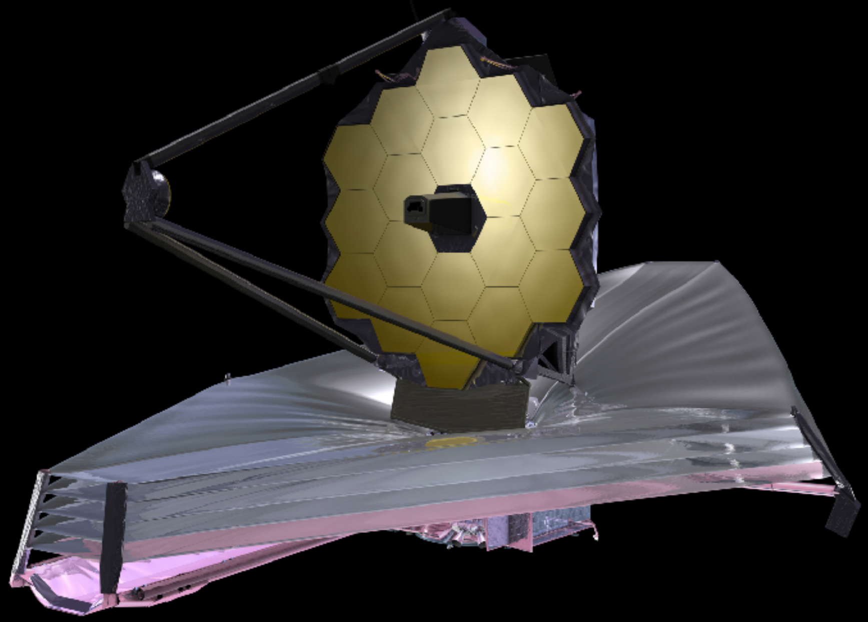

29

The Big Bang

Figure 29.1 Space Telescope of the Future. This drawing shows the James Webb Space Telescope, which is currently planned for launch in 2021. The silver sunshade shadows the primary mirror and science instruments. The primary mirror is 6.5 meters (21 feet) in diameter. Before and during launch, the mirror will be folded up. After the telescope is placed in its orbit, ground controllers will command it to unfold the mirror petals. To see distant galaxies whose light has been shifted to long wavelengths, the telescope will carry several instruments for taking infrared images and spectra. (credit: modification of work by NASA)

Chapter Outline

Thinking Ahead

In previous chapters, we explored the contents of the universe—planets, stars, and galaxies—and learned about how these objects change with time. But what about the universe as a whole? How old is it? What did it look like in the beginning? How has it changed since then? What will be its fate?

Cosmology is the study of the universe as a whole and is the subject of this chapter. The story of observational cosmology really begins in 1929 when Edwin Hubble published observations of redshifts and distances for a small sample of galaxies and showed the then-revolutionary result that we live in an expanding universe—one which in the past was denser, hotter, and smoother. From this early discovery, astronomers developed many predictions about the origin and evolution of the universe and then tested those predictions with observations. In this chapter, we will describe what we already know about the history of our dynamic universe and highlight some of the mysteries that remain.

29.1 The Age of the Universe

Learning Objectives

By the end of this section, you will be able to:

- Describe how we estimate the age of the universe
- Explain how changes in the rate of expansion over time affect estimates of the age of the universe
- Describe the evidence that dark energy exists and that the rate of expansion is currently accelerating
- Describe some independent evidence for the age of the universe that is consistent with the age estimate based on the rate of expansion

To explore the history of the universe, we will follow the same path that astronomers followed historically—beginning with studies of the nearby universe and then probing ever-more-distant objects and looking further back in time.

The realization that the universe changes with time came in the 1920s and 1930s when measurements of the redshifts of a large sample of galaxies became available. With hindsight, it is surprising that scientists were so shocked to discover that the universe is expanding. In fact, our theories of gravity demand that the universe must be either expanding or contracting. To show what we mean, let's begin with a universe of finite size—say a giant ball of a thousand galaxies. All these galaxies attract each other because of their gravity. If they were initially stationary, they would inevitably begin to move closer together and eventually collide. They could avoid this collapse only if for some reason they happened to be moving away from each other at high speeds. In just the same way, only if a rocket is launched at high enough speed can it avoid falling back to Earth.

The problem of what happens in an infinite universe is harder to solve, but Einstein (and others) used his theory of general relativity (which we described in Black Holes and Curved Spacetime) to show that even infinite universes cannot be static. Since astronomers at that time did not yet know the universe was expanding (and Einstein himself was philosophically unwilling to accept a universe in motion), he changed his equations by introducing an arbitrary new term (we might call it a fudge factor) called the **cosmological constant**. This constant represented a hypothetical force of repulsion that could balance gravitational attraction on the largest scales and permit galaxies to remain at fixed distances from one another. That way, the universe could remain still.

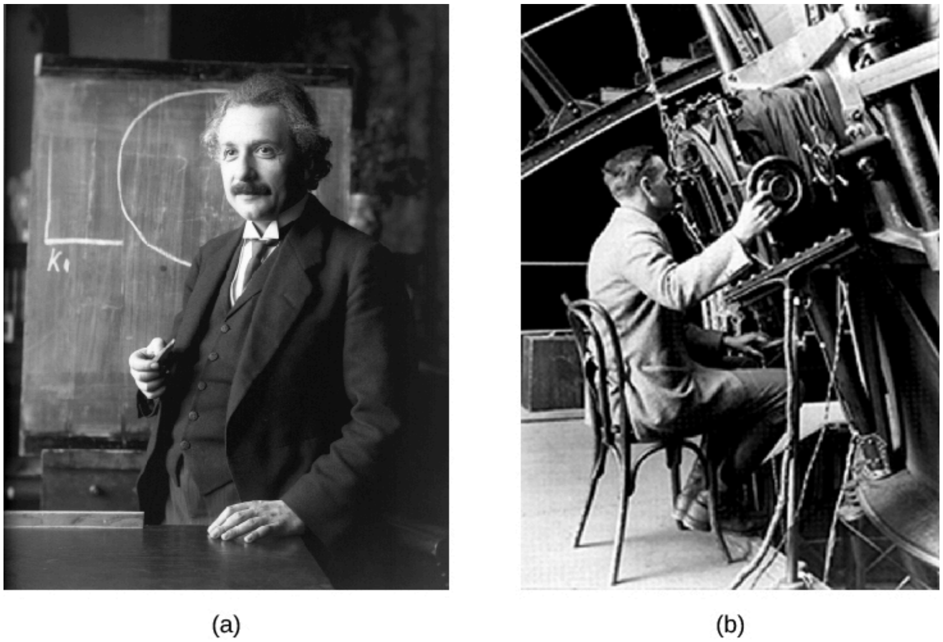
(a) (b)

Figure 29.2 Einstein and Hubble. (a) Albert Einstein is shown in a 1921 photograph. (b) Edwin Hubble at work in the Mt. Wilson Observatory.

About a decade later, Hubble, and his coworkers reported that the universe is expanding, so that no mysterious balancing force is needed. (We discussed this in the chapter on Galaxies.) Einstein is reported to have said that the introduction of the cosmological constant was “the biggest blunder of my life.” As we shall see later in this chapter, however, relatively recent observations indicate that the expansion is *accelerating*. Observations are now being carried out to determine whether this acceleration is consistent with a cosmological constant. In a way, it may turn out that Einstein was right after all.

LINK TO LEARNING

View this web exhibit (https://openstax.org/l/30exhcosmAIPCHP) on the history of our thinking about cosmology, with images and biographies, from the American Institute of Physics Center for the History of Physics.

The Hubble Time

If we had a movie of the expanding universe and ran the film *backward*, what would we see? The galaxies, instead of moving apart, would move *together* in our movie—getting closer and closer all the time. Eventually, we would find that all the matter we can see today was once concentrated in an infinitesimally small volume. Astronomers identify this time with the *beginning of the universe*. The explosion of that concentrated universe at the beginning of time is called the **Big Bang** (not a bad term, since you can’t have a bigger bang than one that creates the entire universe). But when did this bang occur?

We can make a reasonable estimate of the time since the universal expansion began. To see how astronomers do this, let’s begin with an analogy. Suppose your astronomy class decides to have a party (a kind of “Big Bang”) at someone’s home to celebrate the end of the semester. Unfortunately, everyone is celebrating with so much enthusiasm that the neighbors call the police, who arrive and send everyone away at the same moment. You get home at 2 a.m., still somewhat upset about the way the party ended, and realize you forgot to look at your watch to see what time the police got there. But you use a map to measure that the distance between the party and your house is 40 kilometers. And you also remember that you drove the whole trip at a steady speed of 80 kilometers/hour (since you were worried about the police cars following you). Therefore, the trip must have taken:

$$\text{time} = \frac{\text{distance}}{\text{velocity}} = \frac{40\text{ kilometers}}{80\text{ kilometers/hour}} = 0.5\text{ hours}$$

So the party must have broken up at 1:30 a.m.

No humans were around to look at their watches when the universe began, but we can use the same technique to estimate when the galaxies began moving away from each other. (Remember that, in reality, it is space that is expanding, not the galaxies that are moving through static space.) If we can measure how far apart the galaxies are now, and how fast they are moving, we can figure out how long a trip it’s been.

Let’s call the age of the universe measured in this way T_0. Let’s first do a simple case by assuming that the expansion has been at a constant rate ever since the expansion of the universe began. In this case, the time it has taken a galaxy to move a distance, d, away from the Milky Way (remember that at the beginning the galaxies were all together in a very tiny volume) is (as in our example)

$$T_0 = d/v$$

where v is the velocity of the galaxy. If we can measure the speed with which galaxies are moving away, and also the distances between them, we can establish how long ago the expansion began.

Making such measurements should sound very familiar. This is just what Hubble and many astronomers after him needed to do in order to establish the Hubble law and the Hubble constant. We learned in Galaxies that a galaxy’s distance and its velocity in the expanding universe are related by

$$V = H \times d$$

where H is the Hubble constant. Combining these two expressions gives us

$$T_0 = \frac{d}{v} = \frac{d}{(H \times d)} = \frac{1}{H}$$

We see, then, that the work of calculating this time was already done for us when astronomers measured the Hubble constant. The age of the universe estimated in this way turns out to be just the *reciprocal of the Hubble constant* (that is, $1/H$). This age estimate is sometimes called the Hubble time. For a Hubble constant of 20 kilometers/second per million light-years, the Hubble time is about 15 billion years. (By the way, the unit used by astronomers for the Hubble constant is kilometers/second per million parsecs. In these units, the Hubble constant is equal to about 70 kilometers/second per million parsecs, again with an uncertainty of about 5%.)

To make numbers easier to remember, we have done some rounding here. Estimates for the Hubble constant are actually closer to 21 or 22 kilometers/second per million light-years, which would make the age closer to 14 billion years. But there is still about a 5% uncertainty in the Hubble constant, which means the age of the universe estimated in this way is also uncertain by about 5%.

To put these uncertainties in perspective, however, you should know that 50 years ago, the uncertainty was a factor of 2. Remarkable progress toward pinning down the Hubble constant has been made in the last couple of decades.

The Role of Deceleration

The Hubble time is the right age for the universe only if the expansion rate has been constant throughout the time since the expansion of the universe began. Continuing with our end-of-the-semester-party analogy, this is equivalent to assuming that you traveled home from the party at a constant rate, when in fact this may not have been the case. At first, mad about having to leave, you may have driven fast, but then as you calmed down—and thought about police cars on the highway—you may have begun to slow down until you were driving at a more socially acceptable speed (such as 80 kilometers/hour). In this case, given that you were driving faster at the beginning, the trip home would have taken less than a half-hour.

In the same way, in calculating the Hubble time, we have assumed that the expansion rate has been constant throughout all of time. It turns out that this is not a good assumption. Earlier in their thinking about this, astronomers expected that the rate of expansion should be slowing down. We know that matter creates gravity, whereby all objects pull on all other objects. The mutual attraction between galaxies was expected to slow the expansion as time passed. This means that, if gravity were the only force acting (a big *if*, as we shall see in the next section), then the rate of expansion must have been faster in the past than it is today. In this case, we would say the universe has been *decelerating* since the beginning.

How much it has decelerated depends on the importance of gravity in slowing the expansion. If the universe were nearly empty, the role of gravity would be minor. Then the deceleration would be close to zero, and the universe would have been expanding at a constant rate. But in a universe with any significant density of matter, the pull of gravity means that the rate of expansion should be slower now than it used to be. If we use the current rate of expansion to estimate how long it took the galaxies to reach their current separations, we will overestimate the age of the universe—just as we may have overestimated the time it took for you to get home from the party.

