and future. (This is reminiscent of our description of the cosmos, “our world,” in section 1 of chapter 1) So, he aimed to find a solution to the equations of general relativity that described the whole of space-time, not, of course, in its myriad details, but in the broadest possible terms. In particular, matter was treated as smoothed out uniformly across space, although, of course, it is in fact concentrated in great clumps. (For the galaxies are clumps, and within them, the stars are clumps.) By the mid-1920s, solutions of general relativity describing an expanding cosmos (whose matter is smoothed out uniformly across space) had been found, and in the following years, they were investigated and elaborated. In some of these solutions, the expansion began from a very hot, very dense state, which came to be called “the primeval fireball.”

The fourth development was the rise of astrophysics, i.e., the physics of stars. Quantum theory, discovered in the 1920s, was applied to describe in detail, not just how stars shine by burning helium, but also much else: the different types of stars, how other chemical elements are formed in stars (called *nucleosynthesis*, since the stars synthesize i.e., make the nuclei of, elements), and how and why some stars explode and others implode.

By the late 1940s, the third and fourth developments had been combined. For the ideas and results of astrophysics were applied to describe nucleosynthesis in the conjectured primeval fireball. This led to detailed predictions about the cosmic abundances of light elements, like hydrogen, helium, and lithium. It also led to the prediction of a pervasive but very faint electromagnetic radiation with a characteristic wavelength that was a remnant of the fireball.

2.2 The CMB and the Standard Model of Cosmology

This is how matters stood in about 1960. It was the following years that really ushered in the golden age. Again, there were several momentous developments, both observational and theoretical. I shall pick out three that were all well underway by 1980. Developments after 1980, including the proposed multiverse, will be treated in the next section.

Foremost among observations was the discovery (by accident, in 1964) of the predicted remnant radiation. This is the “cosmic microwave background” (or “CMB”) radiation. It was immediately recognized as confirming

general relativity's expanding cosmological solutions. In just a few years, almost all cosmologists accepted that the cosmos originated in a primeval fireball about fourteen billion years ago. (The fireball was soon renamed the "Big Bang," a name that had originally been suggested by sceptics, as a derisive label.) Besides, in the following decades, the CMB has proven to be an extraordinarily rich source of information about the early cosmos.

The second main observational development between 1960 and 1980 was the invention and deployment of several new sorts of telescopes that enabled astronomers and cosmologists to study types of electromagnetic radiation other than visible light. For radio waves (with wavelengths much longer than visible light), there was ground-based radio astronomy, which had been pioneered in the 1940s. For microwave and infrared radiation (i.e., wavelengths a bit longer than visible light) and X-rays (shorter than visible light), one needed to get above the Earth's atmosphere. For these wavelengths, dozens of satellite missions have yielded a profusion of data, both for astrophysics and cosmology: for example, data on the CMB and the cosmic abundances of elements.

The third development was theoretical. After 1960, there was a renaissance in the study of general relativity, both as regards its mathematics and its applications. Here, one high point was a cluster of theorems saying that among the solutions of general relativity (i.e., the space-times that are possible according to the theory), *singularities* are generic, i.e., typical. The idea of a singularity is a breakdown in the smooth structure of space-time, and the theorems showed that such a breakdown must occur under certain conditions. These conditions included when a star whose mass is above a certain limit, having burned all its fuel, implodes under its own weight. This is called the "gravitational collapse" of the star, and it leads to a *black hole* in which the singularity lies. But more relevant to us: among these conditions were conditions that were understood to prevail in the early cosmos. This gave a new perspective on the simple expanding cosmological solutions of general relativity that had been recognized, already in the 1920s, as having an initial singularity: the original "point where the fireball began" (though not a point in space-time itself). Namely, this initial singularity came to be regarded, not as an artefact due to the solution's admittedly very idealized treatment (especially its smoothing out the matter), but as a robust feature of the solution that might well be physically real.

The result of these three developments was that by the mid-1970s, cosmologists had agreed on a model of the history of the cosmos, with an initial

singularity about fourteen billion years ago, followed by a hot primeval fireball that cooled and expanded. It was called “the standard model.” (This is not to be confused with its namesake, the standard model in high-energy physics. That describes the physics of electrons, photons, neutrinos, and quarks, i.e., the constituents of protons and neutrons, and also of unstable particles. It was formulated in the mid-1970s.)

In the last fifty years, this standard model of cosmology has stood up amazingly well (as has its namesake in high-energy physics). Agreed, in addition to elaborating ideas and methods that were already formulated in the 1970s, two major new ideas have had to be added to accommodate observations. One such idea is that much, indeed, the majority, of the mass of the cosmos is of an as-yet unknown form; this is called “dark matter.” Another is that although one would expect the expansion of the cosmos to slow down, i.e., to decelerate (since the stars and galaxies, having mass, pull on each other gravitationally), the expansion is in fact accelerating. This is called “dark energy.”

(Incidentally, these two ideas have prompted the standard model to be re-named the “ Λ -CDM model.” Here, “CDM” means “cold dark matter,” and Λ , i.e., the Greek capital Lambda, represents dark energy. It is the *cosmological constant*, to which we will return in sections 7 and 9.)

But for the most part, these two ideas, dark matter and dark energy, need not concern us, for two reasons. First, their bearing on our topic, the multiverse, is slight. Most proposals about the nature of dark matter and dark energy do not give reasons for or against a multiverse. In effect, they are compatible with inflation’s producing a multiverse, but do not especially support it.

Second, there is good theoretical reason to think that whatever the detailed physics of dark matter and dark energy turns out to be, it will not overturn the main outlines of what the standard model claims to have established.

This is well illustrated by two of the standard model’s “grand narratives” of the history of the cosmos: the thermal history of the cosmos (i.e., its density, temperature, pressure, etc. at successive stages), and the history of the synthesis of elements, both in the primeval fireball and later in the stars. Indeed, it is a striking testimony to how well confirmed this standard model now is, that for both these narratives, the detailed story given in a technical exposition written today matches closely the detailed story in expositions written some fifty years ago.

So, I will end this review of our fortunate golden age in cosmology by taking as an example the thermal history of the universe. This will set the scene for the next section's description of the puzzling features that prompted cosmologists to postulate an even earlier, very brief period of accelerating expansion: inflation.

2.3 The Thermal History of the Universe

I will give just three “snapshots” of what the temperature, density, and relative size of the universe was, at the following times: (1) a millionth of a second after the Big Bang, (2) a hundredth of a second after it, and (3) ten million million seconds, i.e., about 380,000 years, after it. Note that we are now about a hundred thousand million million seconds, i.e., about fourteen billion years, after the Big Bang.

Before I give the numbers, let me adopt the *exponent* (or *index*) *notation*, using a superscript to indicate the number of zeros. So, one hundred is 10^2 , and a million is 10^6 . Similarly, we use negative exponents to represent reciprocals, i.e., 1 divided by a larger number. So, one hundredth is 10^{-2} , and a millionth is 10^{-6} .

This notation raises another important point. It will be helpful (though I admit, it is difficult) to think *logarithmically*, *not arithmetically*: to think, for example, that since the present time is about 10^{17} seconds after the Big Bang, the time $t = 10^{-17}$ seconds before the Big Bang is *as much before* $t = 1$ second, *as we are after it*.

Though this sounds blatantly wrong, the rationale for it is that a great deal of physics is a matter of *scales*. That is, if you change the situation you wish to describe by a factor of about 10, you may well need a very different description, and this is even more likely if you change by a factor of about 100. This trend holds whether the quantity whose value you change is time, distance, energy, or temperature. So, when cosmologists puzzle over what was the state of the universe, at say $t = 10^{-6}$ seconds, or how physical processes changed as a result of the cooling between, say, $t = 10^{-6}$ and $t = 10^{-2}$ seconds, we should not accuse them of straining at gnats, i.e., of foolishly concentrating on events that are so transient that they cannot be very important for the physics. For, agreed, the universe was changing unbelievably rapidly (arithmetically speaking!), but the relevant processes change—and so our description must change—in crucial ways, depending logarithmically on the earlier time.

So here are the three snapshots:

1. $t = 10^{-6}$ seconds after the Big Bang: The temperature was about 10^{13} degrees centigrade. This is when protons and neutrons, i.e., the constituents of atomic nuclei, form; for at yet higher temperatures, they “melt” into their own yet-smaller constituents, quarks. The size of the observable universe relative to its size today was 10^{-12} , and the mass density was about 10^{17} grams per cubic centimetre.
2. $t = 10^{-2}$ seconds after the Big Bang: The temperature was about 10^{11} degrees centigrade. Atomic nuclei form, i.e., at higher temperatures, and they “melt” into their constituent protons and neutrons. The size of the observable universe relative to its size today is 10^{-11} , and the mass density was about 10^9 grams per cubic centimetre.
3. $t = 10^{13}$ seconds (i.e., about 380,000 years) after the Big Bang: This is when atoms formed by free electrons combining with nuclei, to form electrically neutral atoms of the familiar kind. The universe thereby became for the first time transparent to electromagnetic radiation. So, for cosmology, this is a crucially important time. For it means that our direct observations of electromagnetic radiation cannot go back any earlier than this time. (But amazingly, we *do* observe this time: the CMB, the remnant radiation from the Big Bang, dates from exactly this time.) It is known as the “recombination time,” though all agree that “combination” would be a much better name, since the electrons and nuclei were not stably combined at any earlier time. The temperature was about 3,000 degrees centigrade. (By way of comparison, the temperature at the surface of the Sun is about 6,000 degrees.) The size of the observable universe relative to its size today was 10^{-3} , and the mass density was 10^{-21} grams per cubic centimetre.

I said that these claims about the universe’s thermal history were now established. But I agree that when presented with these stupendous figures—so high for temperature and density, so tiny for time and size—one, of course, asks: “Is all this really established as fact?”

I think the answer is “Yes.” Of course, the evidence is technical and varied—and I cannot go into details. But I note that physicists’ description

of even my earliest snapshot, i.e., the description of protons and neutrons “melting” into quarks by the standard model of high energy physics, has been confirmed by terrestrial experiments. Indeed, I could have chosen an earlier snapshot. For it is common nowadays to take the boundary between known and speculative physics to be at about 10-11 seconds after the Big Bang. But agreed, there is a spectrum from caution and confidence (as we discussed in section 4 of chapter 1), and one could reasonably be more cautious, even taking, e.g., one second as the start of what one calls “established.”

3. Inflation . . . Eternally

So much by way of celebrating our golden age of cosmology, and its standard model as formulated in the mid-1970s and developed since then, by e.g., the admission of dark matter. In this section, I first describe, in 3.1, two puzzling features of the model, which were recognized by 1980 and prompted the idea of inflation. I introduce inflation in 3.2 and its conjectured mechanism in 3.3. In 3.4, I describe how this mechanism yields a multiverse.

3.1 Two Problems

The two puzzling features are called the “flatness problem” and the “horizon problem.” (I set aside a third puzzling feature, called “the monopole problem”: not just for brevity, but also because inflation’s treatment of it is similar.) For both of them, the problem is not one of empirical adequacy. That is, the problem is not that the standard model from the mid-1970s gets some observational predictions wrong. The problem is that, according to the model, an empirically measured number amounts to a coincidence so enormous that, as the saying goes, “it cries out for explanation.”

So, first, the flatness problem. The expanding solutions of general relativity fall into three classes:

1. Those in which the matter is on average dense enough that gravitation eventually overcomes the expansion, so that there is a contraction and ultimately a “Big Crunch”. This is called a “closed universe.”
2. Those in which the average density is low enough that expansion goes on forever at some non-zero rate (with, of course, lower densities making for a higher final rate). This is called an “open universe.”

3. Those “between” 1 and 2 in that the average density (a) is low enough that gravitation cannot overcome the expansion, but also (b) is high enough that the final rate of expansion is zero. This is called a “flat universe,” since the instantaneous geometry of space, across the whole universe, gets ever closer to being Euclidean—so “eventually-flat universe” would be a more accurate name.

The density in 3 is special. Not only is it the boundary between the regimes 1 and 2. But once a solution has that density, it will have it forever. It is called the “critical density.”

These ideas are often put in terms of the ratio between the universe’s actual density (of course, as usual, taking the matter as smoothed out over all space) and the critical density. So, this number is a pure number. For it is defined by dividing one density by another, and so it has no units. It is written as the Greek letter “Omega,” i.e., Ω .

So, here is the enormous coincidence. We have measured Ω to be now close to 1, and indeed to have been close to 1 at all times later than about one second after the Big Bang. (This means that ever since that time, the universe has been almost flat: its spatial geometry has been almost Euclidean.) But in these solutions of general relativity, any difference of Ω from 1 in the early universe is very rapidly amplified. For example, if at one second after the Big Bang, $\Omega = 1.08$, then already at ten seconds $\Omega = 2$, and thereafter Ω keeps increasing exponentially. And on the other hand, if at one second after the Big Bang, $\Omega = 0.92$, then already at ten seconds $\Omega = 0.5$, and thereafter Ω keeps decreasing exponentially. In short, in these solutions, $\Omega = 1$ represents an equilibrium—but a very unstable equilibrium. In particular, for Ω to be about 1 today requires that it be stunningly close to this privileged value soon after the Big Bang. For example, at one second after the Big Bang, it has to differ from 1 by at most 10^{-16} . This is about the ratio between the width of a human hair (viz. a tenth of a millimetre) and the average distance between Earth and Mars (viz. 225 million kilometres).

Indeed, this is an enormous coincidence, crying out for explanation.

Second, the horizon problem. It has a similar structure. Namely, although the standard model of the late 1970s is empirically adequate, it requires a feature of the CMB to be “just so” to an extreme degree—indeed, to

a degree so extreme that it is implausible to treat it as a brute fact, without explanation.

The problem arises from the fact that the CMB, dating from 380,000 years after the Big Bang, is *very* uniform across the sky. Its wavelength, amplitude, etc., is almost the same in whatever direction you point your telescope: its wrinkles are minuscule. More precisely, their proportional size is 10^{-5} . That is like having on the surface of a pool of water one metre deep, a wave that is only a hundredth of a millimetre high.

