1 Introduction

This book is about 13.8 billion years of cosmic evolution. We will trace the history of the universe from fractions of a second after the Big Bang until today. Before we begin our journey, however, it will be useful to have a bird's eye view of the subject. In this chapter, I will therefore set the stage by giving a qualitative account of the structure and evolution of the universe, as we now understand it. In the rest of the book, I will then show you where this knowledge comes from.

1.1 Scales of the Universe

The universe is big. In fact, the length scales in cosmology are so enormous that they are hard to grasp (see Table 1.1). Evolution simply hasn't equipped the human brain with the ability to have any intuition for the vastness of the cosmos. Apparently, being able to visualize the size of the universe has not increased the survival chances of our species.

Let us start with the **Solar System**. We live on the third planet from the Sun. Our nearest neighbour—at a distance of 384 400 km (about sixty times the size of the Earth)—is the Moon. Compared to the other objects of the Solar System, the Moon is very close. Although the Apollo spacecraft took three days to travel to the Moon, light makes the trip in just 1.3 seconds. For comparison, the light from the Sun takes over eight minutes to reach us. That's a distance of 150 million km (or 11 780 Earths lined up side to side).

Table 1.1 Important length scales of the universe (in different units)

Object	Size [km]	Size [ly]	Size [Mpc]
Earth	6371	6.7×10^{-10}	2.1×10^{-16}
Distance to Sun	1.5×10^{8}	1.6×10^{-5}	4.8×10^{-12}
Solar System	4.5×10^{9}	4.7×10^{-4}	1.5×10^{-10}
Milky Way Galaxy	1.0×10^{18}	105 700	0.032
Local Group	9×10^{19}	9×10^{6}	3
Local Supercluster	5×10^{21}	5×10^{8}	150
Universe	4.4×10^{23}	46.5 billion	14 000

Cambridge University Press & Assessment
978-1-108-83807-8 — Cosmology
Daniel Baumann
Excerpt
More Information

While it takes light only a few hours to cross the Solar System, the light from even the nearest stars takes several years to reach us. To measure the distances of objects that lie beyond our Solar System, it is therefore convenient to introduce the distance traveled by light in one year. A *light-year* (ly) is

$$1\ \text{ly} \approx 9.5 \times 10^{15}\,\text{m}\,. \tag{1.1}$$

Besides light-years, astronomers also use *parsecs* (pc). As the Earth moves around the Sun, the positions of nearby stars shift ever so slightly, because they are being viewed from different directions. Given the size of an observed shift and the known radius of the Earth's orbit, the distance to a star follows from simple trigonometry. A parsec is defined as the distance at which the radius of the Earth's orbit around the Sun subtends an angle of one arcsecond (i.e. 1/3600 of a degree). Expressed in terms of light-years, this is

$$1\ \text{pc} \approx 3.26\ \text{ly}\,. \tag{1.2}$$

The nearest star, Proxima Centauri, is 4.2 light-years (or 1.3 pc) away.

Our Galaxy, the **Milky Way**, contains about 100 billion stars arranged in a flattened disc. It is about 30 kpc across, which means that it takes light about 100 000 years to cross our Galaxy. The Sun is located at the edge of a spiral arm, 30 000 light-years from the center. It takes 250 million years for the Sun to complete one orbit around the center of the Galaxy.

There are about 100 billion galaxies in the observable universe, each with about 100 billion stars. Our nearest neighbour is the Andromeda Galaxy (M31), a spiral galaxy about 2 million light-years away. The distances to galaxies are therefore measured in Mpc, which will be the unit of choice in cosmology. Andromeda is one of about fifty nearby galaxies that are gravitationally bound together. This arrangement of galaxies, called the **Local Group**, spans about 10 million light-years.

On even larger scales, galaxies arrange themselves into clusters and superclusters, with filamentary structures and giant voids in between them. Although the clusters within the superclusters are gravitationally bound, they are spreading apart as the universe is expanding. The largest such structures, like our **Local Supercluster** (Laniakea), are about 500 million light-years across.

