Part I

Overviews

1

Introduction and overview

Bernard Carr
Astronomy Unit, Queen Mary, University of London

1.1 Introducing the multiverse

Nearly thirty years ago I wrote an article in the journal *Nature* with Martin Rees [1], bringing together all of the known constraints on the physical characteristics of the Universe – including the fine-tunings of the physical constants – which seemed to be necessary for the emergence of life. Such constraints had been dubbed 'anthropic' by Brandon Carter [2] – after the Greek word for 'man' – although it is now appreciated that this is a misnomer, since there is no reason to associate the fine-tunings with mankind in particular. We considered both the 'weak' anthropic principle – which accepts the laws of nature and physical constants as given and claims that the existence of observers then imposes a selection effect on where and when we observe the Universe – and the 'strong' anthropic principle – which (in the sense we used the term) suggests that the existence of observers imposes constraints on the physical constants themselves.

Anthropic claims – at least in their strong form – were regarded with a certain amount of disdain by physicists at the time, and in some quarters they still are. Although we took the view that any sort of explanation for the observed fine-tunings was better than none, many regarded anthropic arguments as going beyond legitimate science. The fact that some people of a theological disposition interpreted the claims as evidence for a Creator – attributing teleological significance to the strong anthropic principle – perhaps enhanced that reaction. However, attitudes have changed considerably since then. This is not so much because the status of the anthropic arguments themselves have changed – as we will see in a later chapter, some of them have become firmer and others weaker. Rather, it is because there has been a fundamental shift in the epistemological status of the anthropic principle. This arises because cosmologists have come to realize that there are many

Universe or Multiverse?, ed. Bernard Carr. Published by Cambridge University Press.

contexts in which our universe could be just one of a (possibly infinite) ensemble of 'parallel' universes in which the physical constants vary. This ensemble is sometimes described as a 'multiverse', and this term is used pervasively in this volume (including the title). However, it must be stressed that many other terms are used – sometimes even in the same context.

These multiverse proposals have not generally been motivated by an attempt to explain the anthropic fine-tunings; most of them have arisen independently out of developments in cosmology and particle physics. Nevertheless, it now seems clear that the two concepts are inherently interlinked. For if there *are* many universes, this begs the question of why we inhabit this particular one, and – at the very least – one would have to concede that our own existence is a relevant selection effect. Indeed, since we necessarily reside in one of the life-conducive universes, the multiverse picture reduces the strong anthropic principle to an aspect of the weak one. For this reason, many physicists would regard the multiverse proposal as providing the most natural explanation of the anthropic fine-tunings.

One reason that the multiverse proposal is now popular is that it seems to be necessary in order to understand the origin of the Universe. Admittedly, cosmologists have widely differing views on how the different worlds might arise. Some invoke models in which a single universe undergoes cycles of expansion and recollapse, with the constants being changed at each bounce [3]. In this case, the different universes are strung out in *time*. Others invoke the 'inflationary' scenario [4], in which our observable domain is part of a single 'bubble' which underwent an extra-fast expansion phase at some early time. There are many other bubbles, each with different laws of low-energy physics, so in this case the different universes are spread out in *space*. As a variant of this idea, Andrei Linde [5] and Alex Vilenkin [6] have invoked 'eternal' inflation, in which each universe is continually self-reproducing, since this predicts that there may be an infinite number of domains – all with different coupling constants. The different universes then extend in *both* space and time.

On the other hand, Stephen Hawking prefers a quantum cosmological explanation for the Universe and has objected to eternal inflation on the grounds that it extends to the infinite past and is thus incompatible with the Hartle–Hawking 'no boundary' proposal for the origin of the Universe [7]. This requires that the Universe started at a finite time but the initial singularity of the classical model is regularized by requiring time to become imaginary there. If one uses the path integral approach to calculate the probability of a particular history, this appears to favour very few expansion e-folds, so the Universe would recollapse too quickly for life to arise.

However, anthropic selection can salvage this, since one only considers histories containing observers [8].

This sort of approach to quantum cosmology only makes sense within the context of the 'many worlds' interpretation of quantum mechanics. This interpretation was suggested by Hugh Everett [9] in the 1950s in order to avoid having to invoke collapse of the quantum mechanical wave-function, an essential feature of the standard Copenhagen interpretation. Instead, our universe is supposed to split every time an observation is made, so one rapidly generates a huge number of parallel worlds [10]. This could be regarded as the earliest multiverse theory. Although one might want to distinguish between classical and quantum multiverses, Max Tegmark [11] has emphasized that there is no fundamental distinction between them.