One naturally asks why it should be so uniform. And this question becomes all the more urgent in the context of the standard model of the late 1970s. For it says that for two directions in the sky with a sufficient angle between them—about one degree or more (the visual width of the moon or more)—the two emission events of the CMB that lie along those directions (about 13 billion years ago) have no *common causal past*. This means that no yet-earlier event could affect both of the two emission events, by influences travelling to each of them at most as fast as light. That is, there is no event that could influence them both via causal, i.e., no-faster-than-light, processes. Hence the phrase “no common causal past.” Relativity theory has some other helpful terminology for this. Given any event, the set of events to its past that *could* affect it by an influence travelling at most as fast as light, is called the event’s *past light cone*. So, the point is that the standard model says that the two past light cones of the two CMB emission events do not overlap.

This makes our question urgent. For this means there could not have been any kind of interaction between events in the past of the first emission event and events in the past of the second emission event. But since the emission events are so strongly correlated—their quantitative properties differ by at most the tiny factor 10^{-5} —one would naturally expect some such interaction. For think of how we explain various systems with uniform properties throughout their extent (such states are called “homogeneous”). For example, a cup of tea with milk throughout it or an iron bar with its temperature equal along its length. We explain these by a past process of interaction. Namely, the system started in a non-uniform (heterogeneous) state, then the milk spread through the tea and the heat spread along the bar. (A process that ends in such an equilibrium state is called “equilibration.”) But here, the standard model forbids such a process of achieving uniformity by an earlier interaction. For it says that no causal process of

any kind could affect the two signals of CMB coming to us from these two directions in the sky.

In short, the standard model tells us to accept these signals' strong correlation as a brute fact, which is encoded in the state of the universe's matter and radiation at times earlier than the recombination time. That is hard to accept.

Besides, it is all the harder to accept when one calculates that the angle between directions sufficient to imply (according to the standard model) no common causal past is very small. It is only about one degree—the visual width of the moon. For this angle being so small means that, according to the standard model, there are a stupendous number of patches of the sky whose CMB radiations are strongly correlated with each other (to within the tiny factor 10^{-5}), even though there were no interactions in their pasts. That is surely incredible.

So, the flatness and horizon problems have a common structure. They each take a certain feature (Ω and the smoothness of the CMB, respectively) to be *just so*. That is, the feature has a value specified to many decimal places (also called “many significant figures”), without the standard model giving any account of why the feature is so exactly specified. This “just-so-ness” is called “fine tuning.” (Later in this chapter, this phrase will get a more specific meaning in the context of *selection effects*.)

3.2 Introducing Inflation

Enter the idea of *inflation*. It turns out that if we change the standard model by “inserting” into it a very brief and very early epoch of rapid, indeed accelerating, expansion, then we can solve both problems.

The basic idea of both solutions is quite simple. It turns out that whatever the value of Ω at the onset of the inflationary epoch, Ω will be driven close to 1 by the end of the epoch and will remain close to 1 for a very long time thereafter, including until now. And recalling that Ω being one is a matter of a flat Euclidean spatial geometry, we can see the simple idea behind this calculation: an expansion of a highly curved surface makes a local patch flatter. Think of blowing up a balloon, or how the fact that the Earth is large makes our local patch of it seem flat.

The situation is similar for the horizon problem. A suitable inflationary epoch changes the space-time geometry in just the right way. Namely, it implies that the past light cones of all emission events of the CMB—even for points on opposite sides of the sky—do in fact overlap. So, with inflation,

the cosmos' very early space-time geometry allowed for a suitable process of equilibration that made the CMB's properties so uniform.

The details of these solutions also work out well, in the sense that when one calculates how much inflation, and when, is needed to solve these two problems, one gets approximately the same results, despite the two problems being so different. Namely, the solutions are quantitatively correct if we postulate:

- (a.) The inflationary epoch ends at about 10^{-34} seconds (which corresponds to a temperature of 10^{28} degrees Celsius).
- (b.) The inflationary expansion is exponential, and started at, for example, 5×10^{-35} seconds with a characteristic expansion time (i.e., the time in which the radius of the universe is multiplied by about 3) of 10^{-36} seconds.

Taking (a) and (b) implies that in the course of the inflationary epoch, the size of the universe expanded by a factor of about 10^{22} .

Agreed, these are dizzying figures, and the epoch is proposed to occur at times and energies very far beyond those we have confirmed in experiments or observed. So, we are undoubtedly in the realm of extreme speculation, and accordingly, caution is in order. It would certainly be reasonable to give low credence to the idea of inflation, and therefore, to the details in the rest of this section and the next. (But if so, the philosophical discussion of explanation from section 5 onward would still stand.)

3.3 The Mechanism?

Having solved the flatness and horizon problems, i.e., avoided two fine-tunings by conjecturing a process of expansion, one naturally asks: "What caused this expansion (what is its mechanism)?" For one might suspect that unless we can cite a plausible cause, we should conclude that it is just a coincidence that the same quantitative details about the expansion solve both problems. In answer to this question, the advocate of inflation has, as the saying goes, good news and bad news.

The good news is that a mechanism has been formulated. Indeed, there are many proposed mechanisms which, needless to say, remain conjectural. Most of them involve postulating a new physical field (called "the inflaton field," and written ϕ), which evolves, i.e., changes over time according to a

postulated potential energy function, written $V(\varphi)$. And, fortunately, from such a field and potential one can deduce some characteristic features of the CMB: characteristic probabilities for the amplitudes and frequencies of the slight wrinkles (unevennesses) in the CMB. These features have been observed by a sequence of increasingly refined instruments, mostly on satellites (the famous acronyms/names are “COBE,” “WMAP,” and “Planck”). So nowadays, these confirmed predictions are regarded as more important evidence that there was a brief epoch of expansion than the epoch’s solving the flatness and horizon problems.

But there is also bad news. The data we now have, and maybe all the data we will ever have, leave wide open which of the many possible mechanisms—which sort of field φ and which potential $V(\varphi)$ —actually occurred. There are two aspects to this. The data leaves open the formal mathematics, e.g., what is the function $V(\varphi)$. And it leaves open what the physical nature of φ is: it is widely believed *not* to be one of the known fields.

So far in this section, we have reviewed two problems that were solved by the idea of an inflationary epoch and broached the question of what mechanism led to that epoch. Now we are ready for the punchline, that is, the punchline for someone interested in the multiverse.

3.4 Spawning Universes

In explaining this punchline, I will expand on the summary I gave in this chapter’s preamble and section 1, but postpone details about string theory until the next section.

It turns out that many models of the inflaton field and its potential involve a *branching structure* in which, during the epoch, countless space-time regions branch off and then expand to yield other universes. Here, “branch off and then expand” means that the region stops its accelerating expansion and expands only slowly, like our observable universe does. (Note that I said “many models,” so again, caution is in order.) As a result, the whole structure gives a multiverse, whose component slowly expanding universes cannot now directly observe (nor otherwise interact with) each other, since they are causally connected only through their common origin during the inflationary epoch. (Incidentally, a pair of neighbouring points on the inflating background space also separate from each other so rapidly as to lose causal contact; so, such pairs also “soon cannot see each other.”)

Besides, in a universe that branches off—called a “bubble,” “domain,” or “pocket universe”—yet another universe can branch off; and also from

that one, there can be a branching . . . and so on. In short, bubbles (domains) spawn more bubbles, endlessly. This idea of open-ended, maybe infinite, branching towards the future is called “eternal inflation.”

Of course, all this is a quantum process. So, the quantum state of the cosmos (in the cosmologists’ sense, explained in section 1) contains components corresponding to (amplitudes for) the different bubbles. And one is committed to the bubbles being real, if one accepts that to a non-zero (or at least large enough) component/amplitude, there corresponds something real. (This acceptance was (i) in section 1.1, and as I said there, it does not imply an Everettian view, since it is compatible with dynamical models of wave-function collapse.)

There are two main types of model that yield eternal inflation, labelled “false vacuum” and “slow roll.” In both types of models, the inflationary expansion coming to an end, i.e., the beginning of a slow expansion, is a matter of the inflaton field evolving, i.e., changing over time, to its state of lowest energy. And as discussed in the preamble and section 1, such states are called “vacuum states,” or for short, “vacua” or “ground states.” (So again, “vacuum” does not mean “nothing” or “no physical system.”)

By and large, physical systems tend to lose energy and evolve to their vacuum states. This is also the case here. Thus, eternal inflation is a matter of the prevention of the inflaton field’s tendency to get into its vacuum state (which would render the expansion slow). The system being prevented in this way is usually called its “being *frustrated*.” Thus, a region where this frustration does *not* occur is where the new bubble branches off. And since the region where there *is* frustration continues to expand exponentially, there is very soon a vastly larger region in tiny patches of which new branching will occur. Hence, eternal inflation.

4. Glimpsing the Landscape of String Theory

So much by way of introducing the idea of eternal inflation. In this section, I turn to string theory, building on the preamble and section 1, but again, necessarily omitting a lot of advanced physics. I will confine myself to just three topics. First, in 4.1, I will state the initial idea of string theory. Then, in 4.2, I will develop the idea that it has many vacuum states (each with an associated “tower” of excited states) that differ in the constants of nature. Then, in 4.3, the dauntingly large number of vacuum states will prompt us to face philosophical questions about explanation and confirmation, which will dominate the rest of the chapter.

4.1 Quantizing a String

String theory is a speculative attempt that began in the mid-1980s to unify general relativity's successful account of gravitation with quantum field theory's successful account of nature's other fundamental forces. (These are the electromagnetic forces between charged particles, and two other forces between subatomic particles such as electrons, neutrinos, and quarks, which are called the “weak” and “strong” forces.) The sense in which string theory aims to unify these forces is like the unification of electric and magnetic forces that Maxwell achieved (section 1 of chapter 4). Roughly speaking, the four apparently diverse forces are to be revealed as aspects of a single force.

We can glimpse why this is hard to achieve by looking at string theory's initial key idea. Namely, it “does for a string, what elementary quantum theory did for a point particle.” (Hence, its name.) Thus, recall from section 2 of chapter 4 that elementary quantum theory replaced the state of a classical point particle—in effect, its single actual position—by an entire function on all possible such positions, mapping each position to a “square root” of a probability (an amplitude) to be found there, if measured. Now, a point particle is extensionless and can be thought of as zero-dimensional, while an infinitely thin line, a mathematical curve, is one-dimensional. So just as we can think of classical point particles as idealizations of tiny spheres, we can think of an infinitely thin line as an idealization of a thin filament—a string, though without the spiral threads. Such a string has, of course, no single position. Each of its constituent points has a position, and the configuration of the string as a whole is the infinite set of those configurations, which we might call a “placement” of the string. So, the quantum replacement of the state of this classical string is a function on all possible placements, mapping each placement to an amplitude . . . No wonder that advanced physics is needed.

4.2 The Landscape

But for this book's purposes, all we need is the upshot: that string theory predicts the system, i.e., the set of all the quantum strings, has very many *different possible vacuum states*. (This was realized around 2000; until then, string theorists had hoped there was only one vacuum state.)

Note that here, the terminology can be confusing. For a state that is not the overall lowest-energy state (the state with energy lower than *all* others), but is only a local minimum with an energy lower than all its near

neighbours in the state space, is often called a “false vacuum.” (As I mentioned at the end of the last section, this term is used in cosmology, as the name of one main type of eternal inflation.) But it is “false” only in the sense that the minimum is *local*. So, this term is rather like calling a valley in a mountain range a “false valley,” just because it is higher above sea level than the lowest valley in the entire mountain range. But this analogy with valleys and peaks has also prompted a more helpful term. In string theory, the whole state space, i.e., the set of states of the quantum strings (varying in energy, with some lower and some higher), is called “the landscape.”

(Incidentally, biology makes an analogous use of “landscape.” In the theory of natural selection, the attributes of an organism or of a population of organisms make it more or less fit, where fitness is, roughly, a matter of having more offspring who live long enough to reproduce. Over time, natural selection, “the survival of the fittest,” increases the proportion of fitter organisms. So, over time, the population (descendants of the original organisms) gets a higher fitness “score.” So, the population “climbs to a peak in the fitness landscape.” So, high fitness is analogous to high energy: the biology-physics difference is that in biology, fitness increases over the generations, while in physics, an individual system tends to a lower energy—to a valley, not a peak. From the philosophical perspective of chapter 3, the interesting point here is, of course, that this is another vivid illustration of our science being up to its neck in modality: almost all positions in the fitness landscape are not inhabited by a real organism or population—it is a realm of possibilities.)

Returning to physics, each of string theory’s vacuum states has a tower of associated higher energy states. The idea goes back to section 3 of chapter 4’s comment that in quantum field theory, particles are really energetic excitations of fields. There, the tower was called “Fock space.” It is a sum of infinitely many subspaces. Namely, first, the space containing the given zero-particle, i.e., lowest-energy state; then, the space of one-particle states; then, the space of two-particle states; then, the space of three-particle states, and so on. But the idea of particles as excitations is more general than this simple sum-of-subspaces structure of Fock space, and it carries over to string theory.

In string theory, all the various higher states generated by exciting a *given* vacuum state will share with it the values of physical parameters, like the electric charge on an electron, the ratio of strengths between the electromagnetic and gravitational forces, or the speed of light. (We tend to call

the charge on an electron a “physical parameter,” not a “physical quantity,” since we think of the value of a quantity, e.g., the position or energy of a system, varying across the states within a single tower, while the value of a parameter is the same for all states in a tower.)

But for different vacua, these values will vary. This is so even for parameters that we take to be constants of nature, i.e., constant in value across the observable universe, like the examples just listed. (This uniformity, this “geographical” unity of the observable universe, is itself very remarkable. *A priori*, there could well be regions of the universe in which the charge on the electrons, or the relative strengths of the forces, or the speed of light signals, is different from what we measure hereabouts.)

4.3 Far Too Many Vacua

This variation prompts an ambitious but alluring project. For we now have a framework so broad that it encompasses scenarios (formally: towers of states above certain vacuum states) that differ from each other in the values of parameters that we usually call “constants of nature,” though this framework means we are now envisaging that they vary across a wider landscape. So, this suggests we invoke this framework to answer the obvious big “Why?” question about such a parameter, namely, “Why does it have the value that it does?” That is, let us try to find features in the framework that, in some sense, favour and thereby explain the value.