A Universal Acceleration

Astronomers spent several decades looking for evidence that the expansion was decelerating, but they were not successful. What they needed were 1) larger telescopes so that they could measure the redshifts of more distant galaxies and 2) a very luminous *standard bulb* (or standard candle), that is, some astronomical object with known luminosity that produces an enormous amount of energy and can be observed at distances of a billion light-years or more.

Recall that we discussed standard bulbs in the chapter on Galaxies. If we compare how luminous a standard bulb is supposed to be and how dim it actually looks in our telescopes, the difference allows us to calculate its distance. The redshift of the galaxy such a bulb is in can tell us how fast it is moving in the universe. So we can measure its distance and motion independently.

These two requirements were finally met in the 1990s. Astronomers showed that supernovae of type Ia (see The Death of Stars), with some corrections based on the shapes of their light curves, are standard bulbs. This type of supernova occurs when a white dwarf accretes enough material from a companion star to exceed the Chandrasekhar limit and then collapses and explodes. At the time of maximum brightness, these dramatic supernovae can briefly outshine the galaxies that host them, and hence, they can be observed at very large distances. Large 8- to 10-meter telescopes can be used to obtain the spectra needed to measure the redshifts of the host galaxies (Figure 29.3).

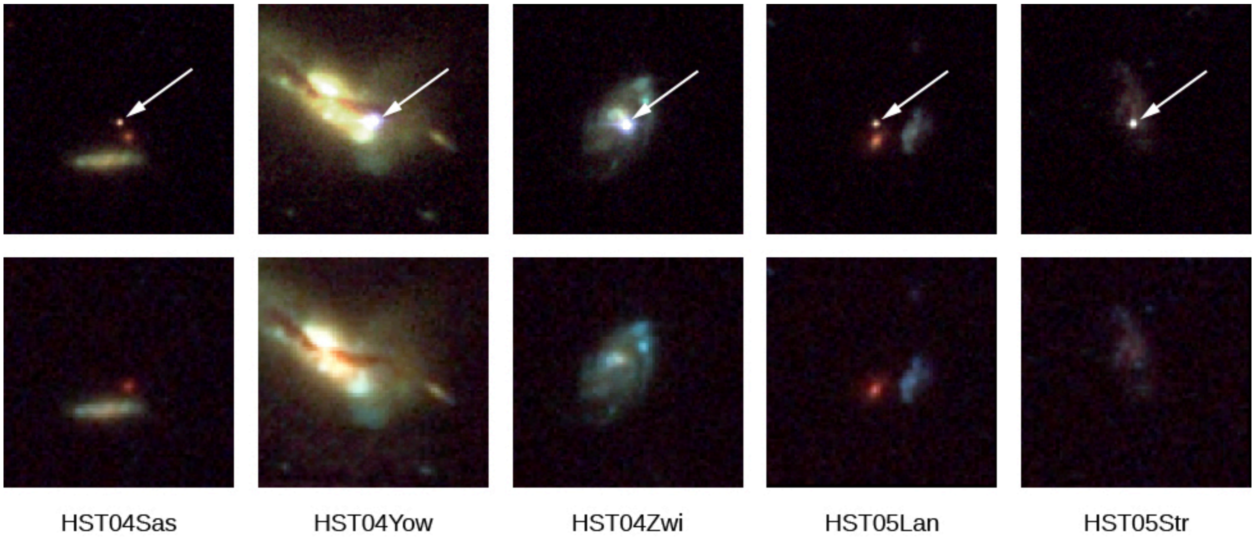

Figure 29.3 Five Supernovae and Their Host Galaxies. The top row shows each galaxy and its supernova (arrow). The bottom row shows the same galaxies either before or after the supernovae exploded. (credit: modification of work by NASA, ESA, and A. Riess (STScI))

The result of painstaking, careful study of these supernovae in a range of galaxies, carried out by two groups of researchers, was published in 1998. It was shocking—and so revolutionary that their discovery received the 2011 Nobel Prize in Physics. What the researchers found was that these type Ia supernovae in distant galaxies were fainter than expected from Hubble’s law, given the measured redshifts of their host galaxies. In other words, distances estimated from the supernovae used as standard bulbs disagreed with the distances measured from the redshifts.

If the universe were decelerating, we would expect the far-away supernovae to be *brighter* than expected. The slowing down would have kept them closer to us. Instead, they were *fainter*, which at first seemed to make no sense.

Before accepting this shocking development, astronomers first explored the possibility that the supernovae might not really be as useful as standard bulbs as they thought. Perhaps the supernovae appeared too faint because dust along our line of sight to them absorbed some of their light. Or perhaps the supernovae at large distances were for some reason intrinsically less luminous than nearby supernovae of type Ia.

A host of more detailed observations ruled out these possibilities. Scientists then had to consider the alternative that the distance estimated from the redshift was incorrect. Distances derived from redshifts assume that the Hubble constant has been truly constant for all time. We saw that one way it might not be constant is that the expansion is slowing down. But suppose neither assumption is right (steady speed or slowing down.)

Suppose, instead, that the universe is *accelerating*. If the universe is expanding faster now than it was billions of years ago, our motion away from the distant supernovae has sped up since the explosion occurred, sweeping us farther away from them. The light of the explosion has to travel a greater distance to reach us than if the expansion rate were constant. The farther the light travels, the fainter it appears. This conclusion would explain the supernova observations in a natural way, and this has now been substantiated by many additional observations over the last couple of decades. It really seems that *the expansion of the universe is accelerating*, a notion so unexpected that astronomers at first resisted considering it.

How can the expansion of the universe be speeding up? If you want to accelerate your car, you must supply energy by stepping on the gas. Similarly, energy must be supplied to accelerate the expansion of the universe. The discovery of the acceleration was shocking because scientists still have no idea what the source of the energy is. Scientists call whatever it is **dark energy**, which is a clear sign of how little we understand it.

Note that this new component of the universe is not the dark matter we talked about in earlier chapters. Dark energy is something else that we have also not yet detected in our laboratories on Earth.

What is dark energy? One possibility is that it is the cosmological constant, which is an energy associated with the vacuum of “empty” space itself. Quantum mechanics (the intriguing theory of how things behave at the atomic and subatomic levels) tells us that the source of this vacuum energy might be tiny elementary particles that flicker in and out of existence everywhere throughout the universe. Various attempts have been made to calculate how big the effects of this vacuum energy should be, but so far these attempts have been unsuccessful. In fact, the order of magnitude of theoretical estimates of the vacuum energy based on the quantum mechanics of matter and the value required to account for the acceleration of the expansion of the universe differ by an incredible factor of at least 10^{120} (that is a 1 followed by 120 zeros)! Various other theories have been suggested, but the bottom line is that, although there is compelling evidence that dark energy exists, we do not yet know the source of that energy.

Whatever the dark energy turns out to be, we should note that the discovery that the rate of expansion has not been constant since the beginning of the universe complicates the calculation of the age of the universe. Interestingly, the acceleration seems not to have started with the Big Bang. During the first several billion years after the Big Bang, when galaxies were close together, gravity was strong enough to slow the expansion. As galaxies moved farther apart, the effect of gravity weakened. Several billion years after the Big Bang, dark energy took over, and the expansion began to accelerate (Figure 29.4).

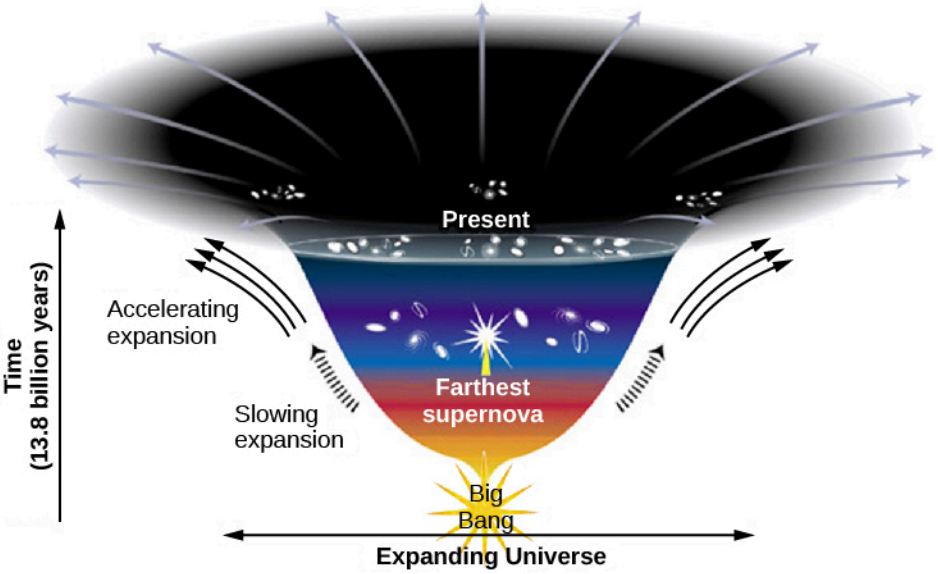

Figure 29.4 Changes in the Rate of Expansion of the Universe Since Its Beginning 13.8 Billion Years Ago. The more the diagram spreads out horizontally, the faster the change in the velocity of expansion. After a period of very rapid expansion at the beginning, which scientists call inflation and which we will discuss later in this chapter, the expansion began to decelerate. Galaxies were then close together, and their mutual gravitational attraction slowed the expansion. After a few billion years, when galaxies were farther apart, the influence of gravity began to weaken. Dark energy then took over and caused the expansion to accelerate. (credit: modification of work by Ann Feild (STScI))

Deceleration works to make the age of the universe estimated by the simple relation $T_0 = 1/H$ seem older than it really is, whereas acceleration works to make it seem younger. By happy coincidence, our best estimates of how much deceleration and acceleration occurred lead to an answer for the age very close to $T_0 = 1/H$. The best current estimate is that the universe is 13.8 billion years old with an uncertainty of only about 100 million years.

Throughout this chapter, we have referred to the Hubble *constant*. We now know that the Hubble constant does change with time. It is, however, constant everywhere in the universe at any given time. When we say the Hubble constant is about 70 kilometers/second/million parsecs, we mean that this is the value of the Hubble constant at the current time.

Comparing Ages

We now have one estimate for the age of the universe from its expansion. Is this estimate consistent with other observations? For example, are the oldest stars or other astronomical objects younger than 13.8 billion years? After all, the universe has to be at least as old as the oldest objects in it.

In our Galaxy and others, the oldest stars are found in the globular clusters (Figure 29.5), which can be dated using the models of stellar evolution described in the chapter Stars from Adolescence to Old Age.

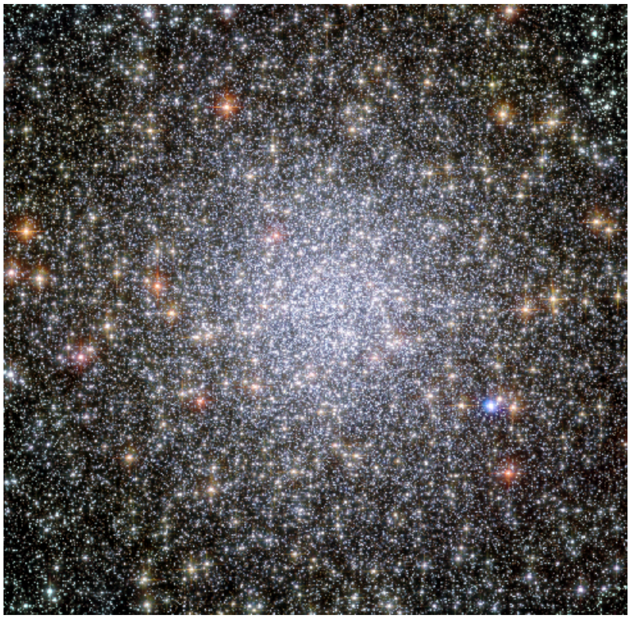

Figure 29.5 Globular Cluster 47 Tucanae. This NASA/ESA Hubble Space Telescope image shows a globular cluster known as 47 Tucanae, since it is in the constellation of Tucana (The Toucan) in the southern sky. The second-brightest globular cluster in the night sky, it includes hundreds of thousands of stars. Globular clusters are among the oldest objects in our Galaxy and can be used to estimate its age. (credit: NASA, ESA, and the Hubble Heritage (STScI/AURA)-ESA/Hubble Collaboration)

The accuracy of the age estimates of the globular clusters has improved markedly in recent years for two reasons. First, models of interiors of globular cluster stars have been improved, mainly through better information about how atoms absorb radiation as they make their way from the center of a star out into space. Second, observations from satellites have improved the accuracy of our measurements of the distances to these clusters. The conclusion is that the oldest stars formed about 12–13 billion years ago.