Quantum theory, of course, originated out of attempts to explain the behaviour of matter on small scales. Recent developments in particle physics have led to the popularity of yet another type of multiverse. The holy grail of particle physics is to find a 'Theory of Everything' (TOE) which unifies all the known forces of physics. Models which unify the weak, strong and electomagnetic interactions are commonly described as 'Grand Unified Theories' (GUTs) and – although still unverified experimentally – have been around for nearly 30 years. Incorporating gravity into this unification has proved more difficult, but recently there have been exciting strides, with superstring theory being the currently favoured model.[1] There are various versions of superstring theory but they are amalgamated in what is termed 'M-theory'.

Unlike the 'Standard Model', which excludes gravity and contains several dozen free parameters, M-theory might conceivably predict all the fundamental constants uniquely [12]. That at least has been the hope. However, recent developments suggest that this may not be the case and that the number of theories (i.e. vacuum states) could be enormous (for example 10^{500} [13]). This is sometimes described as the 'string landscape' scenario [14]. In this case, the dream that all the constants are uniquely determined would be dashed. There would be a huge number of possible universes (corresponding to different minima of the vacuum energy) and the values of the physical constants would be *contingent* (i.e. dependent on which universe we happen to occupy). Trying to predict the values of the constants would then be

1 String theory posits that the fundamental constituents of matter are string-like rather than point-like, with the various types of elementary particle corresponding to different excitation states of these strings. This was originally proposed as a model of strong interactions but in the 1980s it was realized that it could be extended to a version called 'superstring' theory, which also includes gravity.

as forlorn as Kepler's attempts to predict the spacing of the planets in our solar system based on the properties of Platonic solids.

A crucial feature of the string landscape proposal is that the vacuum energy would be manifested as what is termed a 'cosmological constant'. This is a term in the field equations of General Relativity (denoted by Λ) originally introduced by Einstein to allow a static cosmological model but then rejected after the Universe was found to be expanding. For many subsequent decades cosmologists assumed Λ was zero, without understanding why, but a remarkable recent development has been the discovery that the expansion of the Universe is accelerating under the influence of (what at least masquerades as) a cosmological constant. One possibility is that Λ arises through quantum vacuum effects. We do not know how to calculate these, but the most natural value would be the Planck density (which is 120 orders of magnitude larger than the observed value). Indeed in the string landscape proposal, one might expect the value of Λ across the different universes to have a uniform distribution, ranging from minus to plus the Planck value. The observed value therefore seems implausibly small.

There is also another fine-tuning problem, in that the observed vacuum density is currently very similar to the matter density, a coincidence which would only apply at a particular cosmological epoch. However, as first pointed out by Steven Weinberg [15, 16], the value of Λ is constrained anthropically because galaxies could not form if it were much larger than observed. This is not the only possible explanation for the smallness of Λ, but there is a reluctant acceptance that it may be the most plausible one, which is why both string landscape and anthropic ideas are rather popular at present. The crucial issue of whether the number of vacuum states is sufficiently large and their spacing sufficiently small to satisfy the anthropic constraints is still unresolved.

It should be noted that M-theory requires there to be extra dimensions beyond the four familiar ones of space and time. Some of these may be compactified, but others may be extended, in which case, the Universe would correspond to a 4-dimensional 'brane' in a higher-dimensional 'bulk' [17, 18]. In the first versions of this theory, the cosmological constant was negative, which was incompatible with the observed acceleration of the Universe. A few years ago, however, it was realized that M-theory solutions with a positive cosmological constant are also possible [19], and this has revitalized the collaboration between cosmologists and string theorists. The notion that our universe is a brane in a higher-dimensional bulk also suggests another multiverse scenario, since there might be many other branes in the bulk. Collisions between these branes might even generate big bangs of the kind which initiated the expansion of our own universe [20]. Indeed, some people have envisaged successive collisions producing cyclic models, and it has been claimed that this could provide another (non-anthropic) explanation for why Λ naturally tends to a value comparable to the matter density [21].

1.2 Historical perspective

We have seen how a confluence of developments in cosmology and particle physics has led to a dramatic improvement in the credibility of the multiverse proposal. In this section, we will put these developments into a historical perspective, by showing how the notion of the multiverse is just the culmination of attempts to understand the physics of the largest and smallest scales. For what we regard as the 'Universe' has constantly changed as scientific progress has extended observations outwards to ever larger scales and inwards to ever smaller ones. In the process, it has constantly revealed new levels of structure in the world, as well as interesting connections between the laws operating at these different levels. This section will also provide an opportunity to review some of the basic ideas of modern cosmology and particle physics, which may be useful for non-specialists.