Because the universe has a finite age, there is a maximum distance that we can see. This radius of the **observable universe** is

$$46.5\,\text{billion ly} \approx 14\,\text{Gpc} \approx 4.4 \times 10^{26}\,\text{m}\,. \tag{1.3}$$

Notice that this is larger than the naive distance that light traveled in the age of the universe (which is 13.8 billion years). This discrepancy arises because the universe is expanding and the light is carried along with the expansion.

Having developed some sense for the vastness of the universe, let us now discuss what it is made of.

Cambridge University Press & Assessment
978-1-108-83807-8 — Cosmology
Daniel Baumann
Excerpt
More Information

1.2 The Invisible Universe

Most of the matter and energy in the universe is invisible. The stuff that we can see—ordinary atoms—accounts for less than 5% of the total. The rest is in the form of dark matter and dark energy.

The majority of mass in the universe is composed of **dark matter**. This invisible matter is required to explain the stability and growth of structure in the universe. Early evidence for the existence of dark matter was collected by Fritz Zwicky in 1933. When studying galaxies in the Coma cluster, he found that they were moving faster than allowed by the gravity of the visible matter alone and an extra “dunkle materie” seemed to be required to hold the cluster together [1]. It took a long time, however, for the existence of dark matter to become accepted by the astronomical community. Important further evidence for the existence of dark matter came in the 1970s when Vera Rubin and collaborators measured the rotation speeds of hydrogen gas in the outer reaches of galaxies [2]. The large speeds that they found could only be explained if these galaxies were embedded in halos of dark matter.[1]

Today, some of the most striking evidence for dark matter comes from the gravitational lensing of the cosmic microwave background (CMB). As the CMB photons travel through the universe, they get deflected by the intervening large-scale structure, which distorts the hot and cold spots of the CMB. The effect depends on the total amount of matter in the universe and has been measured by the Planck satellite [5]. At the same time, the observed abundances of the light elements (D, He, Li), which were created by Big Bang nucleosynthesis (see below), imply a smaller amount of ordinary baryonic matter. The mismatch between the two measurements points to the existence of non-baryonic dark matter. The same amount of dark matter also explains the pattern of CMB anisotropies and the rate of gravitational clustering. In particular, the small density variations observed in the early universe only grow fast enough if assisted by the gravitational pull of the dark matter. Although the influence of dark matter is now seen over a wide range of scales, the true nature of the putative dark matter remains a pressing open problem in astrophysics and cosmology.

In the 1980s, cosmology was in crisis; the age of a matter-only universe seemed to be shorter than the ages of the oldest stars within it. In addition, observations of the large-scale clustering of galaxies implied that the total matter density was only around 30% of the critical density required for a spatially flat universe [6, 7], in conflict with the theoretical expectation from inflationary cosmology [8]. To account for this “missing energy,” some theorists suggested the existence of a new form of **dark energy** [7, 9, 10], but direct evidence for it was lacking. Both problems were resolved with the discovery of the accelerating universe [11, 12]. By observing distant

---

[1] The history of the dark matter problem is considerably more complex and nuanced than this brief description might suggest. More on the fascinating history of dark matter can be found in [3, 4].

supernova explosions, cosmologists were able to measure both the distances and recession speeds of far-away galaxies. The results showed that the rate of expansion was decelerating at early times (as expected for a universe dominated by ordinary matter), but started to *accelerate* a few billion years ago. As we will see, such an accelerated expansion is possible in Einstein's theory of gravity if the universe is filled with an energy density that doesn't dilute and exerts a negative pressure. The accelerated expansion increases the estimate for the age of the universe, reconciling it with the ages of the oldest stars. Moreover, shortly after the supernova observations, balloon-borne CMB experiments [13, 14] provided strong observational evidence for the spatial flatness of the universe and hence solidified the case that dark energy was needed to explain 70% of the total energy in the universe today.