This age estimate has recently been confirmed by the study of the spectrum of uranium in the stars. The isotope uranium-238 is radioactive and decays (changes into another element) over time. (Uranium-238 gets its designation because it has 92 protons and 146 neutrons.) We know (from how stars and supernovae make elements) how much uranium-238 is generally made compared to other elements. Suppose we measure the amount of uranium relative to nonradioactive elements in a very old star and in our own Sun, and compare the abundances. With those pieces of information, we can estimate how much longer the uranium has been decaying in the very old star because we know from our own Sun how much uranium decays in 4.5 billion years.

The line of uranium is very weak and hard to make out even in the Sun, but it has now been measured in one extremely old star using the European Very Large Telescope (Figure 29.6). Comparing the abundance with that in the solar system, whose age we know, astronomers estimate the star is 12.5 billion years old, with an uncertainty of about 3 billion years. While the uncertainty is large, this work is important confirmation of the ages estimated by studies of the globular cluster stars. Note that the uranium age estimate is completely independent; it does not depend on either the measurement of distances or on models of the interiors of stars.

Figure 29.6 European Extremely Large Telescope, European Very Large Telescope, and the Colosseum. The European Extremely Large Telescope (E-ELT) is currently under construction in Chile. This image compares the size of the E-ELT (left) with the four 8-meter telescopes of the European Very Large Telescope (center) and with the Colosseum in Rome (right). The mirror of the E-ELT will be 39 meters in diameter. Astronomers are building a new generation of giant telescopes in order to observe very distant galaxies and understand what they were like when they were newly formed and the universe was young. (credit: modification of work by ESO)

As we shall see later in this chapter, the globular cluster stars probably did not form until the expansion of the universe had been underway for at least a few hundred million years. Accordingly, their ages are consistent with the 13.8 billion-year age estimated from the expansion rate.

29.2 A Model of the Universe

Learning Objectives

By the end of this section, you will be able to:

- Explain how the rate of expansion of the universe affects its evolution
- Describe four possibilities for the evolution of the universe
- Explain what is expanding when we say that the universe is expanding
- Define critical density and the evidence that matter alone in the universe is much smaller than the critical density
- Describe what the observations say about the likely long-term future of the universe

Let’s now use the results about the expansion of the universe to look at how these ideas might be applied to develop a model for the evolution of the universe as a whole. With this model, astronomers can make predictions about how the universe has evolved so far and what will happen to it in the future.

The Expanding Universe

Every model of the universe must include the expansion we observe. Another key element of the models is that the cosmological principle (which we discussed in The Evolution and Distribution of Galaxies) is valid: on the large scale, the universe at any given time is the same everywhere (homogeneous and isotropic). As a result, the expansion rate must be the same everywhere during any epoch of cosmic time. If so, we don’t need to think about the entire universe when we think about the expansion, we can just look at any sufficiently large portion of it. (Some models for dark energy would allow the expansion rate to be different in different directions, and scientists are designing experiments to test this idea. However, until such evidence is found, we will assume that the cosmological principle applies throughout the universe.)

In Galaxies, we hinted that when we think of the expansion of the universe, it is more correct to think of space itself stretching rather than of galaxies moving through static space. Nevertheless, we have since been discussing the redshifts of galaxies as if they resulted from the motion of the galaxies themselves.

Now, however, it is time to finally put such simplistic notions behind us and take a more sophisticated look at the cosmic expansion. Recall from our discussion of Einstein’s theory of general relativity (in the chapter on Black Holes and Curved Spacetime) that space—or, more precisely, spacetime—is not a mere backdrop to the action of the universe, as Newton thought. Rather, it is an active participant—affected by and in turn affecting the matter and energy in the universe.

Since the expansion of the universe is the stretching of all spacetime, all points in the universe are stretching together. Thus, the expansion began *everywhere at once*. Unfortunately for tourist agencies of the future, there is no location you can visit where the stretching of space began or where we can say that the Big Bang happened.

To describe just how space stretches, we say the cosmic expansion causes the universe to undergo a uniform change in *scale* over time. By scale we mean, for example, the distance between two clusters of galaxies. It is customary to represent the scale by the factor R; if R doubles, then the distance between the clusters has doubled. Since the universe is expanding at the same rate everywhere, the change in R tells us how much it has expanded (or contracted) at any given time. For a static universe, R would be constant as time passes. In an expanding universe, R increases with time.

If it is space that is stretching rather than galaxies moving through space, then why do the galaxies show redshifts in their spectra? When you were young and naïve—a few chapters ago—it was fine to discuss the redshifts of distant galaxies as if they resulted from their motion away from us. But now that you are an older and wiser student of cosmology, this view will simply not do.

A more accurate view of the redshifts of galaxies is that the light waves are stretched by the stretching of the space they travel through. Think about the light from a remote galaxy. As it moves away from its source, the light has to travel through space. If space is stretching during all the time the light is traveling, the light waves will be stretched as well. A redshift is a stretching of waves—the wavelength of each wave increases (Figure 29.7). Light from more distant galaxies travels for more time than light from closer ones. This means that the light has stretched more than light from closer ones and thus shows a greater redshift.

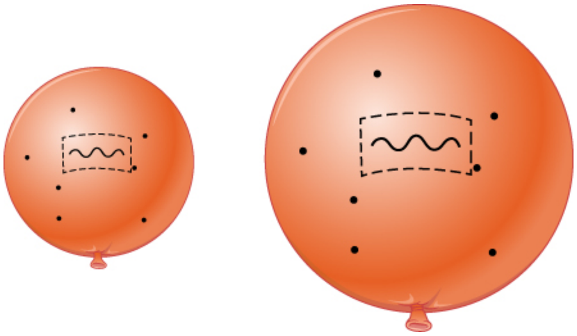

Figure 29.7 Expansion and Redshift. As an elastic surface expands, a wave on its surface stretches. For light waves, the increase in wavelength would be seen as a redshift.

Thus, what the measured redshift of light from an object is telling us is how much the universe has expanded since the light left the object. If the universe has expanded by a factor of 2, then the wavelength of the light (and all electromagnetic waves from the same source) will have doubled.

Models of the Expansion

Before astronomers knew about dark energy or had a good measurement of how much matter exists in the universe, they made speculative models about how the universe might evolve over time. The four possible scenarios are shown in Figure 29.8. In this diagram, time moves forward from the bottom upward, and the scale of space increases by the horizontal circles becoming wider.

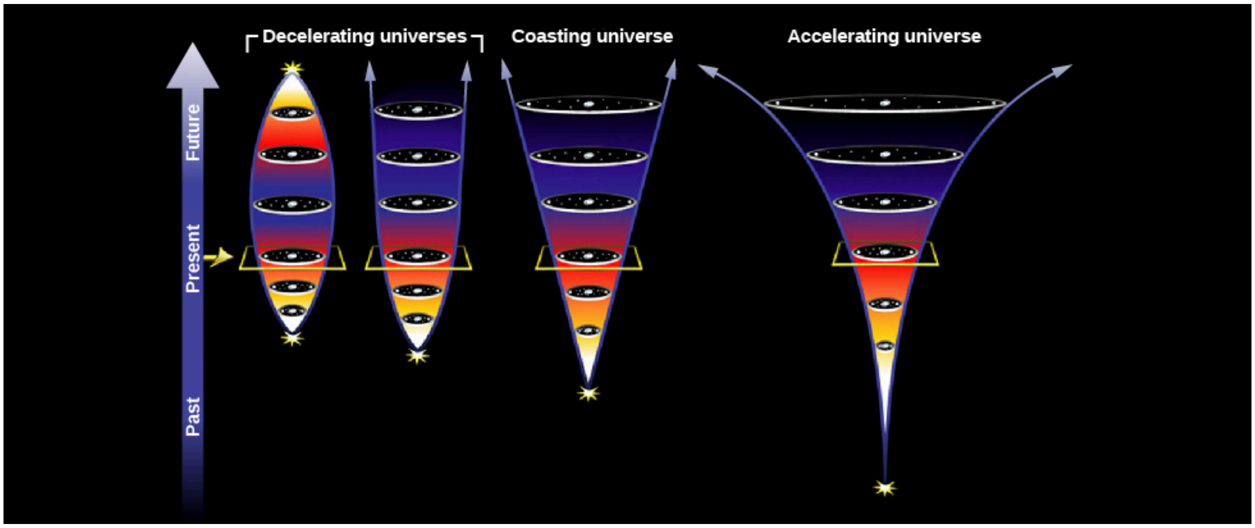

Figure 29.8 Four Possible Models of the Universe. The yellow square marks the present in all four cases, and for all four, the Hubble constant is equal to the same value at the present time. Time is measured in the vertical direction. The first two universes on the left are ones in which the rate of expansion slows over time. The one on the left will eventually slow, come to a stop and reverse, ending up in a "big crunch," while the one next to it will continue to expand forever, but ever-more slowly as time passes. The "coasting" universe is one that expands at a constant rate given by the Hubble constant throughout all of cosmic time. The accelerating universe on the right will continue to expand faster and faster forever. (credit: modification of work by NASA/ESA)

The simplest scenario of an expanding universe would be one in which R increases with time at a constant rate. But you already know that life is not so simple. The universe contains a great deal of mass and its gravity decelerates the expansion—by a large amount if the universe contains a lot of matter, or by a negligible amount if the universe is nearly empty. Then there is the observed acceleration, which astronomers blame on a kind of dark energy.

Let's first explore the range of possibilities with models for different amounts of mass in the universe and for different contributions by dark energy. In some models—as we shall see—the universe expands forever. In others, it stops expanding and starts to contract. After looking at the extreme possibilities, we will look at recent observations that allow us to choose the most likely scenario.

We should perhaps pause for a minute to note how remarkable it is that we can do this at all. Our understanding of the principles that underlie how the universe works on the large scale and our observations of how the objects in the universe change with time allow us to model the evolution of the entire cosmos these days. It is one of the loftiest achievements of the human mind.

What astronomers look at in practice, to determine the kind of universe we live in, is the *average density* of the universe. This is the mass of matter (including the equivalent mass of energy)[1] that would be contained in each unit of volume (say, 1 cubic centimeter) if all the stars, galaxies, and other objects were taken apart, atom by atom, and if all those particles, along with the light and other energy, were distributed throughout all of space with absolute uniformity. If the average density is low, there is less mass and less gravity, and the universe will not decelerate very much. It can therefore expand forever. Higher average density, on the other hand, means there is more mass and more gravity and that the stretching of space might slow down enough that the expansion will eventually stop. An extremely high density might even cause the universe to collapse again.

For a given rate of expansion, there is a **critical density**—the mass per unit volume that will be just enough to slow the expansion to zero at some time infinitely far in the future. If the actual density is higher than this critical density, then the expansion will ultimately reverse and the universe will begin to contract. If the actual density is lower, then the universe will expand forever.

These various possibilities are illustrated in Figure 29.9. In this graph, one of the most comprehensive in all of

---

[1] By equivalent mass we mean that which would result if the energy were turned into mass using Einstein's formula, $E = mc^2$.

science, we chart the development of the scale of space in the cosmos against the passage of time. Time increases to the right, and the scale of the universe, *R*, increases upward in the figure. Today, at the point marked "present" along the time axis, *R* is increasing in each model. We know that the galaxies are currently expanding away from each other, no matter which model is right. (The same situation holds for a baseball thrown high into the air. While it may eventually fall back down, near the beginning of the throw it moves upward most rapidly.)

The various lines moving across the graph correspond to different models of the universe. The straight dashed line corresponds to the empty universe with no deceleration; it intercepts the time axis at a time, T_0 (the Hubble time), in the past. This is not a realistic model but gives us a measure to compare other models to. The curves below the dashed line represent models with no dark energy and with varying amounts of deceleration, starting from the Big Bang at shorter times in the past. The curve above the dashed line shows what happens if the expansion is accelerating. Let's take a closer look at the future according to the different models.

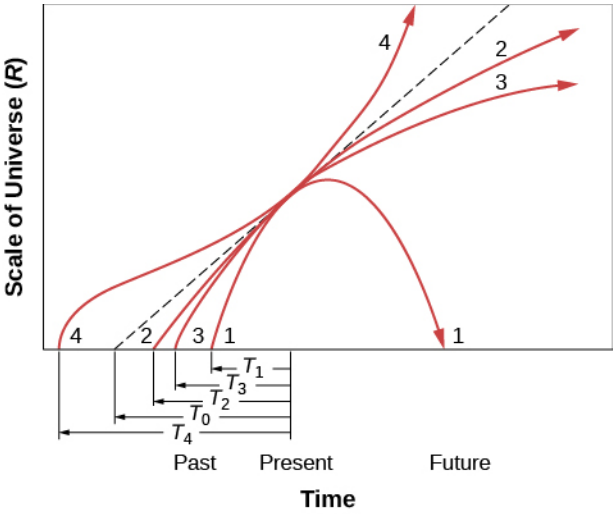

Figure 29.9 Models of the Universe. This graph plots *R*, the scale of the universe, against time for various cosmological models. Curve 1 represents a universe where the density is greater than the critical value; this model predicts that the universe will eventually collapse. Curve 2 represents a universe with a density lower than critical; the universe will continue to expand but at an ever-slower rate. Curve 3 is a critical-density universe; in this universe, the expansion will gradually slow to a stop infinitely far in the future. Curve 4 represents a universe that is accelerating because of the effects of dark energy. The dashed line is for an empty universe, one in which the expansion is not slowed by gravity or accelerated by dark energy. Time is very compressed on this graph.