1.2.1 The outward journey

Geocentric view

Early humans assumed that the Earth was the centre of the Universe. Astronomical events were interpreted as being much closer than they actually are, because the heavens were assumed to be the domain of the divine and therefore perfect and unchanging. The Greeks, for example, believed the Earth was at the centre of a series of 'crystal spheres', these becoming progressively more perfect as one moves outwards. The last one was associated with the immovable stars, so transient phenomena (like meteors and comets) were assumed to be of terrestrial origin. Even the laws of nature (such as the regularity of the seasons) seemed to be human-centred, in the sense that they could be exploited for our own purposes, so it was natural to regard them as a direct testimony to our central role in the world.

Heliocentric view

In 1542 Nicolaus Copernicus argued in *De Revolutionis Orbis* that the heliocentric picture provides a simpler explanation of planetary motions than the geocentric one, thereby removing the Earth from the centre of the Universe. The heliocentric picture had earlier been suggested by Artistarchus, although this was regarded as blasphemous by most of his fellow Greeks, and Nicholas de Cusa, who in 1444 argued that the Universe had no centre and looks the same everywhere. Today this notion is called the Copernican or Cosmological Principle. Then in 1572 Tycho Brahe spotted a supernova in the constellation of Cassiopeia; it brightened suddenly and then dimmed over the course of a year, but the fact that its apparent position did not change as the Earth moved around the Sun implied that it was well beyond the Moon. Because this destroyed the Aristotelian view that the heavens never change, the claim was at first received sceptically. Frustrated by those who had eyes but would not see, Brahe wrote in the preface of *De Nova Stella*: 'O crassa ingenia. O coecos coeli spectators.' (Oh thick wits. Oh blind watchers of the sky.)

Galactocentric view

The next step occurred when Galileo Galilei used the newly invented telescope to show that not even the Sun is special. His observations of sunspots showed that it changes, and in 1610 he speculated in *The Sidereal Message* that the Milky Way – then known as a band of light in the sky but now known to be the Galaxy – consists of stars like the Sun but at such a great distance that they cannot be resolved. This not only cast doubt on the heliocentric view, but also vastly increased the size of the Universe. An equally profound shift in our view of the Universe came a few decades later with Isaac Newton's discovery of universal gravity. By linking astronomical phenomena to those on Earth, Newton removed the special status of the heavens, and the publication of his *Principia* in 1687 led to the 'mechanistic' view in which the Universe is regarded as a giant machine. In the following century, the development of more powerful telescopes – coupled with Newton's laws – enabled astronomers to understand the structure of the Milky Way. In 1750 Thomas Wright proposed that this is a disc of stars, and in 1755 Immanuel Kant speculated that some nebulae are 'island universes' similar to the Milky Way, raising the possibility that even the Galaxy is not so special. However, the galactocentric view persisted for several more centuries, with most astronomers still assuming that the Milky Way comprised the whole Universe. Indeed this was Einstein's belief when he published his theory of General Relativity in 1915 and started to study its cosmological implications.

Cosmocentric view

Then in the 1920s the idea anticipated by Kant – that some of the nebulae are outside the Milky Way – began to take hold. For a while this was a matter of intense controversy. In 1920 Heber Curtis vigorously defended the island universe theory in a famous debate with Harlow Shapley. The controversy was finally resolved in 1924 when Edwin Hubble announced that he had measured the distance to M31 using Cepheid stars. An even more dramatic revelation came in 1929, when Hubble obtained radial velocities and distance estimates for several dozen nearby galaxies, thereby discovering that all galaxies are moving away from us with a speed proportional to their distance. This is now called 'Hubble's law' and it has been shown to apply out to a distance of 10 billion light-years, a region containing 100 billion galaxies. The most natural interpretation of Hubble's law is that space itself is expanding, as indeed had been predicted by Alexander Friedmann in 1920 on the basis of general relativity. Friedmann's model suggested that the Universe began in a state of great compression at a time in the past of order the inverse of the Hubble constant, now known to be about 14 billion years. This is the 'Big Bang' picture, and it received decisive support in 1965 with the discovery that the Universe is bathed in a sea of background radiation. This radiation is found to have the same temperature in every direction and to have a black-body spectrum, implying that the Universe must once have been sufficiently compressed for the radiation to have interacted with the matter. Subsequent studies by the COBE satellite confirmed that it has a perfect black-body spectrum, which firmly established the Big Bang theory as a branch of mainstream physics.

Multiverse view

Further studies of the background radiation – most notably by the WMAP satellite – have revealed the tiny temperature fluctuations associated with the density ripples which eventually led to the formation of galaxies and clusters of galaxies. The angular dependence of these ripples is exactly as predicted by the inflationary scenario, which suggests that our observable domain is just a tiny patch of a much larger universe. This was the first evidence for what Tegmark [11] describes as the 'Level I' multiverse. A still more dramatic revelation has been the discovery – from observations of distant supernovae – that the expansion of the Universe is accelerating. We don't know for sure what is causing this, but it is probably related to the vacuum energy density. As described in Section 1.1, the low value of this density may indicate that there exist many other universes with different vacuum states, so this may be evidence for Tegmark's 'Level II' multiverse.