However, while dark energy solves the age problem, it has led to a new crisis in cosmology. We have given dark energy a name, but we don't know what it is. A natural candidate for dark energy is the energy density of empty space itself, since it doesn't dilute with the expansion of space. In fact, such a "vacuum energy" is predicted by quantum field theory, but its estimated size is many orders of magnitude larger than the observed dark energy density. Explaining this discrepancy remains one of the biggest open problems in cosmology and fundamental physics [15].

1.3 The Hot Big Bang

The universe is expanding [16]. It was therefore denser and hotter in the past. Particles were colliding frequently and the universe was in a state of thermal equilibrium with an associated temperature T. It is convenient to set Boltzmann's constant to unity, $k_\mathrm{B} = 1$, and measure temperature in units of energy. Moreover, we will often use the particle physicists' convention of measuring energies in *electron volt* (eV):

$$\begin{aligned} 1\,\mathrm{eV} &\approx 1.60 \times 10^{-19}\,\mathrm{J} \\ &\approx 1.16 \times 10^{4}\,\mathrm{K}\,. \end{aligned} \tag{1.4}$$

For reference, typical atomic processes are measured in eV, while the characteristic scale of nuclear reactions is MeV. A useful relation between the temperature of the early universe and its age is

$$\frac{T}{1\,\mathrm{MeV}} \approx \left(\frac{t}{1\,\mathrm{s}}\right)^{-1/2} . \tag{1.5}$$

One second after the Big Bang the temperature of the universe was therefore about 1 MeV (or 10^{10} K). While there was very little time available in the early universe, the rates of reactions were extremely high, so that many interesting things happened in a short amount of time (see Table 1.2).

Above 100 GeV (or a trillionth of a second after the Big Bang), all particles of the Standard Model were in equilibrium and were therefore present in roughly equal

Table 1.2 Key events in the history of the universe

Event	Temperature	Energy	Time
Inflation	$< 10^{29}$ K	$< 10^{16}$ GeV	$> 10^{-34}$ s
Dark matter decouples	?	?	?
Baryons form	?	?	?
EW phase transition	10^{15} K	100 GeV	10^{-11} s
Hadrons form	10^{12} K	150 MeV	10^{-5} s
Neutrinos decouple	10^{10} K	1 MeV	1 s
Nuclei form	10^{9} K	100 keV	200 s
Atoms form	3460 K	0.29 eV	290 000 yrs
Photons decouple	2970 K	0.25 eV	370 000 yrs
First stars	50 K	4 meV	100 million yrs
First galaxies	20 K	1.7 meV	1 billion yrs
Dark energy	3.8 K	0.33 meV	9 billion yrs
Einstein born	2.7 K	0.24 meV	13.8 billion yrs

abundances. This state can be viewed as the initial condition for the hot Big Bang. The density at that time was a staggering $10^{36}\,\mathrm{kg\,cm^{-3}}$, which is what you would get if you compressed the mass of the Sun to the size of a marble. In a billionth of a second, the universe expanded by a factor of 10^4. During this expansion, the temperature dropped and the universe went through different evolutionary stages.

At around 100 GeV (or 10^{15} K), the electroweak (EW) symmetry of the Standard Model was broken during the **EW phase transition**. The electromagnetic and weak nuclear forces became distinct entities and the matter particles received their masses. Although the basics of EW symmetry breaking are well understood—and have been experimentally verified by the discovery of the Higgs boson [17, 18]—the detailed dynamics of the EW phase transition and its observational consequences are still a topic of active research.

Once the temperature drops below the mass of a particle species, particles and antiparticles start to annihilate, while the reverse process—the spontaneous creation of particle–antiparticle pairs—becomes inefficient. The first particles to disappear from the universe in this way were the top quarks (the heaviest particles of the Standard Model). W and Z bosons followed. Then the Higgs, the bottom and charm quarks and the tau lepton. Around 150 MeV, the **QCD phase transition** occurred and the remaining quarks condensed into hadrons (mostly protons, neutrons and pions).