Let's start with curve 1 in Figure 29.9. In this case, the actual density of the universe is higher than the critical density and there is no dark energy. This universe will stop expanding at some time in the future and begin contracting. This model is called a **closed universe** and corresponds to the universe on the left in Figure 29.8. Eventually, the scale drops to zero, which means that space will have shrunk to an infinitely small size. The noted physicist John Wheeler called this the "big crunch," because matter, energy, space, and time would all be crushed out of existence. Note that the "big crunch" is the opposite of the Big Bang—it is an *implosion*. The universe is not expanding but rather collapsing in upon itself.

Some scientists speculated that another Big Bang might follow the crunch, giving rise to a new expansion phase, and then another contraction—perhaps oscillating between successive Big Bangs and big crunches indefinitely in the past and future. Such speculation was sometimes referred to as the oscillating theory of the universe. The challenge for theorists was how to describe the transition from collapse (when space and time themselves disappear into the big crunch) to expansion. With the discovery of dark energy, however, it does not appear that the universe will experience a big crunch, so we can put worrying about it on the back burner.

If the density of the universe is less than the critical density (curve 2 in Figure 29.9 and the universe second from the left in Figure 29.8), gravity is never important enough to stop the expansion, and so the universe expands forever. Such a universe is infinite and this model is called an **open universe**. Time and space begin with the Big Bang, but they have no end; the universe simply continues expanding, always a bit more slowly as time goes on. Groups of galaxies eventually get so far apart that it would be difficult for observers in any of them to see the others. (See the feature box on What Might the Universe Be Like in the Distant Future? for more about the distant future in the closed and open universe models.)

At the critical density (curve 3), the universe can just barely expand forever. The critical-density universe has an age of exactly two-thirds T_0, where T_0 is the age of the empty universe. Universes that will someday begin to contract have ages less than two-thirds T_0.

In an empty universe (the dashed line Figure 29.9 and the coasting universe in Figure 29.8), neither gravity nor dark energy is important enough to affect the expansion rate, which is therefore constant throughout all time.

In a universe with dark energy, the rate of the expansion will increase with time, and the expansion will continue at an ever-faster rate. Curve 4 in Figure 29.9, which represents this universe, has a complicated shape. In the beginning, when the matter is all very close together, the rate of expansion is most influenced by gravity. Dark energy appears to act only over large scales and thus becomes more important as the universe grows larger and the matter begins to thin out. In this model, at first the universe slows down, but as space stretches, the acceleration plays a greater role and the expansion speeds up.

The Cosmic Tug of War

We might summarize our discussion so far by saying that a “tug of war” is going on in the universe between the forces that push everything apart and the gravitational attraction of matter, which pulls everything together. If we can determine who will win this tug of war, we will learn the ultimate fate of the universe.

The first thing we need to know is the density of the universe. Is it greater than, less than, or equal to the critical density? The critical density today depends on the value of the expansion rate today, H_0. If the Hubble constant is around 20 kilometers/second per million light-years, the critical density is about 10^{-26} kg/m^3. Let’s see how this value compares with the actual density of the universe.

EXAMPLE 29.1

Critical Density of the Universe

As we discussed, the critical density is that combination of matter and energy that brings the universe coasting to a stop at time infinity. Einstein’s equations lead to the following expression for the critical density (ρ_{crit}):

$$\rho_{\text{crit}} = \frac{3H^2}{8\pi G}$$

where H is the Hubble constant and G is the universal constant of gravity (6.67×10^{-11} Nm^2/kg^2).

Solution

Let’s substitute our values and see what we get. Take an $H = 22$ km/s per million light-years. We need to convert both km and light-years into meters for consistency. A million light-years = $10^6 \times 9.5 \times 10^{15}$ m = 9.5×10^{21} m. And 22 km/s = 2.2×10^4 m/s. That makes $H = 2.3 \times 10^{-18}$ /s and $H^2 = 5.36 \times 10^{-36}$ /s^2. So,

$$\rho_{\text{crit}} = \frac{3 \times 5.36 \times 10^{-36}}{8 \times 3.14 \times 6.67 \times 10^{-11}} = 9.6 \times 10^{-27}\ \text{kg/m}^3$$

which we can round off to the 10^{-26} kg/m^3. (To make the units work out, you have to know that N, the unit of force, is the same as kg × m/s^2.)

Now we can compare densities we measure in the universe to this critical value. Note that density is mass per unit volume, but energy has an equivalent mass of $m = E/c^2$ (from Einstein’s equation $E = mc^2$).

Check Your Learning

a. A single grain of dust has a mass of about 1.1 × 10^{-13} kg. If the average mass-energy density of space is equal to the critical density on average, how much space would be required to produce a total mass-energy equal to a dust grain?

b. If the Hubble constant were twice what it actually is, how much would the critical density be?

Answer:

a. In this case, the average mass-energy in a volume V of space is $E = \rho_{\text{crit}}V$. Thus, for space with critical density, we require that

$$V = \frac{E_{\text{grain}}}{\rho_{\text{crit}}} = \frac{1.1 \times 10^{-13}\ \text{kg}}{9.6 \times 10^{-26}\ \text{kg/m}^3} = 1.15 \times 10^{12}\ \text{m}^3 = (10{,}500\ \text{m})^3 \cong (10.5\ \text{km})^3$$

Thus, the sides of a cube of space with mass-energy density averaging that of the critical density would need to be slightly greater than 10 km to contain the total energy equal to a single grain of dust!

b. Since the critical density goes as the square of the Hubble constant, by doubling the Hubble parameter, the critical density would increase by a factor a four. So if the Hubble constant was 44 km/s per million light-years instead of 22 km/s per million light-years, the critical density would be $\rho_{\text{crit}} = 4 \times 9.6 \times 10^{-27}\ \text{kg/m}^3 = 3.8 \times 10^{-26}\ \text{kg/m}^3$.

We can start our survey of how dense the cosmos is by ignoring the dark energy and just estimating the density of all matter in the universe, including ordinary matter and dark matter. Here is where the cosmological principle really comes in handy. Since the universe is the same all over (at least on large scales), we only need to measure how much matter exists in a (large) representative sample of it. This is similar to the way a representative survey of a few thousand people can tell us whom the millions of residents of the US prefer for president.

There are several methods by which we can try to determine the average density of matter in space. One way is to count all the galaxies out to a given distance and use estimates of their masses, including dark matter, to calculate the average density. Such estimates indicate a density of about 1 to 2 × 10^{-27} kg/m^3 (10 to 20% of critical), which by itself is too small to stop the expansion.

A lot of the dark matter lies outside the boundaries of galaxies, so this inventory is not yet complete. But even if we add an estimate of the dark matter outside galaxies, our total won’t rise beyond about 30% of the critical density. We’ll pin these numbers down more precisely later in this chapter, where we will also include the effects of dark energy.

In any case, even if we ignore dark energy, the evidence is that the universe will continue to expand forever. The discovery of dark energy that is causing the rate of expansion to speed up only strengthens this conclusion. Things definitely do not look good for fans of the closed universe (big crunch) model.

MAKING CONNECTIONS

What Might the Universe Be Like in the Distant Future?

Some say the world will end in fire,
Some say in ice.
From what I’ve tasted of desire
I hold with those who favor fire.

—From the poem “Fire and Ice” by Robert Frost (1923)

Given the destructive power of impacting asteroids, expanding red giants, and nearby supernovae, our species may not be around in the remote future. Nevertheless, you might enjoy speculating about what it would be like to live in a much, much older universe.

The observed acceleration makes it likely that we will have continued expansion into the indefinite future. If the universe expands forever (R increases without limit), the clusters of galaxies will spread ever farther apart with time. As eons pass, the universe will get thinner, colder, and darker.

Within each galaxy, stars will continue to go through their lives, eventually becoming white dwarfs, neutron stars, and black holes. Low-mass stars might take a long time to finish their evolution, but in this model, we would literally have all the time in the world. Ultimately, even the white dwarfs will cool down to be black dwarfs, any neutron stars that reveal themselves as pulsars will slowly stop spinning, and black holes with accretion disks will one day complete their “meals.” The remains of stars will all be dark and difficult to observe.

This means that the light that now reveals galaxies to us will eventually go out. Even if a small pocket of raw material were left in one unsung corner of a galaxy, ready to be turned into a fresh cluster of stars, we will only have to wait until the time that their evolution is also complete. And time is one thing this model of the universe has plenty of. There will surely come a time when all the stars are out, galaxies are as dark as space, and no source of heat remains to help living things survive. Then the lifeless galaxies will just continue to move apart in their lightless realm.

If this view of the future seems discouraging (from a human perspective), keep in mind that we fundamentally do not understand why the expansion rate is currently accelerating. Thus, our speculations about the future are just that: speculations. You might take heart in the knowledge that science is always a progress report. The most advanced ideas about the universe from a hundred years ago now strike us as rather primitive. It may well be that our best models of today will in a hundred or a thousand years also seem rather simplistic and that there are other factors determining the ultimate fate of the universe of which we are still completely unaware.

Ages of Distant Galaxies

In the chapter on Galaxies, we discussed how we can use Hubble’s law to measure the distance to a galaxy. But that simple method only works with galaxies that are not too far away. Once we get to large distances, we are looking so far into the past that we must take into account changes in the rate of the expansion of the universe. Since we cannot measure these changes directly, we must assume one of the models of the universe to be able to convert large redshifts into distances.

This is why astronomers squirm when reporters and students ask them exactly how far away some newly discovered distant quasar or galaxy is. We really can’t give an answer without first explaining the model of the universe we are assuming in calculating it (by which time a reporter or student is long gone or asleep). Specifically, we must use a model that includes the change in the expansion rate with time. The key ingredients of the model are the amounts of matter, including dark matter, and the equivalent mass (according to $E = mc^2$) of the dark energy along with the Hubble constant.

Elsewhere in this book, we have estimated the mass density of ordinary matter plus dark matter as roughly 0.3 times the critical density, and the mass equivalent of dark energy as roughly 0.7 times the critical density. We will refer to these values as the “standard model of the universe.” The latest (slightly improved) estimates for these values and the evidence for them will be given later in this chapter. Calculations also require the current value of the Hubble constant. For Table 29.1, we have adopted a Hubble constant of 67.3 kilometers/second/million parsecs (rather than rounding it to 70 kilometers/second/million parsecs), which is consistent with the

13.8 billion-year age of the universe estimated by the latest observations.

Once we assume a model, we can use it to calculate the age of the universe at the time an object emitted the light we see. As an example, Table 29.1 lists the times that light was emitted by objects at different redshifts as fractions of the current age of the universe. The times are given for two very different models so you can get a feeling for the fact that the calculated ages are fairly similar. The first model assumes that the universe has a critical density of matter and no dark energy. The second model is the standard model described in the preceding paragraph. The first column in the table is the redshift, which is given by the equation $z = \Delta\lambda/\lambda_0$ and is a measure of how much the wavelength of light has been stretched by the expansion of the universe on its long journey to us.

Ages of the Universe at Different Redshifts

Redshift	Percent of Current Age of Universe When the Light Was Emitted (mass = critical density)	Percent of Current Age of Universe When the Light Was Emitted (mass = 0.3 critical density; dark energy = 0.7 critical density)
0	100 (now)	100 (now)
0.5	54	63
1.0	35	43
2.0	19	24
3.0	13	16
4.0	9	11
5.0	7	9
8.0	4	5
11.9	2.1	2.7
Infinite	0	0

Table 29.1

Notice that as we find objects with higher and higher redshifts, we are looking back to smaller and smaller fractions of the age of the universe. The highest observed redshifts as this book is being written are close to 12 (Figure 29.10). As Table 29.1 shows, we are seeing these galaxies as they were when the universe was only about 3% as old as it is now. They were already formed only about 700 million years after the Big Bang.