But this project runs up against a major problem, a problem that suggests the project will stumble, unless one appeals to some philosophically contentious ideas. For it turns out that there is a dauntingly large number of vacuum states. And this problem confronts not only string theory taken alone, regardless of cosmology and inflation. It also confronts inflationary cosmology.

Thus, a recent estimate of the number of string theory vacua is 10^{500} . This is enormously larger than all the numbers in more established branches of physics. For example, the number of elementary particles in our universe, i.e., setting aside the cosmological multiverse, is estimated to be about 10^{100} . So, the number of string theory vacua is larger by a factor, not of 400, but of, 10^{400} . Similarly, in the cosmological multiverse, estimates of the number of bubble universes give vast numbers.

Obviously, to explore this set of states, this landscape—i.e., to understand it in quantitative detail, classifying valleys and peaks—is forever beyond human, or even superhuman, ability.

In the face of this impossibility, the project envisaged above, of explaining the values of the parameters, seems to stumble. For as I said in this chapter's preamble, the obvious overall strategy for getting such an explanation would be to argue that the value we measure is in some precise sense *generic* or *typical* of the values across the multiverse, and is thereby to be expected. One aims to explain the actual value we see by showing that it is typical and to be expected. But how can we do that, without understanding the set of states (the landscape, the towers above the vacua) in quantitative detail?

The rest of this chapter will discuss suggestions for how to do this—albeit contentious ones. Section 5 sets the scene by discussing explanation in general and formulating: first, the obvious strategy above, which I will label “strategy (Gen)” (for “generic”); and another, “strategy (Obs)” (for “observation”), which invokes selection effects. Subsequent sections will treat these strategies in order.

5. Angst About Explanation

Let us, for a moment, take a step back from the details of physics and ask: How does one explain any fact? How does one answer any “Why?” question?

There is a large philosophical literature on explanation, with rival accounts of what an explanation is and what role explanations fulfil in the enterprise of science. But for our purposes, these accounts' agreements matter more than their disagreements, so this section will, for the most part, summarize the agreements. But unfortunately, these agreements will not settle the questions raised at the end of the last section. So, those questions will have to wait for the next two sections.

The accounts agree that in everyday life, what counts as a correct or appropriate answer to a “Why?” question obviously depends strongly on what the enquirer (and no doubt, also the respondent) knows, what their interests are, etc. These accounts also agree that such contextual and pragmatic features also apply to scientific explanation.

They also agree that in both everyday life and science, there is a spectrum of requirements one can impose on the answer, along the lines of whether it must be believed by the respondent, or must be true, or even must be known to be true. Again, it is a contextual and pragmatic matter which requirement lying on this spectrum we should impose on a would-be explanation in order for it to count as a genuine explanation.

These accounts of explanation also agree on some helpful terminology from Latin. The fact to be explained (or the proposition expressing the fact) is called the “explanandum,” and what does the explaining (or the proposition expressing what explains) is called the “explanans.”

More substantively, and relevant to us, they also agree that what facts count as *needing* explanation is a contingent, and often historically determined, matter, no less in science than in everyday life.

These are not just the obvious points that explanation must come to an end somewhere, and that where the respondent’s chain of explanations terminates depends on their state of knowledge, which is a contingent and historically determined matter. (And as every five-year-old who persistently asks “Why?” learns, where the chain of explanations terminates can depend on the parent’s inventiveness or patience.)

There is also the more interesting point that acceptance of a scientific theory (or more loosely, of a research tradition or framework) can influence, even determine, which sorts of facts are taken to need explanation, and which are not. This point merits two examples.

A standard example of this from the history of physics is Kepler’s endeavour (in his *Mysterium Cosmographicum* of 1596) to explain the relative sizes of the planets’ orbits, and the number of known planets (viz. 6), by interpolating the five Platonic solids between the orbits. (The known planets were Mercury, Venus, Earth, Mars, Jupiter, and Saturn.) Thus, Kepler believed that such a major structural feature of our solar system should have a systematic explanation. But nowadays, we accept, not just that there are more planets—Uranus was discovered in 1781 and Neptune in 1846—but that the sizes of the orbits are “merely” accidents of the history of the solar system. They no doubt have a very complicated causal explanation—if only we could know it. The explanation would involve the various radii (i.e., distances from the sun) at which the planets were first formed, how they interacted gravitationally, etc. But these are (at least for the most part) matters of sheer happenstance, about how the solar system happens to have evolved. We do not expect the number of planets, or their orbits’ sizes, to have any general or systematic explanation. (Needless to say, this is not to disparage Kepler’s endeavour. Given his overall world-view, that there are six planets orbiting the Sun is a main, even pre-eminent, fact about the solar system about which it is very natural to ask “Why six?”)

Nor is it only particular facts that can come to be seen as not needing a general or systematic explanation. Very general patterns of behaviour can

also fall out of the purview of explanation. (I say “patterns of behaviour” to set aside controversy about laws of nature (cf. section of chapter 3). But as we will see, the pattern might well be called a “law” of a given theory.) A standard example of this is the idea of “natural motion.”

In the long history, since the ancient Greeks’ geometry and astronomy, of the precise quantitative description of motion, “natural motion” is an inevitably vague term. But the rough idea is motion that needs no explanation, since the body is “moving without interference.” Thus, in Aristotelian cosmology, the natural motion for the element Earth (one of four, the others being air, fire, and water) was downward—towards the Earth. But by the mid-seventeenth century, the mechanical philosophers (cf. section 2 of chapter 2) maintained that natural motion was motion in a straight line at constant speed—it was other motion that was “forced.” For example, a body’s accelerating toward the Earth was due to the Earth’s gravitational force, and a block’s slowing as it slid down an inclined plane was due to friction with the plane. This came to be called “the principle of inertia.” It was given its first clear formulation by Descartes, and later, it was Newton’s First Law of Motion—that a body subject to no force at all moves in a straight line at constant (maybe zero) speed.

Thus, for both the mechanical philosophers and Newton, the motion of a projectile in a straight line and at constant speed (neglecting gravity and air resistance) needs, in a sense, no explanation. Agreed, one can ask what set this projectile moving, i.e., what launched it. I also agree that the motion, once underway, is an instance of the principle of inertia, and so it can be deduced from the principle. But the motion, once underway, needs no explanation in the sense that no *causes* need to be cited. It is enough that the motion instantiates, and can be deduced from, the principle of inertia. (Here, my “it is enough” deliberately echoes Hume’s and Newton’s lowering our sights about the rationalist understanding of nature, discussed in section 5 of chapter 2).

So much by way of examples. The final issue on which these philosophical accounts of explanation also agree is the core idea of explanation: namely, that a *successful explanation shows that the explanandum was to be expected*.

Here, I choose the words “successful,” “shows,” and “to be expected” deliberately. Thus, the first two words signal how my formulation of the agreement between these accounts deliberately steers clear of some controversies

that—irrelevantly for this book’s purposes—dominate the philosophical literature, as follows.

I say “successful” in order to signal flexibility about pragmatic factors such as the enquirer’s interests, and about whether the *explanans* must be true or even known to be true.

I say “shows” in order to signal flexibility about whether (i) there must be an outright *deduction* of the (proposition expressing the) *explanandum* from the propositions comprising the *explanans* (which would be a strong sense of “show”); or (ii) it is sufficient to render the *explanandum* probable (usually in the sense of having a high enough probability, conditional on the *explanans*).

(Of course, there are other controversies I have not touched on. For example, must an explanation of an individual event or fact (as against a general proposition) cite the causes of the event or fact? And is explanation fundamentally contrastive, i.e., about answering “Why A rather than B?” not just “Why A?”?)

On the other hand, my third deliberately chosen phrase, “to be expected,” signals a return to the questions at the end of the last section. I deliberately choose a phrase that is ambiguous; and ambiguous in a way crucial to our concern with multiverse proposals’ endeavour to explain the values of parameters such as constants of nature. For the phrase can be understood as referring to either one of two different strategies for explaining some fact, in particular, explaining some apparent fine-tuning of a parameter’s value. These strategies are:

(Gen): Showing that the fact either is deducible (from the *explanans*) or is generic or typical, i.e., roughly speaking, one of the alternatives that has high enough probability.

(Obs): Showing that although the fact is not deducible and is *not* even generic or typical, it is likely (or has high enough probability) *to be observed*. (As I mentioned in this chapter’s preamble, this strategy invokes “selection effects,” or (a better-known term) “the anthropic principle.”)

The next two sections explore strategy (Gen). The subsequent sections explore strategy (Obs).

6. Expected Because Generic

In this section, I discuss strategy (Gen) in general terms, without considering the multiverse—though with examples from physics. I will first, in 6.1, cast inflation’s answer to the flatness and horizon problems—problems of fine tuning in a single universe—as an example of strategy (Gen). Then, in 6.2, I mention some other examples. This will prompt a more general statement of what fine tuning amounts to. Then, in 6.3, I will give a bit more detail about three ways one can make precise the idea that the value of a parameter is to be expected, because it is generic. I put them under the labels “topology,” “effective field theory,” and “probability.” This section will emphasize the first two of these. But probability will be a large topic for us, also in connection with strategy (Obs). So, although I will introduce it here, the details will be postponed to subsequent sections.

The overall shape of this section will be to start with fine tuning as a problem, and to end with three approaches to answering the problem by saying that the parameters’ value is in fact generic. Thus, the tone of this section’s assessment of strategy (Gen) will be positive. With the examples and approaches considered here, the strategy has successes. But in section 7, the difficulties that the strategy faces will move to centre-stage.

6.1 *The Flatness and Horizon Problems as an Illustration*

First, let us recall the flatness and horizon problems from section 3 above. In both, a certain feature (Ω and the smoothness of the CMB, respectively) had to be *just so*. That is, the feature has a value specified to many decimal places, without our cosmological model (i.e., the model that was standard in the 1970s) giving any account of why the value is so tightly constrained—in short, fine-tuning. As we saw, inflation solved these problems by changing the theoretical context so substantially that the required values could arise through an (admittedly, conjectural) dynamical process, from generic initial states. For a suitable inflationary epoch drives Ω to become close to 1 by the end of the epoch; and thereafter, the standard cosmological model’s (non-inflationary) dynamics makes Ω remain close to 1 for a very long time, including until now. Similarly, a suitable inflationary epoch makes the past light cones of all emission events of the CMB—even for points on opposite sides of the sky—overlap.

We can now construe this discussion in terms of the last section’s ideas about a successful explanation showing that the *explanandum* is to be

expected, either by deduction or by getting a high enough probability, from the *explanans*. And since it is *generic* initial states (i.e., states before the inflationary epoch) that lead to Ω being close to 1, and to the past light cones overlapping, the explanation is *insensitive* to what exactly the initial state is. Such an explanation (or deduction or calculation of high probability) is often called “robust” or “stable” or “resilient.” In short, here, the phrase “to be expected,” in “showing the *explanandum* is to be expected,” has the straightforward sense (Gen) at the end of the last section.

These examples of inflation solving the flatness and horizon problems also illustrate other ideas from the last section, such as that:

- (i.) explanations inevitably come to an end somewhere; and
- (ii.) the theoretical context moulds one’s judgments about what is generic or probable (or probable enough to count as being explained), and what is not.

For after all, one can ask: (i) What explains (and/or what caused) the pre-inflation state, no matter how generic or probable one accepts it to be? And (ii) what justifies one’s judgment about what is a generic or probable enough pre-inflation state? (I will shortly return to this topic of making precise the idea of a state of the whole cosmos being generic or probable.)

But although one can raise these questions, inflation’s account of why Ω is close to 1, namely, as a robust feature of a dynamical mechanism, is generally agreed to be a successful explanation, even though the dynamics is very conjectural. In short, it is a successful example of strategy (Gen).

6.2 The Strategy in General

Besides, there are several other examples in physics where a fine-tuned value gets explained as generic by a suitable change in the theoretical context. One such case was the explanation in the early 1970s of some fine tuning in sub-atomic physics by postulating a new kind of particle (viz. a charmed quark), which was later empirically confirmed.

Indeed, such examples fall into a wider category, of explaining a value (not necessarily fine-tuned) of a parameter by a suitable change in the theoretical context—but not necessarily by showing the value to be generic or typical, e.g., by having high or at least moderate probability.

A famous case of this wider category—with the merit that it is simple enough to describe—is Maxwell’s explanation of the speed of light. When Maxwell formulated his unified theory of electricity and magnetism (mentioned in section 1 of chapter 4), he found that some solutions to his equations described waves of the electric and magnetic fields that his theory postulated. That is, the theory described oscillating patterns of (values of) these fields (vectors located at points in our familiar three-dimensional physical space). These patterns propagated across space at a speed that is a simple function of two fundamental constants (called “permittivity” and “permeability”) that are mentioned in the theories of electricity and magnetism, and whose values were known. When Maxwell calculated this simple function from the two known constants, the answer turned out to be the speed that had already been measured as the speed of light (viz. 300,000 kilometres per second). Maxwell then inferred that light *is* waves of the electric and magnetic fields.

This is often, and rightly, celebrated as a reduction of one field of physics, the theory of light, i.e., optics, to another, the theory of electromagnetism (as we now call it). (Here, “reduction” is meant in the sense of section 1 of chapter 3, viz. deriving the doctrine of one theory from that of another by augmenting the latter with suitable definitions.) But once light is indeed identified as being such waves, Maxwell’s calculation can also be taken as an *explanation* of the speed of light. In effect, the explanation is: “these waves of the electric and magnetic fields must travel at this simple function of the permittivity and permeability constants; and given the actual values of those constants, the speed must therefore be 300,000 kilometres per second—as is observed.”

So much by way of examples of successfully explaining the value of a parameter. I turn now to formulating a bit more generally what fine tuning really amounts to: What is the problematic “just-so-ness” of a parameter’s value? The idea will be that the parameter should not be a function of other parameters, one that depends very sensitively on those other parameters’ values.

To take perhaps the simplest example: The value of a parameter should not be an *arithmetical difference* of two other physically significant numbers that are nearly equal but are both vastly larger in magnitude than the parameter itself.