Particles fall out of thermal equilibrium when their interaction rate drops below the expansion rate of the universe. At that moment, the particles stop interacting with the rest of the thermal bath and a relic abundance will be created. It is likely that the dark matter was created in this way, but the details are still unknown. A known decoupling event is the decoupling of neutrinos about *one second* after the Big Bang, which produced the **cosmic neutrino background** (CνB). During the early radiation-dominated period, these cosmic neutrinos carried about 40% of the total energy density and had a significant effect on the expansion of the universe.

Imprints of the cosmic neutrino background have recently been detected in both the CMB and the clustering of galaxies. These observations provide a window into the universe when it was just one second old.

After the QCD phase transition, the universe was a plasma of mostly free electrons, protons and neutrons, as well as very energetic photons that prevented any heavier nuclei from forming. About *one minute* after the Big Bang, the temperature of the universe had dropped enough for the synthesis of helium-4 and lithium-7 to become efficient [19–21]. The predicted amounts of these elements are consistent with the abundances found in early gas clouds, where very little post-processing of the primordial abundances has taken place. **Big Bang nucleosynthesis** (BBN) produced very few nuclei that are heavier than lithium, because there are no stable nuclei with 5 or 8 nucleons that would be required to sustain the nuclear chain reactions. Heavier elements were instead produced in the interior of stars, where the high densities and long timescales involved allow for three helium-4 nuclei to fuse. These heavier elements include carbon and oxygen which were spread throughout the universe after the stars exploded.

About *370 000 years* after the Big Bang, the universe had cooled enough for the first stable atoms to form in a process called **recombination**. At that moment, light stopped scattering off the free electrons in the plasma and started to propagate freely through the universe. These free-streaming photons are still seen today as an afterglow of the Big Bang. Stretched by 13.8 billion years of cosmic expansion, the universe's first light is observed as a faint microwave radiation [22]—the **cosmic microwave background**. The CMB contains tiny variations in its intensity (as a function of direction on the sky) that reflect perturbations in the density of matter in the early universe. The pattern of these fluctuations contains critical information about the primordial universe. More than any other cosmological probe, the study of the CMB anisotropies has transformed cosmology into a precision science.

What I have described so far are either facts (like BBN and recombination) or theoretical extrapolations that we can make with extremely high confidence (like the EW and QCD phase transitions). However, there are two important events in the history of the early universe that we know must have occurred, but whose details we are much less certain of. The first is **dark matter production**. Some process in the early universe must have led to the abundance of dark matter that we observe in the universe today. There are many ways in which this could have occurred, depending on the precise nature of the dark matter. For example, weakly interacting massive particles (WIMPs) could have decoupled from the primordial plasma at high energies, producing a cosmological abundance of dark matter. Or, a massive boson (maybe an axion) could have started to oscillate around the minimum of its potential when the expansion rate dropped below the mass of the particle, producing a bosonic condensate that acts like dark matter. The nature of the dark matter and its production in the early universe remain important open problems.

Another event that we believe must have occurred in the early universe, but whose details are unknown is **baryogenesis**. This refers to the mechanism by which an asymmetry was created between the amount of matter and antimatter in the universe. The required asymmetry is very small: for every 10^{10} particles of antimatter there must have been one extra matter particle. In this way, the annihilation between matter and antimatter into photons produces the observed matter-to-photon ratio in the universe. It isn't that we have no idea how this asymmetry might have been generated. In fact, there are many models of baryogenesis, but no way, so far, to decide which of these (if any) is the correct one.

1.4 Growth of Structure

The density fluctuations in the early universe eventually grew into all of the structures we see around us. On large scales, the gravitational clustering of matter can be described analytically, while, on small scales, the process becomes highly nonlinear and can only be captured by numerical simulations. The dark matter formed a web-like structure with high-density nodes connected by filaments. The baryonic gas collected in the regions of high dark matter density where it collapsed into stars which then congregated into galaxies.