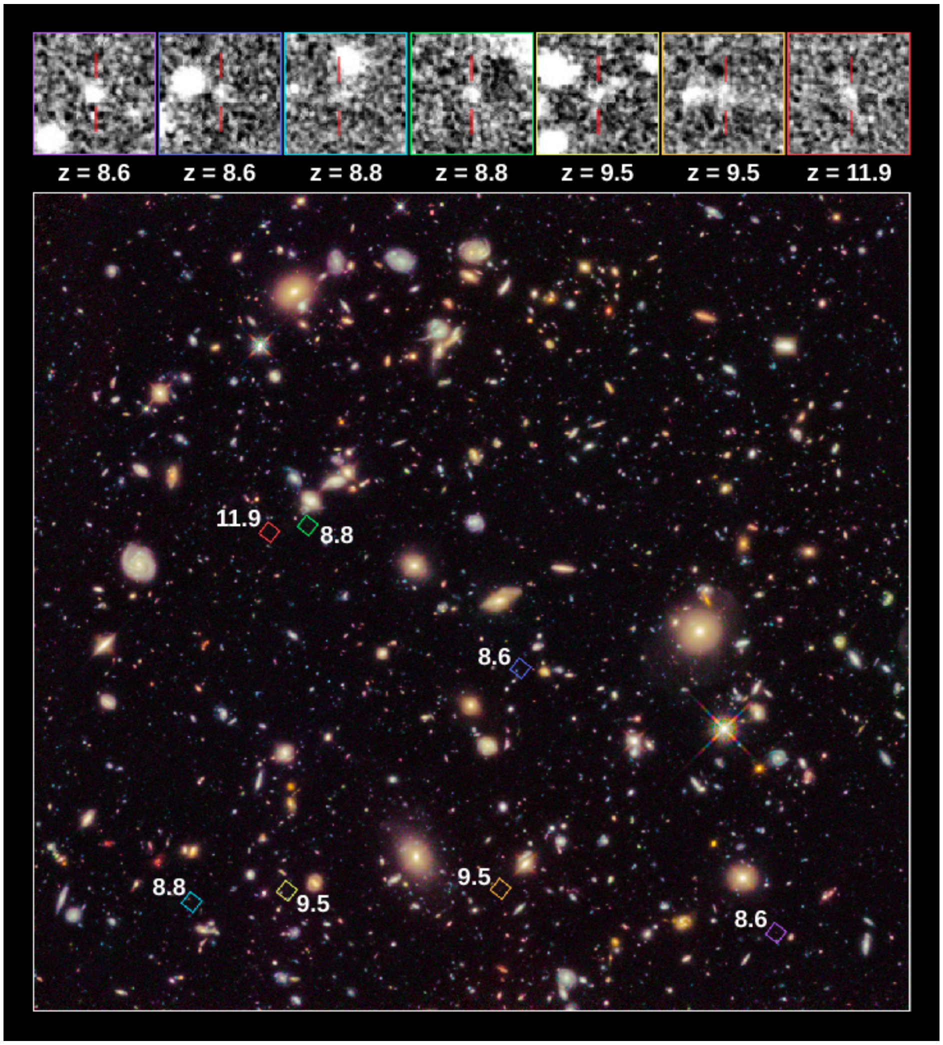

Figure 29.10 Hubble Ultra-Deep Field. This image, called the Hubble Ultra Deep Field, shows faint galaxies, seen very far away and therefore very far back in time. The colored squares in the main image outline the locations of the galaxies. Enlarged views of each galaxy are shown in the black-and-white images. The red lines mark each galaxy's location. The "redshift" of each galaxy is indicated below each box, denoted by the symbol "z." The redshift measures how much a galaxy's ultraviolet and visible light has been stretched to infrared wavelengths by the universe's expansion. The larger the redshift, the more distant the galaxy, and therefore the further astronomers are seeing back in time. One of the seven galaxies may be a distance breaker, observed at a redshift of 11.9. If this redshift is confirmed by additional measurements, the galaxy is seen as it appeared only 380 million years after the Big Bang, when the universe was less than 3% of its present age. (credit: modification of work by NASA, ESA, R. Ellis (Caltech), and the UDF 2012 Team)

29.3 The Beginning of the Universe

Learning Objectives

By the end of this section, you will be able to:

- Describe what the universe was like during the first few minutes after it began to expand
- Explain how the first new elements were formed during the first few minutes after the Big Bang
- Describe how the contents of the universe change as the temperature of the universe decreases

The best evidence we have today indicates that the first galaxies did not begin to form until a few hundred million years after the Big Bang. What were things like before there were galaxies and space had not yet stretched very significantly? Amazingly, scientists have been able to calculate in some detail what was happening in the universe in the first few minutes after the Big Bang.

The History of the Idea

It is one thing to say the universe had a beginning (as the equations of general relativity imply) and quite another to describe that beginning. The Belgian priest and cosmologist Georges Lemaître was probably the first to propose a specific model for the Big Bang itself (Figure 29.11). He envisioned all the matter of the universe starting in one great bulk he called the *primeval atom*, which then broke into tremendous numbers of pieces. Each of these pieces continued to fragment further until they became the present atoms of the universe, created in a vast nuclear fission. In a popular account of his theory, Lemaître wrote, “The evolution of the world could be compared to a display of fireworks just ended—some few red wisps, ashes, and smoke. Standing on a well-cooled cinder, we see the slow fading of the suns and we try to recall the vanished brilliance of the origin of the worlds.”

Figure 29.11 Abbé Georges Lemaître (1894–1966). This Belgian cosmologist studied theology at Mechelen and mathematics and physics at the University of Leuven. It was there that he began to explore the expansion of the universe and postulated its explosive beginning. He actually predicted Hubble’s law 2 years before its verification, and he was the first to consider seriously the physical processes by which the universe began.

LINK TO LEARNING

View a short video (https://openstax.org/l/30Lemaitrevid) about the work of Lemaître, considered by some to be the father of the Big Bang theory.

Physicists today know much more about nuclear physics than was known in the 1920s, and they have shown that the primeval fission model cannot be correct. Yet Lemaître’s vision was in some respects quite prophetic. We still believe that everything was together at the beginning; it was just not in the form of matter we now know. Basic physical principles tell us that when the universe was much denser, it was also much hotter, and that it cools as it expands, much as gas cools when sprayed from an aerosol can.

By the 1940s, scientists knew that fusion of hydrogen into helium was the source of the Sun’s energy. Fusion requires high temperatures, and the early universe must have been hot. Based on these ideas, American physicist George Gamow (Figure 29.12) suggested a universe with a different kind of beginning that involved nuclear **fusion** instead of fission. Ralph Alpher worked out the details for his PhD thesis, and the results were published in 1948. (Gamow, who had a quirky sense of humor, decided at the last minute to add the name of physicist Hans Bethe to their paper, so that the coauthors on this paper about the beginning of things would be Alpher, Bethe, and Gamow, a pun on the first three letters of the Greek alphabet: alpha, beta, and gamma.)

Gamow’s universe started with fundamental particles that built up the heavy elements by fusion in the Big Bang.

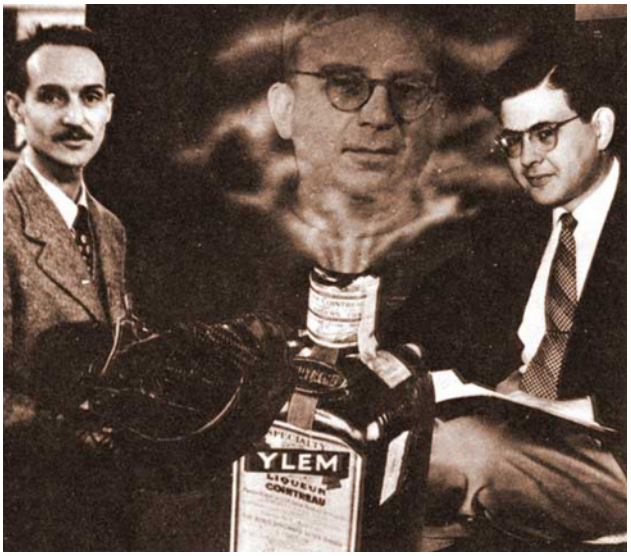

Figure 29.12 George Gamow and Collaborators. This composite image shows George Gamow emerging like a genie from a bottle of ylem, a Greek term for the original substance from which the world formed. Gamow revived the term to describe the material of the hot Big Bang. Flanking him are Robert Herman (left) and Ralph Alpher (right), with whom he collaborated in working out the physics of the Big Bang. (The modern composer Karlheinz Stockhausen was inspired by Gamow’s ideas to write a piece of music called *Ylem*, in which the players actually move away from the stage as they perform, simulating the expansion of the universe.)

Gamow’s ideas were close to our modern view, except we now know that the early universe remained hot enough for fusion for only a short while. Thus, only the three lightest elements—hydrogen, helium, and a small amount of lithium—were formed in appreciable abundances at the beginning. The heavier elements formed later in stars. Since the 1940s, many astronomers and physicists have worked on a detailed theory of what happened in the early stages of the universe.

The First Few Minutes

Let’s start with the first few minutes following the Big Bang. Three basic ideas hold the key to tracing the changes that occurred during the time just after the universe began. The first, as we have already mentioned, is that the universe cools as it expands. Figure 29.13 shows how the temperature changes with the passage of time. Note that a huge span of time, from a tiny fraction of a second to billions of years, is summarized in this diagram. In the first fraction of a second, the universe was unimaginably hot. By the time 0.01 second had elapsed, the temperature had dropped to 100 billion (10^{11}) K. After about 3 minutes, it had fallen to about 1 billion (10^{9}) K, still some 70 times hotter than the interior of the Sun. After a few hundred thousand years, the temperature was down to a mere 3000 K, and the universe has continued to cool since that time.

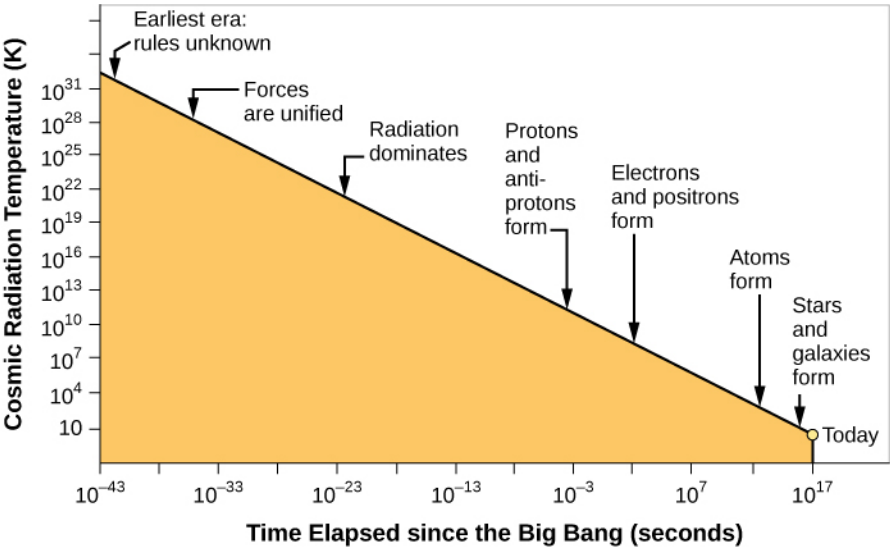

Figure 29.13 Temperature of the Universe. This graph shows how the temperature of the universe varies with time as predicted by the standard model of the Big Bang. Note that both the temperature (vertical axis) and the time in seconds (horizontal axis) change over vast scales on this compressed diagram.

All of these temperatures but the last are derived from theoretical calculations since (obviously) no one was there to measure them directly. As we shall see in the next section, however, we have actually detected the feeble glow of radiation emitted at a time when the universe was a few hundred thousand years old. We can measure the characteristics of that radiation to learn what things were like long ago. Indeed, the fact that we have found this ancient glow is one of the strongest arguments in favor of the Big Bang model.

The second step in understanding the evolution of the universe is to realize that at very early times, it was so hot that it contained mostly radiation (and not the matter that we see today). The photons that filled the universe could collide and produce material particles; that is, under the conditions just after the Big Bang, energy could turn into matter (and matter could turn into energy). We can calculate how much mass is produced from a given amount of energy by using Einstein's formula $E = mc^2$ (see the chapter on The Sun: A Nuclear Powerhouse).

The idea that energy could turn into matter in the universe at large is a new one for many students, since it is not part of our everyday experience. That's because, when we compare the universe today to what it was like right after the Big Bang, we live in cold, hard times. The photons in the universe today typically have far-less energy than the amount required to make new matter. In the discussion on the source of the Sun's energy in The Sun: A Nuclear Powerhouse, we briefly mentioned that when subatomic particles of matter and *antimatter* collide, they turn into pure energy. But the reverse, energy turning into matter and antimatter, is equally possible. This process has been observed in particle accelerators around the world. If we have enough energy, under the right circumstances, new particles of matter (and antimatter) are indeed created —and the conditions were right during the first few minutes after the expansion of the universe began.