Thus, imagine a theoretical framework in which the chosen parameter, which I call p , is a millionth: $p = 10^{-6}$. And imagine this is “because” (i.e.,

because, according to the given framework) p is the difference of two other numbers, q and r , that themselves have some physically significant interpretations, and that are nearly equal but are both vastly larger than p . For example, they might have values $10^6 + 10^{-6}$ and 10^6 . That is, $q = 10^6 + 10^{-6}$ and $r = 10^6$ and $p = q - r = 10^{-6}$.

So, the imagined framework makes the value of the parameter p *fine-tuned*. It is extremely sensitive to the exact values of these other numbers q and r : in my example, sensitive to their thirteenth digit. Had q and r been slightly different (in terms of proportions of their actual values, e.g., in their last digit), then p 's value would have been vastly different (proportionately) from its actual value. In short, the framework, with its equation $p = q - r$, gives us an unsatisfactorily *fragile* derivation of p 's value, not a robust explanation of it.

(Of course, since the value of a parameter usually depends on a human choice of units, these numbers 10^{-6} etc. should be *dimensionless*. That is, they should be pure numbers without a physical unit involved. For example, they could be a ratio of two masses, or of two densities, or of two electric charges, or of the strengths of two forces.)

This example illustrated the idea of sensitivity with an arithmetical difference being tiny and so liable to be vastly (proportionately) changed by a change in the numbers whose difference is being taken. But as I said, fine tuning need not be a matter of an arithmetical difference. The function involved, whose values depend very sensitively on its arguments, could be a function very different (in particular, more complicated) than addition. All such cases of fine tuning prompt strategy (Gen): The value of a parameter should be shown to be generic or typical in some precise sense that is defined by an appropriate theoretical framework. (Of course, this framework is often not the one in which the parameter's value is first observed or known. Recall, for example, how Ω was measured to be close to 1, before the framework of inflation was suggested.)

But can we make precise this idea of being generic, in a more *general* way? That is, can we do so without invoking case studies, with their case-specific functions, like arithmetical difference and case-specific theoretical frameworks that show the value to be generic? ("Being generic" is sometimes called "genericity," a word so ugly that I avoid it.)

6.3 Three Approaches to Implementing the Strategy

In my opinion, mathematics and physics provide three overall approaches to doing so. I suggest the labels *topology*, *effective field theory*, and *probability*. I will discuss the first two, which are comparatively specific to mathematics and physics. Then I will briefly discuss probability, which, of course, extends far beyond mathematics and physics, and which will also occupy us in subsequent sections.

6.3.1 TOPOLOGY

Topology is a major branch of pure mathematics that focuses on the idea of a *continuous transformation*, which means, roughly speaking, a transformation that preserves the nearness relations holding between the objects being transformed. Here, being near need not be a matter of numerical distance. It can be a qualitative relation, and it can come in degrees. Thus, topology uses terms like “closeness” and “neighbourhood,” and a set of objects endowed with such nearness relations is called a “topological space.” Thus, for a transformation T that shifts objects a and b respectively to $T(a)$ and $T(b)$ (in the same or a different topological space), we say that T is *continuous* if, whenever a and b are near, so are $T(a)$ and $T(b)$. On the other hand, T is *discontinuous* if there are objects a and b that “get pulled apart” by T , i.e., are such that $T(a)$ and $T(b)$ are *not* near.

In this way, a discontinuous transformation can express the idea of “sensitive dependence on the inputs” (here: a and b), even without invoking number-valued functions. Or more precisely, it expresses one version of this idea, without invoking number-valued functions. Applying ideas like these, mathematicians have made precise the idea that an object a in a certain set of objects $\{a, b, \dots\}$ is generic in the sense that it is like (in appropriate respects) the other objects in the space that are near it.

Mathematicians have even defined topological spaces whose objects, i.e., elements $a, b, \dots$ of the set, are possible physical systems, each taken as subject to certain forces. (Since specifying a system and the forces on it prompts the traditional format of a physics problem, viz. “For a given initial condition, how will this system change over time?” such spaces are often called “spaces of *problems*.”) In such cases, the elements of the space, i.e., the physical systems are usually described by a mathematical function, such as a potential energy function that encodes the forces on the system. So, nearness of the elements of the space is a matter of the system’s having potential

energy functions that are “nearly the same,” according to some criterion for the approximate equality of functions.

I will not go into these ideas in more detail. But I cannot resist noting that: (i) the terminology of the subject includes the alluring phrases, such as “catastrophe theory” and “structural stability”; and (ii) since the elements of such a space are *possible* physical systems, very few of which are actual, we again see that physics is up to its neck in modality (cf. section 3 of chapter 3).

6.3.2 EFFECTIVE FIELD THEORY

I call the second approach to making precise the idea of being generic, or typical, the *effective field theory approach*. It is like the topological approach, in two ways. It eschews probabilities (which will be the third approach). And it describes a set of complicated entities with such mathematical precision, both about each entity and about their mutual relations, that the set deserves the name “mathematical space.” (As discussed in section 3 of chapter 3 and section 2 of chapter 4, mathematicians call a set that is endowed with various structures, especially structures inspired by geometric or visual intuition, a “space,” even though its elements have nothing to do with points or regions of physical space.)

But there is also a contrast with the topological approach. There, the entities were physical systems, each taken as subject to certain forces (as noted, they are often called “problems”). But in the effective field theory approach, the entities, the elements in the mathematical space, are *physical theories*, where each theory is identified by, roughly speaking, the set of parameters that occur in its specification of the forces on the systems the theory describes. Or a bit more precisely, the forces are encoded in a special function: the Lagrangian. Or in some formulations, they are encoded in a function that is a mathematical “cousin” of the Lagrangian, viz. the Hamiltonian. Both the Lagrangian and the Hamiltonian functions have, as arguments, states in the system’s state space; so, they are functions on state space, and their value is a certain difference, or a certain sum, of different kinds of energies of the state.

(Of course, only a few Lagrangians (Hamiltonians) will be instantiated by an actual physical system. So, most of the elements in this space of theories are not actualized. So, as in my comment (ii) at the end of the topological approach, we are up to our necks in modality.)

The Lagrangian (or Hamiltonian) contains parameters, especially those whose value specifies how strong a force is (called “coupling constants”).

The list of these parameters' values specifies the theory. That is, it specifies the element in the postulated space of theories.

The point of postulating this space of theories—and the link to expressing our topic of being generic—lies in the key idea that the parameters (including, despite their name, the coupling constants) are *not* really constant. For at high energies, they take different values than at low energies.

So, as you mentally traverse a curve in the space of theories, from higher energies to successively lower energies, you can consider the functional dependence of a parameter's values on various other parameters. In the mathematics, traversing such a curve is a matter, in effect, of summarizing the influence of the physical phenomena at higher energies (the physics described by points, i.e., theories, on the curve that we have already traversed) on the lower-energy physical phenomena that are described by the point, i.e., theory you are currently at—the theory you are currently considering. Or a bit more precisely, “influence” here means “mathematical implications,” not “causal effects”; and “summarizing the influence” is a matter of averaging the numerical values implied by the higher-energy phenomena. Or, in yet more technical terms, it is a matter of integrating out higher-energy modes of the system.

Traversing such a curve is called “following the *renormalization group flow*.” For the way that a parameter's value changes as we consider lower energies is described by a mathematical structure called “the renormalization group.” And the approach is called “effective field theory” because in physics, “effective” does not mean “efficacious,” i.e., “having a big or strong or intended effect,” but “approximate in a useful way.” So, in physics, an “effective theory” is a theory which is believed, and often known, to *not* be completely correct, but which is correct to a sufficient approximation that it is useful.

So, the overall idea here is that although we do not know, and may never know, the correct theory of physics at the higher energies that our experiments cannot probe, we can hope to formulate an effective theory of physics at the lower energies that our experiments *can* probe. (The reason for saying “effective field theory,” and not just “effective theory,” is that these ideas were developed (in the period from 1965 to 1975), mainly in the context of quantum field theory, and its topic of renormalization. In these developments, Ken Wilson (1936–2013) was a leading light.)

Besides, traversing a curve from high to low energies, in the above way, amounts to deducing which low-energy theory (low-energy point on

the curve) is implied by the high-energy theory (point) at which the curve began. More precisely, the clause “is implied by . . .” means is implied by (i) the high-energy theory, taken together with (ii) the chosen way of summarizing the influence of higher-energy phenomena—the way that the curve defines.

So, the question arises: Do the curves along which one mentally proceeds, from two different (though close) points, i.e., theories at some high energy, toward lower energies, diverge or converge?

If they diverge, that means that small differences in one or other of the parameters of the high-energy theory from which one started will imply large differences in one or more parameters of the low-energy theory that one arrives at. That is, divergence means low-energy physics, i.e., the physics we can now observe, is extremely sensitive to the values of at least one parameter describing high-energy physics, i.e., the physics we cannot now, and might never, probe with our experiments. So, divergence is bad news. For it means fine tuning of one or more parameters of the low-energy theory, and we might never be able to probe the high-energy physics on which the parameter depends.

On the other hand, convergence of the curves would mean that the values of parameters describing low-energy physics are *robust* to variations in high-energy physics. They stay approximately the same when we envisage different high-energy theories, even substantially different ones. This is good news, in that we can hope to argue, even without knowing the correct high-energy theory, that the value of a parameter at low energies should be (approximately) thus-and-so. In short, we can hope to argue that what we see is *generic*. For whatever the details of the unknown high-energy physics, we would see approximately this value.

6.3.3 PROBABILITY

Finally, I turn to the third approach to making precise the idea of being generic or typical. Of the three, it is by far the oldest; one might even call it “venerable.” The idea is to appeal to *probability*. There should be a probability distribution over the possible values of the parameter, and the actual value should not have too low a probability.

This connects, of course, with statistical inference—in both everyday life and science, far beyond physics. There, it is standard practice to say that if a probability distribution for some variable is hypothesized, then an observation that the value of the variable lies “in the tail of the distribution”—often

called: “has a low likelihood,” i.e., a low probability, conditional on the hypothesis that the distribution is correct—disconfirms the hypothesis that the distribution is the correct one. That is, it disconfirms the hypothesis that the distribution truly governs the variable.

This scheme for understanding typicality seems to me, and most interested parties—be they scientists or philosophers—sensible, perhaps even mandatory, as part of the scientific method. But agreed, questions remain about:

- (i.) how far under the tail of the distribution—how much of an outlier—an observation can be without it disconfirming the hypothesis, i.e., without it being deemed to be atypical;
- (ii.) how in general we should understand “confirm” and “disconfirm,” e.g., whether in Bayesian terms or in traditional (Neyman-Pearson) terms; and relatedly:
- (iii.) whether the probability distribution is subjective or objective; and more generally:
- (iv.) what probability really means (see sections 11 and 12 of chapter 4); and, after Hume’s critique of his predecessors (see sections 4 and 5 of chapter 2), what the philosophical justification for induction is.

But these questions are obviously not specific to physics, let alone our more specific topic of the cosmological multiverse. So, I will not pursue them in general terms. But this is not to suggest that they are easy or irrelevant to our topic. We will see them crop up several times in what follows.

7. Difficulties About Being Generic

The previous section described strategy (Gen), i.e., explaining a parameter by showing it to be generic, and some of its successes. In this section, I report some of the difficulties it faces. I will begin, in 7.1, with some examples, and then move, in 7.2, to general issues. 7.3 will focus on cosmology. These difficulties will prompt us, in the following sections, to consider the other strategy, (Obs).

7.1 Recalcitrant Examples

First, we should note that in recent decades in physics, strategy (Gen)’s “track record” has been mixed. The previous section noted some successes: especially inflation’s explaining Ω and the homogeneity of the CMB, and the charmed quark. But by no means every apparently fine-tuned parameter has been explained along the lines of strategy (Gen). I will report two such recalcitrant examples. Both examples fall under what the last section called the “effective field theory approach.” So, they are examples of dismaying fine tuning. Small differences in one or more parameters of a high-energy theory imply large differences in one or more parameters of the low-energy theory whose predictions we can observe. As in the previous section, one example is from cosmology, the other is from high-energy physics. As we will see in later sections, both examples have prompted some physicists, in particular, cosmologists, to shift to strategy (Obs).

The first example is the *cosmological constant*. Introduced by Einstein as a possible emendation of his field equations for general relativity and written as Λ (i.e., the capital Greek letter, lambda), this constant represents a repulsive force between two masses. So, in a cosmological context, Λ amounts to a cause or tendency for the universe to expand; so, it opposes the gravitational force that tends to make matter clump together. And the eventual destiny of a universe that, like ours, is in fact expanding—whether to expand forever, or to come to a stop and re-contract—will be determined by the balance between Λ and gravitation. The current evidence is that the universe’s expansion is accelerating. This means that we measure Λ to be positive.

However, we have no good explanation, following strategy (Gen), of the value of Λ , or even of its approximate value. Worse, theoretical estimates of the value of Λ (using the framework of quantum field theory) are wildly wrong. For, unless we assume a lot of fine tuning, the estimates are wrong by very many factors of ten. In some estimates, the error is a factor of 10^{120} (that is, 120 factors of ten, called “120 orders of magnitude”). This discrepancy is called “the cosmological constant problem.” It is, of course, agreed to be a major problem for physics. Regardless of one’s philosophical views about explanation, and, in particular, about strategy (Gen), it suggests a fundamental conflict between quantum theory and general relativity. (But as mentioned, we will see that strategy (Obs) fares better in dealing with it.)

The second example is the mass of the *Higgs boson*. (This particle, first postulated in the mid-1960s, was discovered in 2012 at the particle

accelerator at CERN, Geneva.) This is another dismaying example of fine tuning. Thus, recall the scenario in the middle of the last section: a parameter p is defined as a tiny arithmetical difference of two other numbers q and r , each of which is vastly larger than p and also has an appropriate physical interpretation. This scenario implies that the value of p is very sensitive to (changes enormously with) changes in q and r . My toy example was $q = 10^6 + 10^{-6}$ and $r = 10^6$, so that $p = q - r = 10^{-6}$.