The first stars—called **Population III stars**[2]—formed when the universe was about *100 million years* old. Computer simulations suggest that these stars were very massive, about a few hundred times more massive than the Sun. They were also very luminous and, hence, burned up their fuel rapidly. Although the first stars were short-lived, their impact on the universe was significant. They emitted ultraviolet light which heated and ionized the surrounding gas. The dynamics of this process of **reionization** are still not completely known and are actively investigated through numerical simulations. The first stars also may have provided the seeds for the growth of supermassive black holes which are found at the centers of most galaxies. And, finally, they created the first heavy elements in their interiors which were dispersed throughout the cosmos when the stars exploded. Enriched with these heavy elements, the baryonic gas cooled more efficiently, allowing smaller and more long-lived stars to be formed.

The first **galaxies** started to appear about *one billion years* after the Big Bang. Over time, these galaxies formed clusters and superclusters, a process that is still ongoing today. In the future, however, the growth of structure will stop as dark energy starts to dominate the universe. The details of galaxy formation are intricate and still an active area of research. In this book, we will be more interested in galaxies as tracers of the underlying distribution of dark matter, which in turn is determined by the seed fluctuations in the early universe. This distribution isn't random but has interesting spatial correlations which have been measured in large galaxy surveys.

---

[2] This terminology displays the full range of the weirdness of astronomical nomenclature. All elements heavier then helium are called "metals." The youngest metal-rich stars are called Population I stars, while older metal-poor stars are called Population II. By extension, the first stars, containing no metals, are called Population III stars.

1.5 Cosmic Palaeontology

Cosmology is famously an observational rather than an experimental science. No experimentalists were present in the early universe, and the birth and subsequent evolution of the universe cannot be repeated. Instead, we can only measure the spatial correlations between cosmological structures at late times. A central challenge of modern cosmology is to construct a consistent "history" of the universe that explains these correlations. This cosmological history is a narrative, a story we tell to give a rational accounting of the patterns that we see in the cosmological correlations.

This parallels the way palaeontologists infer the history of the Earth by studying the pattern of fossilized bones in the ground today. Like astronomical objects, these fossils are not randomly spread throughout space, but display interesting correlations which we try to explain by invoking past events. In much the same way, cosmologists study the pattern of cosmological structures observed today to infer the history of the early universe.

A remarkable feature of the observed correlations in the CMB is that they span scales that are larger than the distance traveled by light between the beginning of the hot Big Bang and the time when the CMB was created. This is in conflict with causality, unless the correlations were generated *before* the hot Big Bang. Indeed, there is now growing evidence that the hot Big Bang was not the beginning of time, but that the primordial density fluctuations were produced during an earlier period of accelerated expansion called **inflation** [8, 23, 24]. Small quantum fluctuations during inflation were amplified by the rapid expansion of the space and became the seeds for the large-scale structure of the universe [25–29].

If inflation really occurred, it was a rather dramatic event in the history of the universe. In just a billionth of a trillionth of a trillionth of a second, the universe doubled in size about 80 times. A region of space the size of a mosquito got stretched to the size of an entire galaxy. The entire observable universe then originated from a microscopic, causally-connected region of space. The correlations observed in the afterglow of the Big Bang were inherited from the correlations of the quantum fluctuations during inflation. While this picture provides an elegant explanation for the initial conditions of the primordial universe, it must be emphasized that inflation is *not* yet a fact—at the same level that, for example, BBN is a fact. The theoretical framework for inflation, however, is sufficiently well developed to justify including it in an introductory textbook on standard cosmology. Moreover, many new observations of the primordial correlations are being carried out—or are in the planning stages—that will subject the inflationary paradigm to further tests.

Further Reading

In writing this book, I have drawn on many excellent earlier treatments of cosmology. In the following, I will list some of the sources that I have found particularly useful. At the end of each chapter, I will give additional reading suggestions, including pointers to the research literature.