Our third key point is that the hotter the universe was, the more energetic were the photons available to make matter and antimatter (see Figure 29.13). To take a specific example, at a temperature of 6 billion (6×10^9) K, the collision of two typical photons can create an electron and its antimatter counterpart, a positron. If the temperature exceeds 10^{14} K, much more massive protons and antiprotons can be created.

The Evolution of the Early Universe

Keeping these three ideas in mind, we can trace the evolution of the universe from the time it was about 0.01 second old and had a temperature of about 100 billion K. Why not begin at the very beginning? There are as yet no theories that allow us penetrate to a time before about 10^{-43} second (this number is a decimal point followed by 42 zeros and then a one). It is so small that we cannot relate it to anything in our everyday experience. When the universe was that young, its density was so high that the theory of general relativity is not adequate to describe it, and even the concept of time breaks down.

Scientists, by the way, have been somewhat more successful in describing the universe when it was older than 10^{-43} second but still less than about 0.01 second old. We will take a look at some of these ideas later in this chapter, but for now, we want to start with somewhat more familiar situations.

By the time the universe was 0.01 second old, it consisted of a soup of matter and radiation; the matter included protons and neutrons, leftovers from an even younger and hotter universe. Each particle collided rapidly with other particles. The temperature was no longer high enough to allow colliding photons to produce neutrons or protons, but it was sufficient for the production of electrons and positrons (Figure 29.14). There was probably also a sea of exotic subatomic particles that would later play a role as dark matter. All the particles jiggled about on their own; it was still much too hot for protons and neutrons to combine to form the nuclei of atoms.

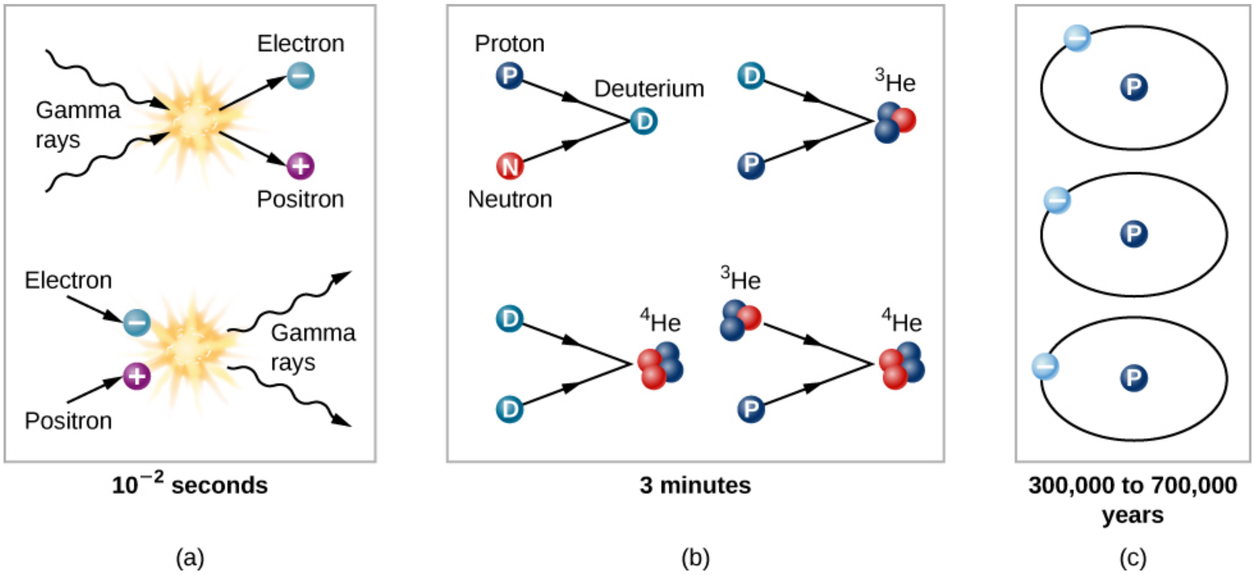

Figure 29.14 Particle Interactions in the Early Universe. (a) In the first fractions of a second, when the universe was very hot, energy was converted into particles and antiparticles. The reverse reaction also happened: a particle and antiparticle could collide and produce energy. (b) As the temperature of the universe decreased, the energy of typical photons became too low to create matter. Instead, existing particles fused to create such nuclei as deuterium and helium. (c) Later, it became cool enough for electrons to settle down with nuclei and make neutral atoms. Most of the universe was still hydrogen.

Think of the universe at this time as a seething cauldron, with photons colliding and interchanging energy, and sometimes being destroyed to create a pair of particles. The particles also collided with one another. Frequently, a matter particle and an antimatter particle met and turned each other into a burst of gamma-ray radiation.

Among the particles created in the early phases of the universe was the ghostly neutrino (see The Sun: A Nuclear Powerhouse), which today interacts only very rarely with ordinary matter. In the crowded conditions of the very early universe, however, neutrinos ran into so many electrons and positrons that they experienced frequent interactions despite their “antisocial” natures.

By the time the universe was a little more than 1 second old, the density had dropped to the point where neutrinos no longer interacted with matter but simply traveled freely through space. In fact, these neutrinos should now be all around us. Since they have been traveling through space unimpeded (and hence unchanged) since the universe was 1 second old, measurements of their properties would offer one of the best tests of the Big Bang model. Unfortunately, the very characteristic that makes them so useful—the fact that they interact so weakly with matter that they have survived unaltered for all but the first second of time—also renders them unable to be measured, at least with present techniques. Perhaps someday someone will devise a way to capture these elusive messengers from the past.

Atomic Nuclei Form

When the universe was about 3 minutes old and its temperature was down to about 900 million K, protons and neutrons could combine. At higher temperatures, these atomic nuclei had immediately been blasted apart by interactions with high-energy photons and thus could not survive. But at the temperatures and densities reached between 3 and 4 minutes after the beginning, **deuterium** (a proton and neutron) lasted long enough that collisions could convert some of it into helium, (Figure 29.14). In essence, the entire universe was acting the way centers of stars do today—fusing new elements from simpler components. In addition, a little bit of element 3, **lithium**, could also form.

This burst of cosmic fusion was only a brief interlude, however. By 4 minutes after the Big Bang, more helium was having trouble forming. The universe was still expanding and cooling down. After the formation of helium and some lithium, the temperature had dropped so low that the fusion of helium nuclei into still-heavier elements could not occur. No elements beyond lithium could form in the first few minutes. That 4-minute period was the end of the time when the entire universe was a fusion factory. In the cool universe we know today, the fusion of new elements is limited to the centers of stars and the explosions of supernovae.

Still, the fact that the Big Bang model allows the creation of a good deal of helium is the answer to a long-standing mystery in astronomy. Put simply, there is just too much helium in the universe to be explained by what happens inside stars. All the generations of stars that have produced helium since the Big Bang cannot account for the quantity of helium we observe. Furthermore, even the oldest stars and the most distant galaxies show significant amounts of helium. These observations find a natural explanation in the synthesis of helium by the Big Bang itself during the first few minutes of time. We estimate that *10 times more helium* was manufactured in the first 4 minutes of the universe than in all the generations of stars during the succeeding 10 to 15 billion years.

Learning from Deuterium

We can learn many things from the way the early universe made atomic nuclei. It turns out that all of the deuterium (a hydrogen nucleus with a neutron in it) in the universe was formed during the first 4 minutes. In stars, any region hot enough to fuse two protons to form a deuterium nucleus is also hot enough to change it further—either by destroying it through a collision with an energetic photon or by converting it into helium through nuclear reactions.

The amount of deuterium that can be produced in the first 4 minutes of creation depends on the density of the universe at the time deuterium was formed. If the density were relatively high, nearly all the deuterium would have been converted into helium through interactions with protons, just as it is in stars. If the density were relatively low, then the universe would have expanded and thinned out rapidly enough that some deuterium would have survived. The amount of deuterium we see today thus gives us a clue to the density of the universe when it was about 4 minutes old. Theoretical models can relate the density then to the density now; thus, measurements of the abundance of deuterium today can give us an estimate of the current density of the universe.

The measurements of deuterium indicate that the present-day density of ordinary matter—protons and neutrons—is about 5×10^{-28} kg/m^3. Deuterium can only provide an estimate of the density of ordinary matter because the abundance of deuterium is determined by the particles that interact to form it, namely protons and neutrons alone. From the abundance of deuterium, we know that not enough protons and neutrons are present, by a factor of about 20, to produce a critical-density universe.

We do know, however, that there are dark matter particles that add to the overall matter density of the universe, which is then higher than what is calculated for ordinary matter alone. Because dark matter particles do not affect the production of deuterium, measurement of the deuterium abundance cannot tell us how much dark matter exists. Dark matter is made of some exotic kind of particle, not yet detected in any earthbound laboratory. It is definitely not made of protons and neutrons like the readers of this book.

29.4 The Cosmic Microwave Background

Learning Objectives

By the end of this section, you will be able to:

- Explain why we can observe the afterglow of the hot, early universe
- Discuss the properties of this afterglow as we see it today, including its average temperature and the size of its temperature fluctuations
- Describe open, flat, and curved universes and explain which type of universe is supported by observations
- Summarize our current knowledge of the basic properties of the universe including its age and contents

The description of the first few minutes of the universe is based on theoretical calculations. It is crucial, however, that a scientific theory should be testable. What predictions does it make? And do observations show those predictions to be accurate? One success of the theory of the first few minutes of the universe is the correct prediction of the amount of helium in the universe.

Another prediction is that a significant milestone in the history of the universe occurred about 380,000 years after the Big Bang. Scientists have directly observed what the universe was like at this early stage, and these observations offer some of the strongest support for the Big Bang theory. To find out what this milestone was, let’s look at what theory tells us about what happened during the first few hundred thousand years after the Big Bang.

The fusion of helium and lithium was completed when the universe was about 4 minutes old. The universe then continued to resemble the interior of a star in some ways for a few hundred thousand years more. It remained hot and opaque, with radiation being scattered from one particle to another. It was still too hot for electrons to “settle down” and become associated with a particular nucleus; such free electrons are especially effective at scattering photons, thus ensuring that no radiation ever got very far in the early universe without having its path changed. In a way, the universe was like an enormous crowd right after a popular concert; if you get separated from a friend, even if he is wearing a flashing button, it is impossible to see through the dense crowd to spot him. Only after the crowd clears is there a path for the light from his button to reach you.

The Universe Becomes Transparent

Not until a few hundred thousand years after the Big Bang, when the temperature had dropped to about 3000 K and the density of atomic nuclei to about 1000 per cubic centimeter, did the electrons and nuclei manage to combine to form stable atoms of hydrogen and helium (Figure 29.14). With no free electrons to scatter photons, the universe became transparent for the first time in cosmic history. From this point on, matter and radiation interacted much less frequently; we say that they *decoupled* from each other and evolved separately. Suddenly, electromagnetic radiation could really travel, and it has been traveling through the universe ever since.

Discovery of the Cosmic Background Radiation

If the model of the universe described in the previous section is correct, then—as we look far outward in the universe and thus far back in time—the first “afterglow” of the hot, early universe should still be detectable. Observations of it would be very strong evidence that our theoretical calculations about how the universe evolved are correct. As we shall see, we have indeed detected the radiation emitted at this **photon decoupling time**, when radiation began to stream freely through the universe without interacting with matter (Figure 29.15).