The mass of the Higgs boson is just such a parameter p . The exact value of the exponent depends on which higher value of energy one envisages the (unknown) higher-energy theory being valid; in other words, it depends on how far beyond energies that we can observe one expects new physics to “kick in.” The higher this value of energy, the larger the exponent. For example, if one sets it very high, at the Planck energy—an energy at which the need to reconcile quantum field theory with general relativity becomes acute—then the exponent is 34. That is, the Higgs mass is a difference of two numbers that, written in decimal notation, match in their first thirty-four digits, and then differ in the thirty-fifth digit. Indeed, dismaying.

Again, strategy (Gen) stumbles here. For the most popular version of strategy (Gen) for this case is to appeal to an idea that extends the standard model of high-energy physics, which (as mentioned in section 2) was consolidated in the mid-1970s. And this version of strategy (Gen) turns out to have a fine-tuning problem of its own.

The idea being appealed to here is called “supersymmetry.” We do not need details about it, but can just note the following. Supersymmetry comes in various versions. The “good news” is that some versions imply that the observed mass of the Higgs boson is generic in strategy (Gen)’s sense: the observed mass lies in a range where it is expected to lie. But the trouble is that in order to have this implication, these versions also imply that there are other particles with a mass similar to the Higgs, particles that have not been observed. These particles, predicted by supersymmetry, are supersymmetric partners of known particles, and are called “superpartners.” But as I say, no such particles have been observed, not even with masses far from that of the Higgs.

We can put the problem a bit more precisely. The only way that supersymmetry can avoid the embarrassing implication that there are superpartners with a mass similar to the Higgs is to postulate higher masses for the superpartners, so high that our particle accelerators—more generally, our

experiments—cannot detect them. But to postulate this requires . . . fine tuning the masses of the superpartners.

(A note about terminology: in physics, especially high-energy physics, “naturalness” is used to mean, in effect, the opposite of “fine tuning.” So, in this terminology, the hope was that supersymmetry would show the Higgs mass to be natural. But the problem is that if all masses are natural, then the masses of the superpartners should be similar to that of the Higgs.)

7.2 The Difficulties in General Terms

So much by way of reporting examples where strategy (Gen) stumbles. I now move to a general statement of the difficulties strategy (Gen) faces. As I see matters, it is clearest to distinguish:

- (i.) A group of difficulties, each of which is not specific to cosmology, but arises from the variety, and often, the context-dependence or subjectivity, of the considerations that determine whether something counts as generic and/or as not needing explanation, so that it can provide, or at least contribute to, an explanans. These are difficulties that we have already seen at various places in our discussion.
- (ii.) Difficulties that are specific to cosmology, i.e., that arise from the fact that the system we are concerned with is the entire cosmos. These are difficulties we have touched on but not focused on.

I will treat (i) here, and (ii) in 7.3 below, which will lead into the next section.

As to (i), here are three main ways in which we have seen variety, or context-dependence, or subjectivity, about what counts as generic and/or as not needing explanation. I present them in the order in which they came up in the previous two sections.

First, in both sections, we discussed how judgments about what is generic or not needing explanation are often moulded by the context of enquiry and by pragmatic or even subjective factors. For example, recall Kepler’s effort to explain why there were six (known) planets in terms of Platonic solids (cf. the start of section 5 above), and the judgment that an initial state (“initial condition” in the terminology of physics) is generic at least in the

sense that it provides an *explanans* for a later state (cf. (i) and (ii) at the start of section 6.1 above).

Second, in what I called the topological and the effective field theory approaches to making “generic” precise, there is again variety and context-dependence of judgments. This may seem surprising, since these approaches’ definitions of notions like a topology or a renormalization group and its flow (on a space of physical systems, problems, or theories) are, after all, mathematical.

But, of course, being mathematical does not imply being unique. In general, a set can have many different topologies defined on it, and similarly, for defining renormalization flows on a set of Lagrangians (or Hamiltonians). Agreed, of the many mathematically consistent definitions, only some will be natural or significant, from the point of view of physics. But in general, the notion of “physically natural or significant” is too vague and/or ambiguous to pick out a unique definition. So, there will be a choice to be made, depending on context or aims.

Third, on the probabilistic approach to making “generic” precise, there are similar difficulties. For a set can have many different probability distributions (in more mathematical terminology: probability measures) defined on it. So, the obvious question arises: What justifies the one being used? And even after accepting, for whatever reason, one distribution as correct, or at least as correct for one’s purposes, there are the questions that I listed as (i) and (ii) at the end of the section 6. Namely, how far under the tail of the distribution—how much of an outlier—must an observation, for example, of the value of a variable, be to count as not generic, as atypical? (And so, to count as disconfirming the hypothesis that the distribution is correct.) And even after accepting an answer to this, there is the more general question of what framework of statistical inference (Bayesian? Neyman-Pearson?) we should adopt. That is, how should we infer from observations to hypotheses about what the correct probability distribution is?

7.3 Difficulties Specific to Cosmology

So much for my group (i) of difficulties. I turn to my second group (ii): Difficulties that arise from the fact that the system we are concerned with is the entire cosmos. So, we return to focus on the main question we first formulated at the end of section 4 above. Namely, suppose we are given a cosmological multiverse of many bubble universes (domains), across which the value of a fundamental physical parameter (normally called a “constant

of nature”!), such as the cosmological constant or the electric charge on an electron, varies. Then the question is: How can we implement our explanatory strategy (Gen) to explain the value as “to be expected”?

As I see matters, there are two main points to make about this question.

The first is the obvious point that in this cosmological setting, the difficulties in group (i) are exacerbated. For the cosmological multiverse, the theoretical context is so speculative that rigorous definitions are hardly to be had. So, the difficulties are not just about having to appeal to context, pragmatic factors, etc., so as to single out your preferred precise definition of being generic. Also, the daunting complexity of the relevant state space (recall the 10^{500} vacua at the end of section 4 above) makes it hard to give rigorous definitions of topologies, or renormalization flows, or probability measures.

We can express this point as a further comment on the problem that we first admitted back at the beginning of chapter 4’s discussion of the Everettian multiverse (section 6 of chapter 4). Namely, no Everettian knows how to write down the state of the cosmos—the symbol Ψ , with its honorific use of a capital letter, is a promissory note. We also admitted in this chapter’s section 1 that eternal inflation aggravated this problem in several ways, although in order to keep the discussion as simple as possible, I proposed that we take the envisaged quantum state of the cosmological multiverse to be a sum or superposition, over the countless bubbles (domains), of each of their Everettian states Ψ in the sense of chapter 4 (especially section 8’s sketch definition of “world”). That is, the multiverse’s state is a sum of states Ψ_i , where the label “i” on the summands labels the different universes. And recall from the end of this chapter’s section 4, that there might be dauntingly many states in the sum, dauntingly many values of the label “i” in the context of string theory: 10^{500} !

So much by way of gathering previous discussions’ threads about our not knowing how to write down the state of the cosmos. The present point is the further comment that not only are we unable to write down the state. Also, we cannot rigorously define such notions as the state being generic or the appropriate probability distribution on states.

The second main point is about probability, and more specifically about *confirmation*, a topic that will be developed in the following sections. Suppose that despite the difficulties above, we could define various probability distributions on the states of the multiverse and make sense of the

idea that one of them is correct. Still, we would face the question: “How can we gather evidence about which one is correct?”

The problem is obvious. We presumably cannot get empirical data about bubble universes other than the one we are in. For the spatiotemporal connection between “our bubble universe” and any others is through the inflationary epoch, which for us is long gone. (And even apart from inaccessibility, its extreme conditions, for example, of temperature, put it so far beyond established physics that we could hardly expect to get interpretable data from it.) But without such data, it is very unclear how we could gather evidence about which probability distribution is correct. After all, our understanding of the phrase “correct probability” derives from cases where there is a set of actual systems or events (tosses of a coin or coins, rolls of a die or dice, etc.) that are, or are believed to be, suitably similar. We then estimate probabilities by counting the proportions, the relative frequencies, with which certain features (e.g., heads or tails or scores on a die) occur. Agreed, there is debate about how best to make these estimates from observed frequencies, a debate addressed by theories of statistical inference—recall questions (i) to (iv) at the end of section 6. But all parties agree that there are deep connections between probabilities and frequencies. So, if we do not know any frequencies, how can we guess probabilities?

8. Biased Sampling: Eddington’s Net

In the last two sections, I discussed the successes, and then the difficulties, of what (at the end of section 5) I labelled “strategy (Gen)” for explaining a fact: namely, by deducing it, or showing it is generic, given some appropriate framework or *explanans*. In that discussion, the sort of fact to be explained has been, since the end of section 4, facts about the value of a cosmological parameter, or a constant of nature—though we now envisage that such “constants of nature” may vary from one bubble universe to another.

So, now I turn to what I labelled “strategy (Obs)”: explaining a fact that one admits may well *not* be deducible or generic, by showing that it is likely (or at least has high enough probability) *to be observed*. In this section, I discuss this strategy and how it differs from strategy (Gen) in everyday terms, regardless of cosmology, to which the next section will return.

The distinction between these two strategies lies in the fact that what is most probable to occur is not necessarily what is most probable to be observed. That is, we need to distinguish (a) having a high, or high enough,

probability (or frequency) in a *total population* of cases, from (b) having a high, or high enough, probability (or frequency) in the *subpopulation* that we observe. The distinction between (a) and (b) arises from something familiar from very elementary applications of probability theory: biased sampling.

For example, when you take a sample from a set, say, ten adults from a population of 10,000, to estimate the average height, your sample might be *biased* in the sense that the frequency of the attributes of interest (here, being such-and-such metres tall) within the sample is different from its frequency in the total population of 10,000. Agreed, some difference in frequencies is to be expected. Almost always, the sample frequency does not exactly equal the population frequency (or population probability)—this is called “stochastic variation.” And what counts as a large enough difference to earn the label “bias” is a matter of how big a difference counts as significant for one’s purposes—and so is partly a matter of judgment. For example, the sample might be biased, with large heights more frequent (as a proportion) than in the total population, simply because you chose the ten people from your local basketball club. And whether your consequent overestimate of the population’s average height is large enough to matter will depend on your purposes. For example, a five-centimetre overestimate would matter if you planned to sell shirts to the population, but not if you planned to sell them lottery tickets.

So far, so obvious. But our interest lies in cases where the sample is biased, not as a matter of coincidence (as might well occur in the example of the basketball club), but as a systematic effect of the method of observation or data-gathering. This is called a “selection effect” (or an “effect of observational selection”).

A famous example occurred in the 1936 U.S. Presidential election. The incumbent Democratic President, Roosevelt, beat his Republican challenger, Landon, by a large margin. But one magazine had predicted that Landon would win on the grounds that it posted questionnaires to ten million subscribers, of whom about two million responded, mostly favouring Landon. But this was a selection effect. The subscribers were disproportionately Republican, compared with the nation at large, and the subscribers with the interest to send back a response were even more disproportionately Republican.

There is also a famous and vivid metaphor for selection effects, invented by the British physicist Arthur Eddington (1882–1944). In his *The Philosophy of Physical Science* (1938), he wrote:

Let us suppose that an ichthyologist is exploring the life of the ocean. He casts a net into the water and brings up a fishy assortment. Surveying his catch, he proceeds in the usual manner of a scientist to systematise what it reveals. He arrives at two generalisations: (1) No sea creature is less than two inches long. (2) All sea creatures have gills. These are both true of his catch, and he assumes tentatively that they will remain true however often he repeats it.

To sum up, fishermen whose net has a mesh of, say, two inches, and who therefore observe that all the fish *in their catch* are longer than two inches, should not infer that all the fish *in the lake* are also longer than two inches. (Incidentally, Eddington intended his metaphor to teach a different and more contentious moral than just “Beware selection effects.” Namely, his moral was about the relation between physics and philosophy, which we will return to in the “Notes and Further Reading” section at the end of this chapter and chapter 6.)

So far, we have thought of selection effects as a bug, as a hindrance to making good estimates of the probability or frequency of an attribute in the total population from which we take our sample. That is right: they are a hindrance, especially if we do not know the details of how our sample is biased. If we do know those details, we can try to “build in” the details to the procedure by which we make an estimate, so as to compensate for the bias.

How to do this is a topic in the statistical theory of estimation. The rough idea is, of course, to conditionalize one’s probabilities on a description of the sampling process. Similarly, if we know only some details, or have only probabilistic information about how the sampling is or might be biased, we “build in” what we know about the process. In short, this is familiar ground in the practice of statistical inference, specifically, in the theory of estimation. There may be practical difficulties about learning the details of the sampling process, and debate within statistical theory about how best to “build in” those details to the procedure for making an estimate. But there is no general or philosophical problem about the fact that we need to allow for these details.

But there is also another way to think of a selection effect. Namely, as *explaining* the frequency of an attribute that we observe (the height of a human, or their political views, or the length of a fish), despite that frequency being different from those for the total population. It is, of course, this perspective that is encapsulated by the strategy (Obs).

Again, I think there is no general or philosophical problem about this strategy, although there may well be difficulties about the details of the sampling process. As above, these could include, first, practical difficulties about learning the details. For example, does the fishermen's two-inch mesh really prevent any fish longer than two inches, if such there be, from getting caught? And there could be theoretical difficulties about how the calculation of an estimate should allow for such details. For example, how should my estimate of average height allow for my having sampled heights from a basketball club? Being told to conditionalize on the proposition, *All the people in my sample play basketball* gives very little guidance, if I do not know any details about how much playing basketball favours the tall.

I do not mean to downplay these difficulties, whether in physics or in other sciences. In all sciences, the observational process is indeed liable to be biased, i.e., the value of the variable we wish to observe may be correlated with the process of observation, and it can be a hard and complicated matter to recognize this and to understand it in enough detail to compensate for the bias. Just think of the care that goes into calibrating scientific instruments. But the point is that this is familiar ground in the practice of science and statistics: there is no general or philosophical problem hereabouts.

Or rather, there is no such problem, outside cosmology. But when we consider cosmology, there may be such problems, as I discuss in the next section.