The following textbooks are all excellent and may be consulted for further details or alternative points of view:

- Dodelson, *Modern Cosmology*
  This book is already a modern classic. It is fantastically written and all explanations are exceptionally clear.
- Mukhanov, *Physical Foundations of Cosmology*
  Written by one of the architects of inflation, this book is especially good for early universe cosmology.
- Weinberg, *Cosmology*
  Sadly Steven Weinberg died during the writing of this book. He was a giant of theoretical physics and a hero of so many of us. His cosmology textbook is a great resource, especially for the more subtle aspects of the subject.
- Peter and Uzan, *Primordial Cosmology*
  This is a very comprehensive book, which despite its title also covers important parts of late-time cosmology.
- Kolb and Turner, *The Early Universe*
  Despite its advanced age, this book is still a very useful resource. Kolb and Turner's treatment of the thermal history of the universe remains unparalleled.
- Ryden, *Introduction to Cosmology*
  This is one of the few good cosmology books for undergraduates. Its level is lower than that of this book, but it is beautifully written and contains many very clear explanations.

Problems

1.1 The range of length scales involved in cosmology is hard to grasp. The best we can do is to consider relative distances and compare them to something more familiar. In this exercise, we will make some attempt at obtaining a more intuitive understanding of the vastness of the cosmos.

1. Consider shrinking the Earth to the size of a basketball. What would then be the size of the Moon and its orbit around the Earth?
2. Now imagine scaling the Earth down to the size of a peppercorn. What would then be the size of the Sun and the Earth's orbit? How far away would the most distant planet in the Solar System be?
3. The "Solar Neighbourhood" is a collection of about fifty nearby stars, spread across about 65 light-years, that travel together with the Sun. Scaling this region down, so that it fits inside a basketball court, what would be the size of the Solar System?
4. Shrinking our Galaxy to the size of the basketball court, what would now be the size of the Solar Neighbourhood?
5. The "Local Group" comprises about fifty nearby galaxies, spread across about 10 million light-years. If we squeezed this region into the size of the basketball court, what would be the size of our Milky Way galaxy?
6. The largest structures in the universe, like our "Local Supercluster," are about 500 million light-years across. Scaling these superclusters down to the dimensions of the basketball court, what would be the size of our Local Group?
7. The radius of the observable universe is 46.5 billion light-years. Compressing the observable universe to the size of the basketball court, what would be the size of the largest superclusters?

1.2 A key parameter in cosmology is the Hubble constant

$$H_0 \approx 70\,\mathrm{km\,s^{-1}Mpc^{-1}}\,.$$

In the following, you will use the measured value of the Hubble constant to estimate a few fundamental scales of our universe.

1. What is the Hubble time $t_{H_0} \equiv H_0^{-1}$ in years? This is a rough estimate of the age of the universe.
2. What is the Hubble distance $d_{H_0} \equiv cH_0^{-1}$ in meters? This is a rough estimate of the size of the observable universe.
3. The average density of the universe today is

$$\rho_0 = \frac{3H_0^2}{8\pi G}\,,$$

where $G = 6.67 \times 10^{-11}\,\mathrm{m^3\,kg^{-1}\,s^{-2}}$ is Newton's constant. What is ρ_0 in $\mathrm{g\,cm^{-3}}$? How does this compare to the density of water?
4. Let us assume that the universe is filled with only hydrogen atoms. What is then the total number of atoms in the universe? How does this compare to the number of hydrogen nuclei in your brain?

Hint: Assume that the brain is mostly water. Use $m_{\mathrm{H}} \approx 2 \times 10^{-24}$ g and $m_{\mathrm{H_2O}} \approx 3 \times 10^{-23}$ g.

5. The maximal energy scale probed by the Large Hadron Collider (LHC) is $E_{\rm max} \sim$ TeV. What length scale $\ell_{\rm min}$ does this correspond to? How does this compare to the size of the universe d_{H_0}?