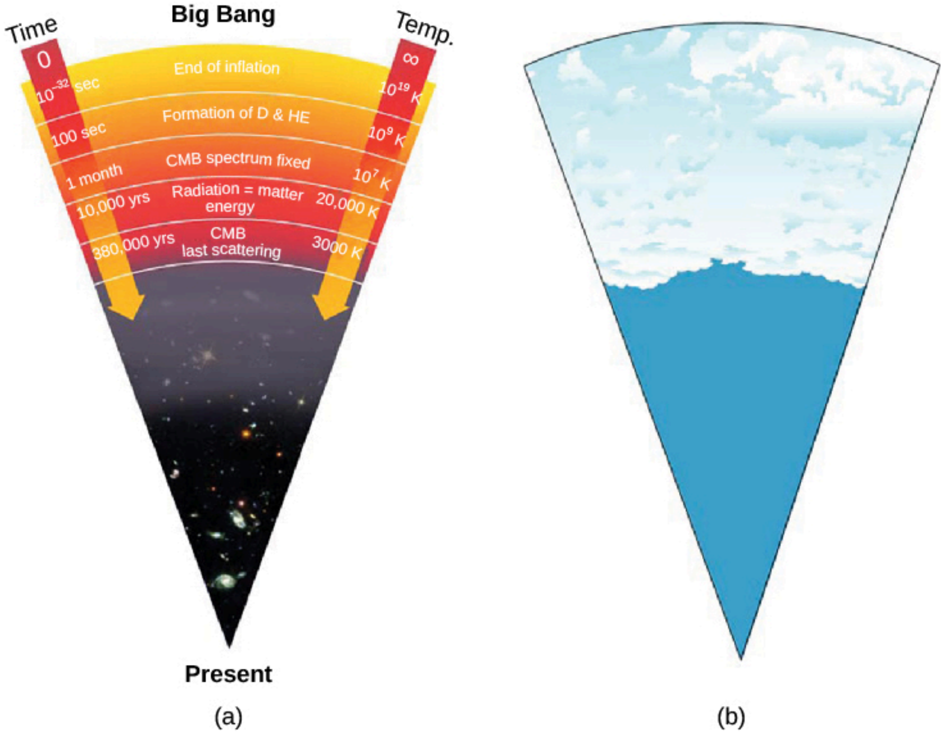

Figure 29.15 Cosmic Microwave Background and Clouds Compared. (a) Early in the universe, photons (electromagnetic energy) were scattering off the crowded, hot, charged particles and could not get very far without colliding with another particle. But after electrons and photons settled into neutral atoms, there was far less scattering, and photons could travel over vast distances. The universe became transparent. As we look out in space and back in time, we can't see back beyond this time. (b) This is similar to what happens when we see clouds in Earth's atmosphere. Water droplets in a cloud scatter light very efficiently, but clear air lets light travel over long distances. So as we look up into the atmosphere, our vision is blocked by the cloud layers and we can't see beyond them. (credit: modification of work by NASA)

The detection of this afterglow was initially an accident. In the late 1940s, Ralph Alpher and Robert Herman, working with George Gamow, realized that just before the universe became transparent, it must have been radiating like a blackbody at a temperature of about 3000 K—the temperature at which hydrogen atoms could begin to form. If we could have seen that radiation just after neutral atoms formed, it would have resembled radiation from a reddish star. It was as if a giant fireball filled the whole universe.

But that was nearly 14 billion years ago, and, in the meantime, the scale of the universe has increased a thousand fold. This expansion has increased the wavelength of the radiation by a factor of 1000 (see Figure 29.7). According to Wien's law, which relates wavelength and temperature, the expansion has correspondingly lowered the temperature by a factor of 1000 (see the chapter on Radiation and Spectra). The cosmic background behaves like a blackbody and should therefore have a spectrum that obeys Wien's Law.

Alpher and Herman predicted that the glow from the fireball should now be at radio wavelengths and should resemble the radiation from a blackbody at a temperature only a few degrees above absolute zero. Since the fireball was everywhere throughout the universe, the radiation left over from it should also be everywhere. If our eyes were sensitive to radio wavelengths, the whole sky would appear to glow very faintly. However, our eyes can't see at these wavelengths, and at the time Alpher and Herman made their prediction, there were no instruments that could detect the glow. Over the years, their prediction was forgotten.

In the mid-1960s, in Holmdel, New Jersey, Arno Penzias and Robert Wilson of AT&T's Bell Laboratories had built a delicate microwave antenna (Figure 29.16) to measure astronomical sources, including supernova remnants like Cassiopeia A (see the chapter on The Death of Stars). They were plagued with some unexpected background noise, just like faint static on a radio, which they could not get rid of. The puzzling thing about this radiation was that it seemed to be coming from all directions at once. This is very unusual in astronomy: after all, most radiation has a specific direction where it is strongest—the direction of the Sun, or a supernova remnant, or the disk of the Milky Way, for example.

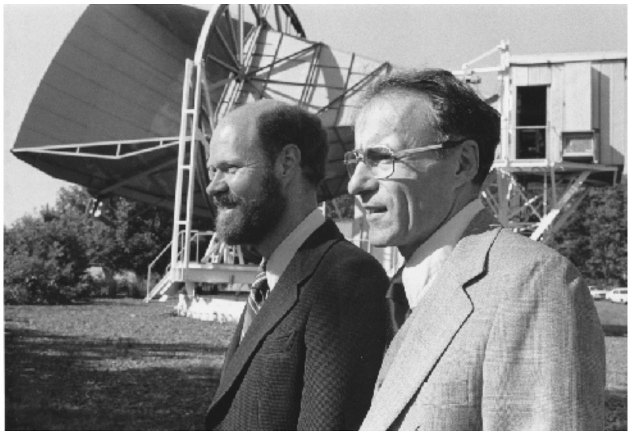

Figure 29.16 Robert Wilson (left) and Arno Penzias (right). These two scientists are standing in front of the horn-shaped antenna with which they discovered the cosmic background radiation. The photo was taken in 1978, just after they received the Nobel Prize in physics.

Penzias and Wilson at first thought that any radiation appearing to come from all directions must originate from inside their telescope, so they took everything apart to look for the source of the noise. They even found that some pigeons had roosted inside the big horn-shaped antenna and had left (as Penzias delicately put it) “a layer of white, sticky, dielectric substance coating the inside of the antenna.” However, nothing the scientists did could reduce the background radiation to zero, and they reluctantly came to accept that it must be real, and it must be coming from space.

Penzias and Wilson were not cosmologists, but as they began to discuss their puzzling discovery with other scientists, they were quickly put in touch with a group of astronomers and physicists at Princeton University (a short drive away). These astronomers had—as it happened—been redoing the calculations of Alpher and Herman from the 1940s and also realized that the radiation from the decoupling time should be detectable as a faint afterglow of radio waves. The different calculations of what the observed temperature would be for this **cosmic microwave background (CMB)**[2] were uncertain, but all predicted less than 40 K.

Penzias and Wilson found the distribution of intensity at different radio wavelengths to correspond to a temperature of 3.5 K. This is very cold—closer to absolute zero than most other astronomical measurements—and a testament to how much space (and the waves within it) has stretched. Their measurements have been repeated with better instruments, which give us a reading of 2.73 K. So Penzias and Wilson came very close. Rounding this value, scientists often refer to “the 3-degree microwave background.”

Many other experiments on Earth and in space soon confirmed the discovery by Penzias and Wilson: The radiation was indeed coming from all directions (it was isotropic) and matched the predictions of the Big Bang theory with remarkable precision. Penzias and Wilson had inadvertently observed the glow from the primeval fireball. They received the Nobel Prize for their work in 1978. And just before his death in 1966, Lemaître learned that his “vanished brilliance” had been discovered and confirmed.

LINK TO LEARNING

You may enjoy watching *Three Degrees*, a 26-minute video (https://openstax.org/l/30threedegvid) from Bell Labs about Penzias and Wilson’s discovery of the cosmic background radiation (with interesting historical footage).

---

[2] Recall that microwaves are in the radio region of the electromagnetic spectrum.

Properties of the Cosmic Microwave Background

One issue that worried astronomers is that Penzias and Wilson were measuring the background radiation filling space through Earth’s atmosphere. What if that atmosphere is a source of radio waves or somehow affected their measurements? It would be better to measure something this important from space.

The first accurate measurements of the CMB were made with a satellite orbiting Earth. Named the Cosmic Background Explorer (COBE), it was launched by NASA in November 1989. The data it received quickly showed that the CMB closely matches that expected from a blackbody with a temperature of 2.73 K (Figure 29.17). This is exactly the result expected if the CMB was indeed redshifted radiation emitted by a hot gas that filled all of space shortly after the universe began.

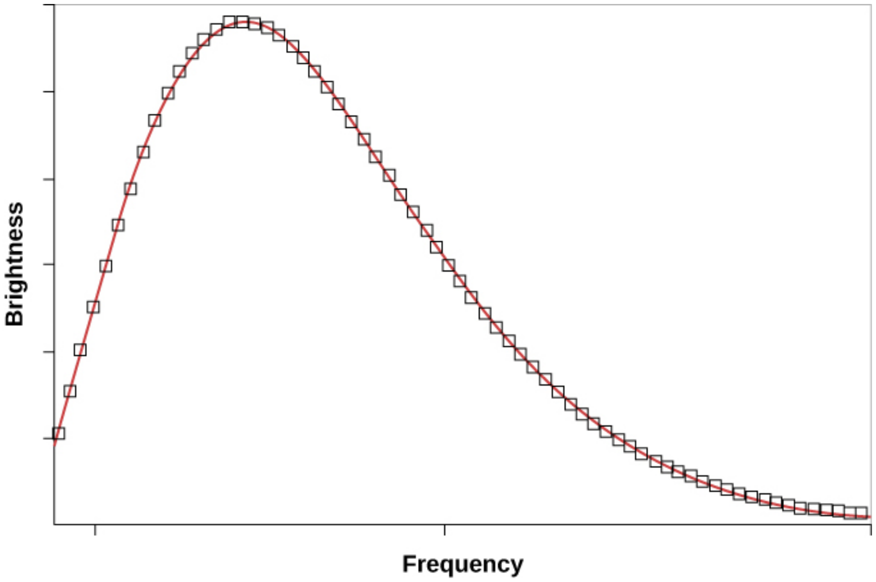

Figure 29.17 Cosmic Background Radiation. The solid line shows how the intensity of radiation should change with wavelength for a blackbody with a temperature of 2.73 K. The boxes show the intensity of the cosmic background radiation as measured at various wavelengths by COBE’s instruments. The fit is perfect. When this graph was first shown at a meeting of astronomers, they gave it a standing ovation.

The first important conclusion from measurements of the CMB, therefore, is that the universe we have today has indeed evolved from a hot, uniform state. This observation also provides direct support for the general idea that we live in an evolving universe, since the universe is cooler today than it was in the beginning.

Small Differences in the CMB

It was known even before the launch of COBE that the CMB is extremely *isotropic*. In fact, its uniformity in every direction is one of the best confirmations of the cosmological principle— that the universe is homogenous and isotropic.

According to our theories, however, the temperature could not have been *perfectly* uniform when the CMB was emitted. After all, the CMB is radiation that was scattered from the particles in the universe at the time of decoupling. If the radiation were completely smooth, then all those particles must have been distributed through space absolutely evenly. Yet it is those particles that have become all the galaxies and stars (and astronomy students) that now inhabit the cosmos. Had the particles been completely smoothly distributed, they could not have formed all the large-scale structures now present in the universe—the clusters and superclusters of galaxies discussed in the last few chapters.

The early universe must have had tiny density fluctuations from which such structures could evolve. Regions of higher-than-average density would have attracted additional matter and eventually grown into the galaxies and clusters that we see today. It turns out that these denser regions would appear to us to be colder spots, that is, they would have lower-than-average temperatures.

The reason that temperature and density are related can be explained this way. At the time of decoupling, photons in a slightly denser portion of space had to expend some of their energy to escape the gravitational force exerted by the surrounding gas. In losing energy, the photons became slightly colder than the overall average temperature at the time of decoupling. Vice versa, photons that were located in a slightly less dense portion of space lost less energy upon leaving it than other photons, thus appearing slightly hotter than average. Therefore, if the seeds of present-day galaxies existed at the time that the CMB was emitted, we should see some slight variations in the CMB temperature as we look in different directions in the sky.

Scientists working with the data from the COBE satellite did indeed detect very subtle temperature differences—about 1 part in 100,000—in the CMB. The regions of lower-than-average temperature come in a variety of sizes, but even the smallest of the colder areas detected by COBE is far too large to be the precursor of an individual galaxy, or even a supercluster of galaxies. This is because the COBE instrument had “blurry vision” (poor resolution) and could only measure large patches of the sky. We needed instruments with “sharper vision.”

The most detailed measurements of the CMB have been obtained by two satellites launched more recently than COBE. The results from the first of these satellites, the Wilkinson Microwave Anisotropy Probe (WMAP) spacecraft, were published in 2003. In 2015, measurements from the Planck satellite extended the WMAP measurements to even-higher spatial resolution and lower noise (Figure 29.18).