9. Selection Effects in Cosmology: The Anthropic Principle and the Cosmological Constant

So, we return to the main question that we first formulated at the end of section 4. Namely, given a cosmological multiverse of many bubble universes (domains), across which the value of a fundamental physical parameter, such as the cosmological constant or the electric charge on an electron, varies: How can we explain the value that we measure? And relatedly, since one confirms a scientific theory by its predicting—and one hopes, explaining—results of measurements and observations: How, if at all, can we confirm

(or disconfirm) a theory postulating such a multiverse? I begin, in 9.1, with general remarks, and then, in 9.2, I turn to specifics.

9.1. *The Daunting Questions*

In general terms, our trouble about addressing these two questions is that it is not enough to say that (i) a cosmological theory will assign differing probabilities to various values of such a parameter, of which each bubble universe exhibits one value; and that (ii) this enables us to assess the theory by ordinary statistical inference, along the lines that if the observed value is too much of an outlier, i.e., in the tail of the probability distribution, then we should conclude that the theory is disconfirmed. Indeed, it is not enough for two reasons. We spelled out the first reason at the end of section 7. Namely, we measure and observe only our own bubble universe, our own “cosmic parish.” So, for a parameter that describes an entire such universe, like a “constant of nature,” we only get one number—we cannot count frequencies. That is a miserably meagre basis on which to judge which probability distribution is correct. Indeed, it is too meagre even for estimating the average value of the attribute in question; imagine trying to estimate the average height of a population of 10,000 by measuring the height of just one person. And for the cosmological multiverse, we expect the number of bubble universes to be vastly larger than ten thousand.

The second reason is, of course, selection effects, which prompt our explanatory strategy (Obs), viz. that we explain a value of a parameter by its being probable (or at least is probable enough) *to be observed*. For in the context of measuring fundamental parameters in a multiverse, biased sampling threatens to be a significant problem. The problem is not just that, as in the first reason, each bubble universe exhibiting only one value means the sample size is so small as to be useless, thanks to stochastic variation. Also, established theories in both physics and chemistry show that many of the parameters at issue, such as the cosmological constant and the charge on the electron, *are* indeed correlated with what the last section called the “process of observation.” That is, they are correlated with facts that underpin humans’ being able to measure the parameter, such as the fact that life on earth depends on a suitable abundance of carbon and oxygen, or that stars exist and have planets orbiting them, for times long enough for the complex carbon chemistry of life to evolve on a planet.

With this mention of how the process of observation involves humans (or their complex carbon chemistry), we thus arrive at last at the well-known

phrase “the anthropic principle.” For some fifty years, it has been the topic of heated debate in both cosmology and philosophy. (The phrase was suggested in 1973 by Brandon Carter (1942–), a theoretical astrophysicist. People also talk of “anthropic reasoning.”)

9.2 Fine Tuning: The Cosmological Constant

The general idea of these correlations is that modern astronomy and cosmology (as reported in section 2) have shown our universe to be very unified. This is not just in what I called the “geographical” sense that a parameter, such as the charge on the electron, takes the same value throughout our universe (section 4), but also in the sense that what happens to its smaller parts, such as a star or galaxy, depends on truly global, i.e., universe-wide, features.

A good example is the parameter I mentioned first in this chapter: the density parameter Ω , which is the ratio of the universe’s density to the critical density that would make the final rate of expansion zero. (Cf. sections 2 and 3. But here we are concerned not with the speculative inflationary phase perhaps explaining Ω ’s fine tuning, but with Ω ’s value much later on, and so within established cosmology, say, after the recombination time 380,000 years after the Big Bang.) If the early universe were very dense, i.e., Ω was much greater than 1, the universe would have recollapsed in far less than thirteen billion years, so that life would not have had time to evolve on planets, while if it was much less than 1, no stars, and therefore no planets, would have formed.

Another example, which I will return to later in this section, is the *cosmological constant* Λ (introduced at the start of section 7). Since it represents a universal expansion, having a much larger value than it actually does is like Ω being much less than 1. Namely, a much larger value would have made the universe’s expansion too fast for stars to form.

A third example is the other parameter I mentioned, the charge on the electron, or more precisely, the ratio of the strengths of electromagnetic to gravitational forces (which is greater, the greater the charge on the electron). If this had been much smaller than it is, gravitation would have been comparatively stronger, and the universe would have recollapsed in far less than thirteen billion years, so that life would not have had time to evolve on planets.

These correlations are often “tight” in the sense that they are not probabilistic. They are not a matter of the probability of one proposition, the

parameter's value, being altered by conditionalizing on another proposition, such as the proposition that there is abundant carbon and oxygen. They are a matter of one value, or a range of values, being mathematically dependent on, i.e., a function of, another value or range of values. So these mathematical dependences can be, and often are, summed up in what section 7 of chapter 3 called a "counterfactual conditional" along the lines of: "If the parameter had taken a different (or different enough) value than its actual one, then there would be no observations—at least, no observations by humans (understood as having such-and-such carbon chemistry)." Witness the examples above.

Besides, the correlations are in several cases "tight" in a distinct numerical sense. Namely, only a very narrow range of values of the parameter, such as a few percentage points around its actual value, is compatible with a fact like there being abundant carbon and oxygen. Hence, this is also called "fine tuning."

Agreed, this is not stupendous fine tuning to many decimal places, such as we saw in the flatness and horizon problems (section 3) and in the problem of the Higgs mass (section 7). (The fine tuning of Ω in the flatness problem was 10^{-16} , i.e., a hundred-million-millionth of one per cent.) But each of these stupendous fine tunings occurred within a single theoretical framework, while here, many such frameworks are in play. The physical and chemical facts and processes that connect global features—like cosmic expansion or the relative strengths of the fundamental forces—to local features—such as the existence of heavier elements like carbon or the presence of life on rocky planets orbiting stars—are highly diverse. They range from nucleosynthesis in the early universe and in stars, through planet formation and the chemistry of water, to the evolution of life. So, it is indeed very striking that our established theories of these diverse facts and processes, conjoined together, provide a "patchwork description" of the facts and processes that—despite its diverse ingredients—implies these numerical quantitative links, constrained to within a few percentage points.

But this is not to suggest that it is straightforward to spell out these implications, that it is straightforward to quantify the correlation. As I said at the end of the last section, in connection with calibrating scientific instruments, even within a single scientific theory, compensating for the fact that the observational process is biased can be a hard and complicated matter. This is all the more so when there are several theories or frameworks in play,

and when the parameter in question is correlated via various mechanisms with various aspects of our making observations.

For example, there are many different necessary conditions, each scientifically describable, of our observing the charge on the electron. The observer is alive, and life requires, one may well argue, complex carbon chemistry. Carbon requires stellar nucleosynthesis. And the complex chemistry of life requires, one may argue, that a planet orbit a star at a suitable distance (neither too hot nor too cold, like Goldilocks' porridge), and for a long enough time, so that life can evolve. All these correlations and the mechanisms underpinning them, and these mechanisms' mutual relations, are very hard to disentangle. And this is so, even if we somehow settle on some exact definition of "observation" or "life." Besides, this is so even for a single parameter, such as the charge on the electron, let alone all the physical parameters of interest.

But despite the complexities just mentioned, some examples *are* comparatively straightforward to calculate. So, I end this section with a bit more detail about one such example, namely, Weinberg's renowned explanation of the value of the *cosmological constant* as an observation selection effect. Recall from the start of section 7 that the cosmological constant Λ represents a repulsive force between two masses. So, in a cosmological context, Λ amounts to a tendency for the universe to expand; so, it opposes the gravitational force that tends to make matter clump together.

Weinberg recognized that the requirement that life evolves in an expanding universe of the type considered in the standard model of cosmology is correlated with the value of the cosmological constant, through a single, and comparatively simple, mechanism. Thus, he wrote:

In a continually expanding universe, the cosmological constant (unlike charges, masses etc.) can affect the evolution of life in only one way. Without undue anthropocentrism, it seems safe to assume that in order for any sort of life to arise in an initially homogeneous and isotropic universe, it is necessary for sufficiently large gravitationally bound systems to form first . . . However, once a sufficiently large gravitationally bound system has formed, a cosmological constant would have no further effect on its dynamics, or on the eventual evolution of life.

So, the idea is that the evolution of life constrains the cosmological constant in a simple way, because we can think of (a positive value of) the constant as a long-range repulsive (“anti-gravity”) force. Thus, one assumes that (i) life can only exist on planets, and (ii) life takes a long time, say billions of years, to evolve. Since (i) requires that matter has the chance to clump together under gravity so as to form planets, the initial expansion cannot be too powerful. That is, (i) implies that there is an upper bound on the cosmological constant (i.e., a number that it must be less than). On the other hand, (ii) means that the universe must last long enough for life to evolve. So, gravity cannot be so powerful (the initial expansion cannot be so weak, Λ cannot be so small or negative) that gravity overcomes the initial expansion in a Big Crunch that happens so early that life does not have time enough to evolve. Thus, (ii) implies that there is also a lower bound on the cosmological constant.

Indeed, the calculation along these lines in 1997, by Weinberg and co-authors, amounted to showing that the observed value of Λ (tentative in 1997, but confirmed a year later) fell “safely” between this lower and upper bound. So, it is natural to see this calculation as explaining the observed value of Λ as an observation selection effect.

10. Confirming a Theory of the Multiverse

The last section’s report of the fine-tuned correlations between cosmological parameters and facts about observers focused on the one universe we are in. But now let us consider these correlations in the context of the cosmological multiverse.

I shall first state—without repeating all the difficulties presented in sections 7 to 9—the basic predicament that besets observers in a single bubble universe, trying to confirm a cosmological theory. Then I shall sketch a scheme for overcoming this predicament, a scheme that I find clarifying. In 10.1, I report the scheme’s general ideas. In 10.2, I report how it incorporates ideas from both the strategies (Gen) and (Obs). But I admit, of course, that the scheme is fiendishly difficult to apply, i.e., calculate with, except in simple toy models of such theories.

The basic predicament is that when we observe our universe, we are like Eddington’s fishermen. Our observations of a physical parameter (e.g., the cosmological constant) are like measurements of the length of fish in the catch. And so, we should not infer that in unobserved bubble universes—in

domains other than ours—the parameter takes, or is likely to take, a value close to what we observe. For our established theories describe how the parameter’s value is correlated with whether the domain has observers in it. So, if in some other bubble, those theories are true, or approximately true, and there is no observer there, the value would be different. And if in this other bubble, those theories are badly wrong, i.e., not even approximately true, then anyway, all bets are off about the parameter’s value.

10.1 Hartle and Srednicki’s Scheme

The scheme I favour was first proposed by Hartle and Srednicki (about twenty years ago). It has, of course, been developed since then, with proposals by such authors as Aguirre, Azhar, Hertog, Tegmark, and by Hartle and Srednicki themselves. But for simplicity and brevity, I will sketch a very simplified version (for references, see the “Notes and Further Reading” section at the end of this chapter).

The scheme aims to incorporate appropriately the ideas from both strategies, (Gen) and (Obs): the idea of being generic and the idea of being probable (or probable enough) to be observed. And it does this in a Bayesian way.

In a bit more detail, this means that the scheme prescribes probabilities for data D (say, the value of a cosmological parameter such as Ω or Λ) to be observed, conditional on the cosmological theory T , and other propositions, for example, propositions encoding selection effects. (Probabilities like these, i.e., probabilities of data or evidence conditional on a theory or hypothesis, are often called “likelihoods.”) One then uses Bayes’ theorem to calculate the probability of the conjunction of T with the other propositions, conditional on the data D .

Here, we need not go into any detail about Bayes’ theorem. For us, it is sufficient that the theorem provides a way to calculate from the conditional probability $P(A/B)$ of a proposition A conditional on a probability of B , the “opposite” conditional probability, $P(B/A)$. (The calculation uses the values of other probabilities, additional to $P(A/B)$.) The basic idea for how to use the theorem to make inferences (called “Bayesianism” or “Bayesian statistical inference”) is then: We take B to be some hypothesis or theory we wish to assess (confirm or disconfirm), while A is some proposition reporting evidence such as an experimental result.

Thus, we imagine that the hypothesis B prescribes a value of $P(A/B)$: the probability, assuming B is true, of the evidence A . $P(A/B)$ might be high, or it might be low, with A “lying under the tails” of the probability distribution

$P(\cdot/B)$ prescribed by B . Then the “magic” of Bayes’ theorem is that we can then calculate $P(B/A)$. And then the Bayesian statistician tells us that if we, in fact, learn the evidence or result A —and so set our credence (subjective probability) in A equal to 1—we should adjust our credence in the hypothesis B , from our initial $P(B)$, to be equal to $P(B/A)$. (So, $P(B/A)$ is often called the “posterior probability” of B .)

Also, the Bayesian says that A confirms B provided that this prescription makes our credence go up, i.e., provided that $P(B/A)$ is larger than $P(B)$. This Bayesian account of how evidence confirms or disconfirms hypotheses has many merits. In many cases, it gives the intuitively right verdicts about confirmation. In particular, observing some evidence A that “lies under the tails” of the probability distribution $P(\cdot/B)$ prescribed by B will, generally speaking, disconfirm B .

Returning to cosmology, the idea of Srednicki and Hartle’s scheme will thus be that after receiving the data D (about say, the value of Ω or Λ) we should apply Bayes’ theorem to the probability of D conditional on the conjunction of the theory T and propositions about selection effects, etc., to set our credence in the conjunction equal to the posterior probability.

10.2 The Details

To convey how ideas from both my strategies, (Gen) and (Obs), get incorporated in the scheme, I will begin by summarizing the problems one faces in extracting predictions from cosmological theories of the kinds currently envisaged (including inflation). This will lead to details about Srednicki and Hartle’s proposal, which they call a “framework.”

I summarize the problems under three headings. These headings will echo the difficulties I presented in sections 7 to 9, about such matters as (i) the definition of a probability function on a very large space of possibilities, and (ii) how to specify the “fact about life or observation” on which we need to conditionalize so as to accommodate selection effects, etc. (My three headings are also a simplification, or amalgamation, of an analysis by Aguirre. He lists seven problems, or headings, rather than my three; again, the “Notes and Further Reading” section of this chapter give references.)

I call the headings “Measure,” “Conditionalization,” and “Typicality.” They are as follows.