This brief historical review of developments on the outer front illustrates that the longer we have studied the Universe, the larger it has become. Indeed, the multiverse might be regarded as just one more step in the sequence of expanding vistas opened up by cosmological progress (from geocentric to heliocentric to galactocentric to cosmocentric). More conservative cosmologists might prefer to maintain the cosmocentric view that ours is the only Universe, but perhaps the tide of history is against them.

1.2.2 The inward journey

Equally dramatic changes of perspective have come from revelations on the inward front, with the advent of atomic theory in the eighteenth century, the discovery of subatomic particles at the start of the twentieth century and the advent of quantum theory shortly thereafter. The crucial achievement of the inward journey is that it has revealed that everything in the Universe is made up of a few fundamental particles and that these interact through four forces: gravity, electromagnetism, the weak force and the strong force. These interactions have different strengths and characteristics, and it used to be thought that they operated independently. However, it is now thought that some (and possibly all) of them can be unified as part of a single interaction.

Figure 1.1 illustrates that the history of physics might be regarded as the history of this unification. Electricity and magnetism were combined by Maxwell's theory of electromagnetism in the nineteenth century. The electromagnetic force was then combined with the weak force in the (now experimentally confirmed) electroweak theory in the 1970s. Theorists have subsequently merged the electroweak force with the strong force as part of the Grand Unified Theory (GUT), although this has still not been verified experimentally. As discussed in Section 1.1, the final (and as yet incomplete) step is the unification with gravity, as attempted by string theory or M-theory.

A remarkable feature of these theories is that the Universe may have more than the three dimensions of space that we actually observe, with the extra dimensions being compactified on the Planck scale (the distance of 10^{-33} cm at which quantum gravity effects become important), so that we do not notice them. In M-theory itself, the total number of dimensions (including time) is eleven, with 4-dimensional physics emerging from the way in which the extra dimensions are compactified (described by what is called a Calabi–Yau manifold). The discovery of dark dimensions through particle physics shakes our view of the nature of reality just as profoundly as the discovery of dark energy through cosmology. Indeed, we saw in Section 1.1 that there may be an intimate link between these ideas.

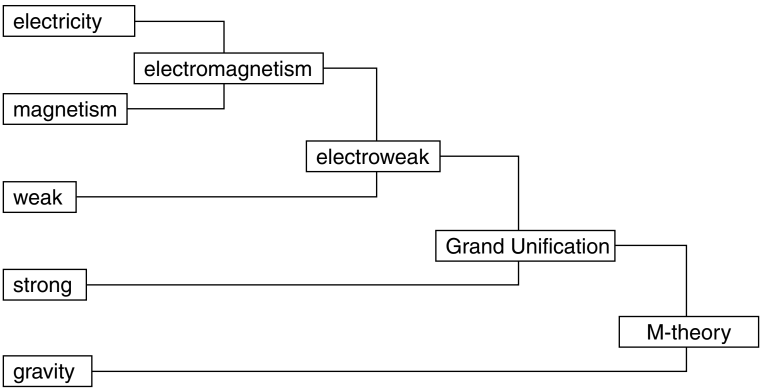

Fig. 1.1. This shows the successive steps by which physics has attempted to unify the four known forces of nature. Time runs to the right.

1.2.3 The cosmic uroborus

Taken together, scientific progress on both the outer and inner fronts can certainly be regarded as a triumph. In particular, physics has revealed a unity about the Universe which makes it clear that everything is connected in a way which would have seemed inconceivable a few decades ago. This unity is succinctly encapsulated in the image of the uroborus (i.e. the snake eating its own tail). This is shown in Fig. 1.2 (adapted from a picture originally presented by Sheldon Glashow) and demonstrates the intimate link between the macroscopic domain (on the left) and the microscopic domain (on the right).

The pictures drawn around the snake represent the different types of structure which exist in the Universe. Near the bottom are human beings. As we move to the left, we encounter successively larger objects: a mountain, a planet, a star, the solar system, a galaxy, a cluster of galaxies and finally the entire observable Universe. As we move to the right, we encounter successively smaller objects: a cell, a DNA molecule, an atom, a nucleus, a quark, the GUT scale and finally the Planck length. The numbers at the edge indicate the scale of these structures in centimetres. As one moves clockwise from the tail to the head, the scale increases through 60 decades: from the smallest meaningful scale allowed by quantum gravity (10^{-33} cm) to the scale of the visible Universe (10^{27} cm). If one expresses these scales in units of the Planck length, they go from 0 to 60, so the uroborus provides a sort of 'clock' in which each 'minute' corresponds to a factor of 10 in scale.