1. *Measure*. As discussed in sections 6 and 7, what are the elements of the set (called in probability theory “the sample

space”) on which the probability distribution (called “measure”) is to be defined? Should they be domains, i.e., bubble universes, even though these vary greatly in volume? Or should the elements be space-time regions of equal volume? Or some other option? And once a sample space is defined, which measure on it should we adopt?

2. *Conditionalization*. As discussed in sections 8 and 9, we need to allow for selection effects. But how exactly should we characterize our observational situation? How detailed should the proposition describing it, on which we will conditionalize, be?
3. *Typicality*. As discussed in sections 6 and 7, there are various problems with how to make precise the idea of a fact (in particular, our *explanandum*) being generic or typical. In particular, how much “under the tails” of a probability measure can our observation turn out to be, without our then inferring that the theory is disconfirmed?

This trio of headings leads directly to Srednicki and Hartle’s proposed Bayesian scheme for discussing the confirmation of cosmological theories. Srednicki and Hartle define a “framework” as a conjunction of:

- (i.) A cosmological theory T (though often cosmologists will say “model”). This is taken to solve the Measure problem described by 1) above. So, we write a probability $P(\text{ }/T)$, where the argument-place, i.e., the gap, will be filled by a proposition about the value of a physical parameter, e.g., the cosmological constant.
- (ii.) A “selection proposition” that describes our observational situation, which is called a “conditionalization scheme,” and labelled as C . So, we conditionalize on C as well as on T , and we consider $P(\text{ }/T, C)$, where we expect the argument-place to be filled by a proposition D about “our seeing” some specific data.

- (iii.) A probability distribution, denoted by the Greek letter ξ (pronounced “xi”), and called the “xerographic distribution” by Srednicki and Hartle. This is defined on those domains, i.e., bubble universes, that have a non-zero measure according to $P(/T, C)$. ξ encodes typicality assumptions: as Srednicki and Hartle discuss, it need not be a flat distribution.

So, the idea is that (i) to (iii) jointly implement our envisaged solutions to the problems listed under (i) to (iii) above. The upshot is that we are to consider $P(/T, C, \xi)$. Srednicki and Hartle call $P(D / T, C, \xi)$ the “first-person” likelihood of seeing data D . It is a probability “to be observed,” rather than a probability “to be,” thanks to its conditioning on C and on ξ , i.e., its encoding our observational situation and also the typicality assumption we are making.

Srednicki and Hartle then propose a Bayesian framework to compute degrees of confirmation of the framework, i.e., the conjunction of T , C , and ξ . That is, they use Bayes’ theorem to calculate $P(T, C, \xi / D)$. They (and other authors, such as Azhar) go on to give examples of the framework in action. They show, for various “toy” cosmological models/theories T (e.g., with finitely many bubble universes, so as to assume the Measure problem has been solved), how various conditionalization schemes C , and typicality assumptions ξ , fare in the light of various data D .

To sum up, I suggest that the Srednicki-Hartle scheme of frameworks is a clear and convincing for handling both selection effects and assumptions about being generic (typicality). In other words, it handles both the idea of being probable enough to be observed and the idea of being generic. It gives both ideas an appropriate role in the endeavour of confirming a theory postulating a cosmological multiverse. So, we should look forward to it being applied to more realistic cosmological models.

11. Notes and Further Reading

As with chapters 2, 3, and 4, the literature relevant to this chapter is dauntingly large. As in those chapters, I therefore recommend two general strategies:

- (i) Consult the internet encyclopedias and archives, together with the accessible books, listed under the “Notes and Further Reading” section for chapter 1. All of these sources cover the cosmological multiverse.
- (ii) Read the seminal works. Many of the classic contributions are remarkably readable.

In section 1 below, I give more detail about selected items falling under these two headings. I begin with accessible, even popular, works, and then turn to academic (though still readable) books. Section 2 lists research articles, ordered to follow the sequence of topics in this chapter.

11.1 Books

11.1.1 INTERNET ENCYCLOPEDIAS

I recommend three entries in the *Stanford Encyclopedia of Philosophy*, on fine-tuning, the philosophy of cosmology, and cosmology and theology:

- Friederich, Simon. “Fine-Tuning.” Available at: <https://plato.stanford.edu/entries/fine-tuning/>
- Smeenk, Chris, and George F. R. Ellis. “Philosophy of Cosmology.” Available at: <https://plato.stanford.edu/entries/cosmology/#Mult>
- Halvorson, Hans and Helga Kragh. “Cosmology and Theology.” Available at: <https://plato.stanford.edu/entries/cosmology-theology/#5.2>

11.1.2 SEMINAL WORKS

I recommend three especially accessible books, the first two by physicists and the third by a philosopher.

- Barrow, John, and Frank Tipler. *The Anthropic Cosmological Principle*. Oxford University Press, 1986. This remains the

locus classicus for many topics of this chapter. Although published soon after inflation was proposed, it does cover inflation. I especially recommend chapters 4–8, with chapters 6 and 7 focusing on inflationary and quantum-theoretic aspects.

- Carr, Bernard, ed. *Universe or Multiverse?* Cambridge University Press, 2007. This volume collects articles from four conferences (2001–2005) devoted to the chapter’s themes. Almost all the articles do not require knowledge of advanced physics (not even within cosmology). Most of the authors are prominent researchers in cosmology. Some of them advocate the cosmological multiverse and endorse anthropic explanations of parameters, such as the cosmological constant (e.g. Linde, Rees, Susskind, Tegmark, and Weinberg), while some criticize both these tenets (e.g., Ellis and Smolin). So, this collection is an invaluable resource, and in my list of research articles below, I will cite (for more specific reasons) the articles by Aguirre, Vilenkin, and Weinberg.
- Friederich, Simon. *Multiverse Theories: A Philosophical Perspective*. Cambridge University Press, 2021. This is an excellent recent philosophical assessment. Chapters 7 to 9 focus on confirmation of multiverse theories, while chapter 10 briefly (and sceptically) discusses the logical and quantum multiverses of chapters 3 and 4. It is available at: <https://www.cambridge.org/core/books/multiverse-theories/68CE18BE78DE31550C67855107A57942>

11.1.3 POPULAR MONOGRAPHS

There are many excellent popular books addressing all or some of this chapter’s topics. Here are four popular books about all of the topics, ordered by how sharply they focus on this Chapter’s concerns, to the exclusion of other themes.

- Tegmark, Max. *The Mathematical Universe*. Vintage Books, 2014. This offers a detailed popular advocacy of the cosmological multiverse. As noted in chapter 1 and in the “Notes and Future Reading” section of chapter 3, I

stand by my criticisms of this book, especially Tegmark's Pythagoreanism. But these criticisms do not undermine his multiverse advocacy.

- Rees, Martin. *Before the Beginning*. Simon & Schuster, 1997. And Rees, Martin. *Just Six Numbers*. Weidenfeld & Nicolson, 1999. Both books also give a lot of detail about astrophysics and cosmology, apart from the multiverse: for example, about the early universe, the cosmic background radiation, stars, galaxies, and black holes—thus filling out my section 2's review of the current golden age of cosmology.
- Barrow, John. *The Book of Universes*. Bodley Head, 2011. As the title hints, this gives a lot of detail about the “other” universes (space-times) that are admitted as possible solutions by general relativity (i.e., by Einstein's theory of gravity), but are not part of the standard Big Bang cosmological model. But its chapters 9 and 10 discuss inflation and the multiverse.

Here are four other popular science books, each giving a detailed account of one topic of this chapter. I list them in the chronological order in which their topic took centre-stage in cosmology, which mostly corresponds to the topics' order in this chapter's sections.

- Singh, Simon. *The Big Bang: The Origin of the Universe*. Fourth Estate, 2004. This focuses on how the standard Big Bang cosmological model was confirmed by about 1980, largely as a result of the discovery of the CMB (cf. section 2).
- Guth, Alan. *The Inflationary Universe: The Quest for a New Theory of Cosmic Origins*. Penguin, 1998. This masterpiece of popular science is by one of the inventors of inflationary cosmology (and so also an advocate of the multiverse) (cf. section 3).
- Greene, Brian. *The Elegant Universe: Superstrings, Hidden Dimensions and the Quest for the Ultimate Theory*. Random House, 1999. This is an excellent exposition of string theory, though without emphasis on the “landscape” of many vacuum states (cf. section 4).

- Hertog, Thomas. *On the Origin of Time*. Penguin, 2023. This focuses on quantum cosmology, including times even earlier than the putative inflationary epoch, and therefore, on the no-boundary proposal for the initial state of the cosmos, mentioned in section 1.

A Cautionary Note on “Nothing”

This last topic, the initial state of the cosmos, prompts me to issue a warning.

Recall (from this chapter’s preamble and section 1) that in physics, “vacuum” does not mean “nothing” or “no physical system.” It means “state of lowest energy” (hence the synonym “ground state”). There is a widespread tendency in popular physics books, even by physicists, to forget this, and thereby to suggest that by postulating a vacuum state as the initial state of the cosmos, we can explain the creation of the cosmos—out of nothing! In this widespread but pernicious mistake, the phrase “vacuum fluctuation” gets abused similarly to “vacuum” state.

So far as I know, the most egregious example of this mistake is L. Krauss’ *The Universe from Nothing* (Free Press, 2011)—which I therefore do *not* recommend.

But I do recommend two antidotes to this sort of error.

- The devastating review of Lawrence Krauss by David Albert, in *The New York Times* (March 25, 2012), which is available at: <https://www.nytimes.com/2012/03/25/books/review/a-universe-from-nothing-by-lawrence-m-krauss.html> (accessed January 13, 2026).
- Weatherall, James. *Void: The Strange Physics of Nothing*. Yale University Press, 2016. This is a magisterial, but very readable, account of physics’ changing conception of vacuum (void, empty space) from the time of Descartes and Newton, through the rise of classical field theories such as electromagnetism and of quantum field theories, until today, including the landscape of countless string vacua, and thus, the cosmological multiverse.

11.1.4 ACADEMIC TEXTS

I turn to recommending a few academic books. They are (1) cosmology textbooks, (2) philosophy of cosmology books, and (3) histories of twentieth-century cosmology.

(1) Cosmology

Here are three cosmology textbooks, in roughly ascending order of difficulty.

- Longair, Malcolm. *Our Evolving Universe*. Cambridge University Press, 1996. Its final chapter, “The Origin of the Universe,” is a very readable “immediate successor” to this chapter.
- Rowan-Robinson, Michael. *Cosmology*. Oxford University Press, 2004 (fourth edition). Its Epilogue, “Twenty Controversies in Cosmology,” covers inflation. It is available at: <https://academic.oup.com/book/52969?login=false>
- Liddle, Andrew. *An Introduction to Modern Cosmology*. Wiley, 2015 (third edition). This is slightly more theoretical than Rowan-Robinson’s book, and it also covers inflation.

And here is an accessible review article, which was written for the celebratory “Einstein’s Legacy” issue of *Science* magazine in 2005:

- Guth, Alan (an inventor of inflationary cosmology) and David Kaiser’s “Inflationary Cosmology: Exploring the UV from the Smallest to the Largest Scales.” Available at: <https://arxiv.org/abs/astro-ph/0502328>; <https://www.science.org/doi/10.1126/science.1107483>

(2) Philosophy of Cosmology

Here are four books on the philosophy of cosmology. The first is a collection of invited articles. The second is a magisterial monograph covering many topics in the philosophy, or foundations, of general relativity as well as inflationary cosmology (its chapter 5). The third is Friederich’s recent monograph devoted to the multiverse, which I already recommended as “seminal” above, especially for its discussion of how we could confirm a

multiverse theory (i.e., the topic of my section 5 onward, especially section 10). The fourth is a collection of invited articles discussing in detail fine tuning of the conditions for complex chemistry and life, as well as for planet formation, etc. (i.e., the topic of my section 9).

- Chamcham, Khalil, Joseph Silk, John Barrow, and Simon Saunders, eds. *The Philosophy of Cosmology*. Cambridge University Press, 2017. This has many good articles. I will cite those by Smeenk, Hartle, and Hertog in my list of research articles below. The book is available at: <https://www.cambridge.org/core/books/philosophy-of-cosmology/2E9F97DDF98A672256D35B46C3F574B4>
- Earman, John. *Bangs, Crunches, Whimpers and Shrieks: Singularities and Acausalities in Relativistic Space-Times*. Oxford University Press, 1995. The book is at Oxford Scholarship Online: <https://academic.oup.com/book/49463?login=false>. It is also available, with almost all Earman's other work in philosophy of physics, at: <https://sites.pitt.edu/~jeerman/> (last modified January 16, 2013).
- Friederich, Simon. *Multiverse Theories: A Philosophical Perspective*. Cambridge University Press, 2021. The book is available at Cambridge University Press Core: <https://www.cambridge.org/core/books/multiverse-theories/68CE18BE78DE31550C67855107A57942>
- Sloan, David, Rafael Batista, Michael Hicks, and Roger Davies, eds. *Fine-Tuning in the Physical Universe*. Cambridge University Press, 2020. Available at: <https://www.cambridge.org/core/books/finetuning-in-the-physical-universe/DAAE3182CBC72F012EFF589E67178F1C>

(3) History of Twentieth-Century Cosmology

Here are three histories of twentieth-century cosmology. The first book narrates the establishment of the standard Big Bang model of cosmology. (The stress is on its defeat of its main rival, the steady-state theory, which my section 2 did not mention.) The second book is more technical; it also

covers the entire twentieth century and astrophysics as much as cosmology. Its chapter 16 focuses on this chapter's topics. The third book is a recent collection of articles.

- Kragh, Helga. *Cosmology and Controversy: The Historical Development of Two Theories of the Universe*. Princeton University Press, 1996.
- Longair, Malcolm. *The Cosmic Century: A History of Astrophysics and Cosmology*. Cambridge University Press, 2006.
- Kragh, Helga, and Malcolm Longair, eds. *The Oxford Handbook of the History of Modern Cosmology*. Oxford University Press, 2019; available at: <https://academic.oup.com/edited-volume/34295>

11.2 Research Articles

I now list some research articles. Their order mostly corresponds to this chapter's sequence of sections. The first three groups, 2.1 to 2.3, relate to this chapter's preamble and section 1. Groups 2.4 and 2.5 are about the difficulties of confirmation—in philosophical terminology: the underdetermination of theory by data—for cosmology in general, and for inflationary cosmology in particular. So, these correspond to sections 2, 3, and 5–7. Groups 2.6 and 2.9 are about observation selection effects, in general and in cosmology, and correspond to sections 7 to 9 of this chapter. Finally, 2.10 gives more details about the Hartle-Srednicki proposal of frameworks for confirming multiverse theories.

11.2.1 THE RELATION OF EVERETTIAN AND COSMOLOGICAL UNIVERSES

The relation of the Everettian and cosmological multiverses was a topic in this chapter's preamble and section 1. Again, there is a large literature. As a place to begin, I recommend the articles by Carter (the inventor of the phrase “anthropic principle”) and Mukhanov in Carr's edited collection *Universe or Multiverse?* listed at the start of (1) above. (Both authors advocate an Everettian approach.)

Examples of research articles about this relation—in fact, proposing that the two multiverses are the same—are:

- Nomura, Yasunori. “Physical Theories, Eternal Inflation, and the Quantum Universe.” *Journal of High Energy Physics (JHEP)*, volume 11, 063; available at: <https://arxiv.org/abs/1104.2324>
- Bousso, Raphael, and Leonard Susskind. “The Multiverse Interpretation of Quantum Mechanics.” *Physical Review D* 85 (2012); available at: <https://arxiv.org/abs/1105.3796>

11.2.2 DYNAMICAL REDUCTION AND COSMOLOGY

In section 1 (and more briefly in the preamble and section 3), I mentioned how cosmology confronts the quantum measurement problem on a cosmic scale. In particular, there was the idea that a peak in (the amplitude of) the wrinkles in the CMB is a seed for later gravitational clumping, and thereby for the later existence of galaxies and stars. I warned that “seed” is a metaphor, since the transition is from a quantum amplitude to a classical event of aggregation—a “collapse of the wave function.”

Recall also my warning, in 11.1.3 above, of a widespread mistaken tendency to slur over this transition, especially with the buzzwords of “vacuum state” and “vacuum fluctuation.”

In the light of this, I recommend some research tying the dynamical reduction programme—i.e., the effort to suitably modify the Schrödinger equation: cf. the citation of Pearle in subsection 1 of the “Notes and Further Reading” section of chapter 4—to cosmology. In this endeavour, work by Sudarsky and co-authors has been prominent. So here are three such articles.

- Perez, Alejandro, Hanno Sahlmann, and Daniel Sudarsky. “On the Quantum of Cosmic Structure,” *Classical and Quantum Gravity* (2006); available at <https://arxiv.org/abs/gr-qc/0508100>
- Berjon, Javier, Elias Okon, and Daniel Sudarsky. “Critical Review of Prevailing Explanations for the Emergence of Classicality in Cosmology,” *Physical Review D*, (2021); available at: <https://arxiv.org/pdf/2009.09999>
- Lechuga, Rosa-Laura, and Daniel Sudarsky. “Eternal Inflation and Collapse Theories.” *Journal of Cosmology and Astroparticle Physics* (2024); available at: <https://arxiv.org/abs/2308.01383>

11.2.3 QUANTUM COSMOLOGY: FORMULATING AN INITIAL QUANTUM STATE OF THE COSMOS

My section 1 also mentioned quantum cosmology's efforts to formulate an initial quantum state of the cosmos, including Hartle and Hawking's no-boundary proposal of 1983. This is a central topic of the popular book *On the Origin of Time* by T. Hertog, cited in 1.2 and 1.4 above. In a large body of literature, a very fine recent research article is:

- Halliwell, Jonathan, Jim Hartle, and Thomas Hertog. "What is the No-Boundary Wave Function of the Universe?" *Physical Review D* (2019); available at: <https://arxiv.org/abs/1812.01760>

11.2.4 COSMOLOGY AND THE UNDERDETERMINATION OF THEORY BY DATA

The stupendous achievements of modern cosmology, reviewed in section 2, naturally prompt the question: "Can all these details about the physics in places and times so very distant from us now really be established?" The worry, in philosophical terminology, is the underdetermination of theory by data; and *prima facie*, it seems this must be a big problem for cosmology.

About this, I will here cite three survey articles (two by me). Below, I cite more specialized articles, especially about the problems of confirming inflation.

- Butterfield, Jeremy. "On Under-determination in Cosmology." *Studies in the History and Philosophy of Modern Physics*, 46 (2014), 57–69; available at: <https://arxiv.org/abs/1406.4747>; <https://philsci-archive.pitt.edu/9866/>; <https://www.sciencedirect.com/science/article/abs/pii/S1355219813000476>
- Azhar, Feraz, and Jeremy Butterfield. "Scientific Realism and Primordial Cosmology"; available at <https://arxiv.org/abs/1606.04071>; <https://philsci-archive.pitt.edu/12192/>
- Smeenk, Chris. "Philosophical Aspects of Cosmology." *The Oxford Handbook of the History of Modern Cosmology*. Edited by Kragh, Helga and Malcolm Longair. Oxford University Press, 2019, cited in (1.3) above; available at: <https://doi.org/10.1093/oxfordhb/9780198817666.013.13>

11.2.5 CONFIRMING INFLATION

Some articles about the specific difficulties of confirming inflation (and doubts about its providing explanations), in chronological order:

- Earman, John. *Bangs, Crunches, Whimpers and Shrieks: Singularities and Acausalities in Relativistic Space-Times*. Oxford University Press, 1995; chapter 5. Cited in (1.2) above.
- Smeenk, Chris. “Predictability Crisis in Early Universe Cosmology.” *Studies in History and Philosophy of Modern Physics* 46 (2014), 122–133
- McCoy, Casey. “Does Inflation Solve the Hot Big Bang Model’s Fine-Tuning Problems?” *Studies in History and Philosophy of Modern Physics* 51 (2015), 23–36. Available at: <https://www.sciencedirect.com/science/article/abs/pii/>
- Smeenk, Chris. “Testing Inflation.” Edited by Khalil Chamcham et al., cited in (1.2) above. The article is available at <https://www.cambridge.org/core/books/philosophy-of-cosmology/testing-inflation/4095B5F8D7991E344D203CCBE0B369C8>
- Koberinski, Adam, and Chris Smeenk. “Establishing a Theory of Inflationary Cosmology.” *British Journal for the Philosophy of Science* 2024; available at: <https://doi.org/10.1086/733886>
- Azhar, Feraz, and Niels Linnemann. “Rethinking the Anthropic Principle,” in *Philosophy of Science* 2025; available at: <https://www.cambridge.org/core/journals/philosophy-of-science/article/rethinking-the-anthropic-principle/D043C5DE013A2E48700E306658D4E9AF>; <https://philsci-archive.pitt.edu/23934/>

11.2.6 CONFIRMING THEORIES OF DARK MATTER AND DARK ENERGY

Finally, here are two articles that also cover the problems of confirming theories of dark matter and dark energy (mentioned in section 2 above).

- Longair, Malcolm, and Chris Smeenk. “Inflation, Dark Matter, and Dark Energy.” Edited by Kragh, Helga and Malcolm Longair, cited in (1.2) above; available at: <https://doi.org/10.1093/oxfordhb/9780198817666.013.11>
- Pedro G. Ferreira, William J. Wolf, and James Read. “The Spectre of Underdetermination in Modern Cosmology”; available at: <https://arxiv.org/abs/2501.06095>; <https://philsci-archive.pitt.edu/24537/>

11.2.7 OBSERVATION SELECTION EFFECTS (IN GENERAL)

About observation selection effects in general, here are two items. The first is about supersymmetry, whose confirmational difficulties were a topic in section 7. It is a Bayesian analysis, and so far as I know, the most recent one.

- Dawid, Richard, and James Wells. “A Bayesian Model of Credence in Low Energy Supersymmetry.” (2024), available at: <https://philsci-archive.pitt.edu/24172/>

The second item raises a fun historical point. It returns us to Eddington’s famous metaphor of the net, in his 1938 book, *The Philosophy of Physical Science* (Cambridge University Press). As I quoted it in section 8:

Let us suppose that an ichthyologist is exploring the life of the ocean. He casts a net into the water and brings up a fishy assortment. Surveying his catch, he proceeds in the usual manner of a scientist to systematise what it reveals. He arrives at two generalisations: (1) No sea creature is less than two inches long. (2) All sea creatures have gills. These are both true of his catch, and he assumes tentatively that they will remain true however often he repeats it.

Although that passage is famous, philosophers should also take notice of—and take encouragement from!—what Eddington goes on to say just afterwards, which is almost never quoted. For Eddington takes the net to stand, not just for our means of observation in the specific science “ichthyology” (so that naively, we might infer that all fishes are longer than two inches), but also for our scientific method as a whole. Thus, Eddington’s moral is not just the obvious one I stressed in section 8, viz. “conditionalize your credence on your means of observation,” but also that we should allow for types of knowledge inaccessible to the scientific method. This open-mindedness is bound to be welcome to a philosopher . . . Thus, Eddington writes:

In applying this analogy, the catch stands for the body of knowledge which constitutes physical science, and the net for the sensory and intellectual equipment which we use in obtaining it. The casting of the net corresponds to observation; for knowledge which has not been or could not be obtained by observation is not admitted into physical science. An onlooker may object that the first generalisation is wrong. “There are plenty of sea-creatures under two inches long, only your net is not adapted to catch them.” The ichthyologist dismisses this objection contemptuously. “Anything uncatchable by my net is *ipso facto* outside the scope of ichthyological knowledge. In short, “what my net can’t catch isn’t fish.” Or—to translate the analogy—“If you are not simply guessing, you are claiming a knowledge of the physical universe discovered in some other way than by the methods of physical science, and admittedly unverifiable by such methods. You are a metaphysician. Bah!”

11.2.8 OBSERVATION SELECTION EFFECTS IN COSMOLOGY

About observation selection effects in cosmology (section 9), I begin with two survey articles. The first is philosophical and sceptical about inferring from fine tuning to a multiverse. The second is scientific (focused on fine-tuning of conditions for stars, planets, and life) and more optimistic about such inferences.

- Landsman, Klaas. “The Fine-Tuning Argument.” In *The Challenge of Chance*. Springer, 2016. Edited by Landsman, Klaas and E. van Wolde. Available at: <https://library.oapen.org/bitstream/handle/20.500.12657/27974/1/1002025.pdf#page=115>; <https://www.math.ru.nl/~landsman/eprints.html> (accessed January 13, 2026).
- Livio, Mario, and Martin Rees. “Fine-Tuning, Complexity, and Life in the Multiverse.” In *Fine-Tuning in the Physical Universe*. Edited by Sloan, David, et al., cited in (11.1.2) above. Available at: <https://arxiv.org/abs/1801.06944>

11.2.9 THE COSMOLOGICAL CONSTANT

I turn to the cosmological constant Λ and anthropic explanations of it (as in the second half of section 9). Here are some important sources:

- Earman, John “Lambda: The Constant that Refuses to Die.” *Archive for History of the Exact Sciences* (2001). This is a superb historico-philosophical overview of the cosmological constant. It is with almost all Earman’s other outstanding work in philosophy of physics, available at: <https://sites.pitt.edu/~jearman/> (last modified January 16, 2013).
- Weinberg, Steven et al. provide an anthropic explanation of the value of Λ . It is discussed in two articles in the collection, *Universe or Multiverse?* Edited by Carr, Bernard (Cambridge University Press, 2007), which I recommended at the start of 11.1.2 above.
- The first article is by Weinberg himself: “Living in the Multiverse,” which is also available at: <https://www.cambridge.org/core/books/philosophy-of-cosmology/testing-inflation/4095B5F8D7991E344D203CCBE0B369C8>. (My quotation from Weinberg in section 9 is from his 1987 paper in “Physical Review Letters.”)

- The second article (more detailed than Weinberg's) is Vilenkin, A. "Anthropic Predictions: The Case of the Cosmological Constant." It is available at: <https://arxiv.org/abs/astro-ph/0407586>

11.2.10 CONFIRMING MULTIVERSE THEORIES

Here are some details about Hartle and Srednicki's proposals for how to confirm multiverse theories. I outlined this in section 10. My three headings were "Measure," "Conditionalization," and "Typicality" were for the problems that such confirmation faces.

- They were an amalgamation of the seven problems listed by Aguirre, Anthony, in his excellent analysis: "Making Predictions in a Multiverse: Conundrums, Dangers, Coincidences," in the collection, *Universe or Multiverse?* (Edited by Carr, Bernard.) I recommended this at the start of 11.1.2 above.

Hartle and Srednicki's first two papers are:

- Hartle, James, and Mark Srednicki. "Are we typical?" *Physical Review D* 75: 123523 (2007); available at: <https://arxiv.org/abs/0704.2630>
- Hartle, James, and Mark Srednicki. "Science in a Very Large Universe." *Physical Review D* 81: 123524 (2010); available at: <https://arxiv.org/abs/0906.0042>

Later work along the same lines, with T. Hertog, includes:

- Hartle, James, and Thomas Hertog. "The Observer Strikes Back." In *The Philosophy of Cosmology* (2015), edited by Chamcham, Khalil et al., cited in (1.2) above, and available at: <https://arxiv.org/abs/1503.07205>
- Hartle, James, and Thomas Hertog. "One Bubble to Rule Them All." *Physical Review D* 95, 123502, (2017); available at <https://arxiv.org/abs/1604.03580>

- Hartle, James, and Thomas Hertog also connect an anthropic explanation of the cosmological constant Λ (cf. Weinberg et al. in (11.2.9)) with the no-boundary proposal discussed in (11.2.3), in their paper: “Anthropic Bounds on Λ from the No-Boundary Quantum State.” *Physical Review D* 88: 123516. It is available at: <https://arxiv.org/abs/1309.0493>

Finally, here is a (positive) philosophical assessment of Hartle and Srednicki’s proposals:

- Azhar, Feraz. “Three Aspects of Typicality in Multiverse Cosmology.” In *EPSA 2015* (European Philosophy of Science Association 2015 Conference Proceedings), edited by Romeijn, Jan-Willem, Michela Massimi, and Gerhard Schurz; Springer, 2016; available at: <https://arxiv.org/abs/1609.02586>