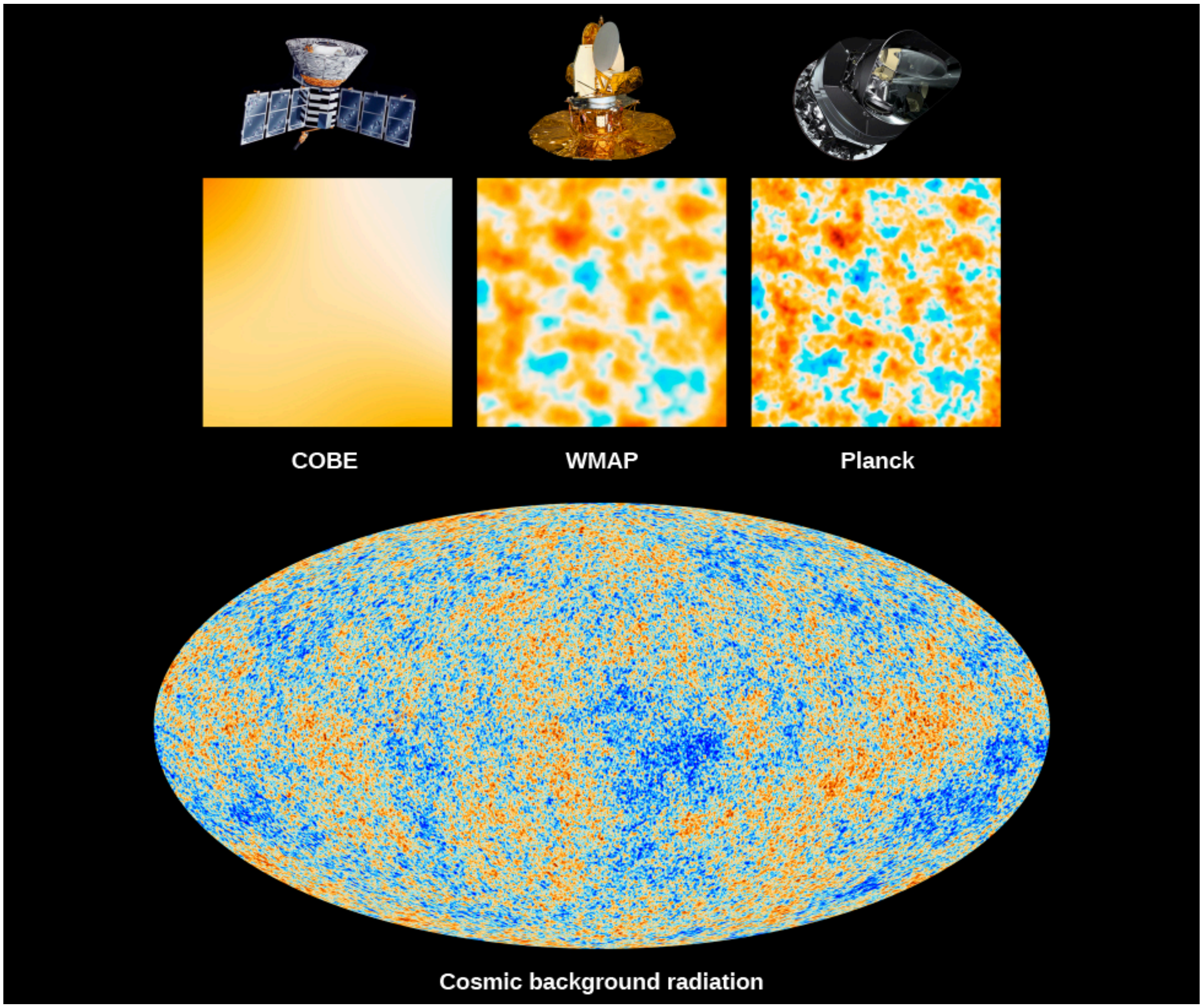

Figure 29.18 CMB Observations. This comparison shows how much detail can be seen in the observations of three satellites used to measure the CMB. The CMB is a snapshot of the oldest light in our universe, imprinted on the sky when the universe was just about 380,000 years old. The first spacecraft, launched in 1989, is NASA’s Cosmic Background Explorer, or COBE. WMAP was launched in

2001, and Planck was launched in 2009. The three panels show 10-square-degree patches of all-sky maps. This cosmic background radiation image (bottom) is an all-sky map of the CMB as observed by the Planck mission. The colors in the map represent different temperatures: red for warmer and blue for cooler. These tiny temperature fluctuations correspond to regions of slightly different densities, representing the seeds of all future structures: the stars, galaxies, and galaxy clusters of today. (credit top: modification of work by NASA/JPL-Caltech/ESA; credit bottom: modification of work by ESA and the Planck Collaboration)

Theoretical calculations show that the sizes of the hot and cold spots in the CMB depend on the geometry of the universe and hence on its total density. (It's not at all obvious that it should do so, and it takes some pretty fancy calculations—way beyond the level of our text—to make the connection, but having such a dependence is very useful.) The total density we are discussing here includes both the amount of mass in the universe and the mass equivalent of the dark energy. That is, we must add together mass and energy: ordinary matter, dark matter, and the dark energy that is speeding up the expansion.

To see why this works, remember (from the chapter on Black Holes and Curved Spacetime) that with his theory of general relativity, Einstein showed that matter can curve space and that the amount of curvature depends on the amount of matter present. Therefore, the total amount of matter in the universe (including dark matter and the equivalent matter contribution by dark energy), determines the overall geometry of space. Just like the geometry of space around a black hole has a curvature to it, so the entire universe may have a curvature. Let's take a look at the possibilities (Figure 29.19).

If the density of matter is higher than the critical density, the universe will eventually collapse. In such a closed universe, two initially parallel rays of light will eventually meet. This kind of geometry is referred to as spherical geometry. If the density of matter is less than critical, the universe will expand forever. Two initially parallel rays of light will diverge, and this is referred to as hyperbolic geometry. In a critical-density universe, two parallel light rays never meet, and the expansion comes to a halt only at some time infinitely far in the future. We refer to this as a **flat universe**, and the kind of Euclidean geometry you learned in high school applies in this type of universe.

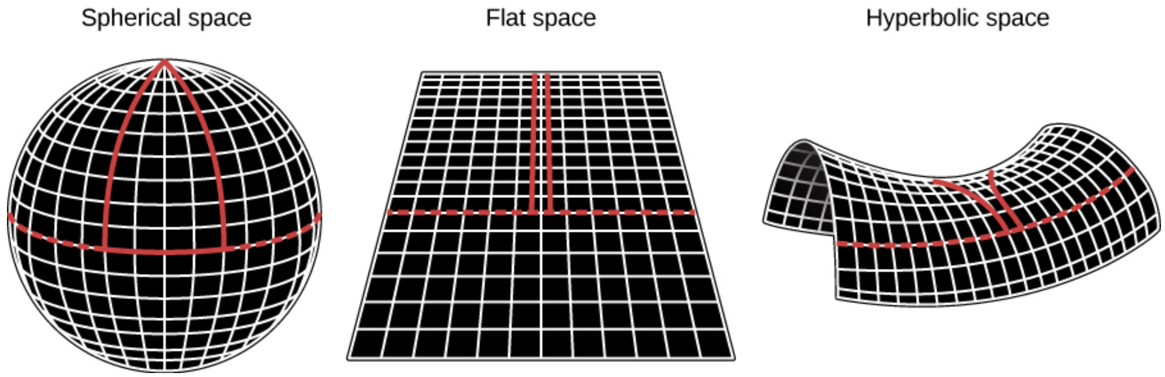

Figure 29.19 Picturing Space Curvature for the Entire Universe. The density of matter and energy determines the overall geometry of space. If the density of the universe is greater than the critical density, then the universe will ultimately collapse and space is said to be *closed* like the surface of a sphere. If the density exactly equals the critical density, then space is *flat* like a sheet of paper; the universe will expand forever, with the rate of expansion coming to a halt infinitely far in the future. If the density is less than critical, then the expansion will continue forever and space is said to be *open* and negatively curved like the surface of a saddle (where more space than you expect opens up as you move farther away). Note that the red lines in each diagram show what happens in each kind of space—they are initially parallel but follow different paths depending on the curvature of space. Remember that these drawings are trying to show how space for the entire universe is "warped"—this can't be seen locally in the small amount of space that we humans occupy.

If the density of the universe is equal to the critical density, then the hot and cold spots in the CMB should typically be about a degree in size. If the density is greater than critical, then the typical sizes will be larger than one degree. If the universe has a density less than critical, then the structures will appear smaller. In Figure 29.20, you can see the differences easily. WMAP and Planck observations of the CMB confirmed earlier experiments that we do indeed live in a flat, critical-density universe.

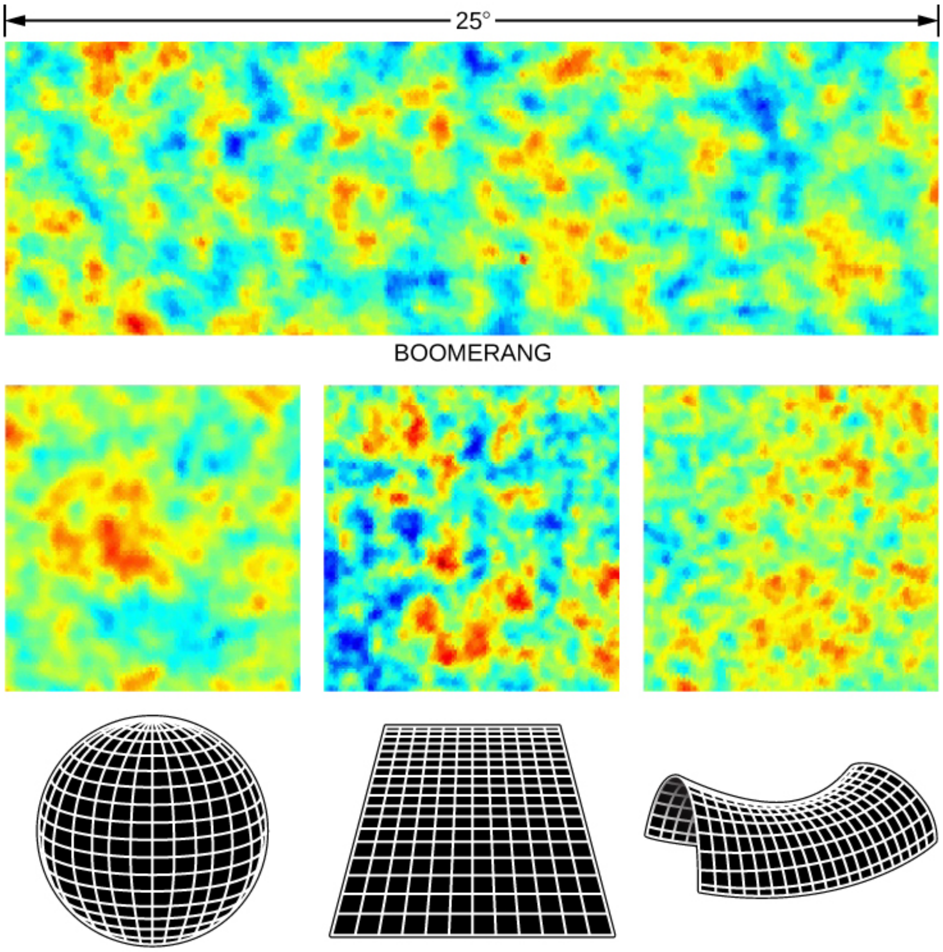

Figure 29.20 Comparison of CMB Observations with Possible Models of the Universe. Cosmological simulations predict that if our universe has critical density, then the CMB images will be dominated by hot and cold spots of around one degree in size (bottom center). If, on the other hand, the density is higher than critical (and the universe will ultimately collapse), then the images' hot and cold spots will appear larger than one degree (bottom left). If the density of the universe is less than critical (and the expansion will continue forever), then the structures will appear smaller (bottom right). As the measurements show, the universe is at critical density. The measurements shown were made by a balloon-borne instrument called BOOMERanG (Balloon Observations of Millimetric Extragalactic Radiation and Geophysics), which was flown in Antarctica. Subsequent satellite observations by WMAP and Planck confirm the BOOMERanG result. (credit: modification of work by NASA)

Key numbers from an analysis of the Planck data give us the best values currently available for some of the basic properties of the universe:

- Age of universe: 13.799 ± 0.038 billion years (Note: That means we know the age of the universe to within 38 million years. Amazing!)
- Hubble constant: 67.31 ± 0.96 kilometers/second/million parsecs (in the units we've been using, 20.65 kilometers/second/million light-years)
- Fraction of universe's content that is "dark energy": 68.5% ± 1.3%
- Fraction of the universe's content that is matter: 31.5% ± 1.3%

Note that this value for the Hubble constant is slightly smaller than the value of 70 kilometers/second/million parsecs that we have adopted in this book. In fact, the value derived from measurements of redshifts is 73 kilometers/second/million parsecs. So precise is modern cosmology these days that scientists are working hard to resolve this discrepancy. The fact that the difference between these two independent measurements is so small is actually a remarkable achievement. Only a few decades ago, astronomers were arguing about whether the Hubble constant was around 50 kilometers/second/million parsecs or 100 kilometers/second/million parsecs.

Analysis of Planck data also shows that ordinary matter (mainly protons and neutrons) makes up 4.9% of the total density. Dark matter plus normal matter add up to 31.5% of the total density. Dark energy contributes the remaining 68.5%. The age of the universe at decoupling—that is, when the CMB was emitted—was 380,000 years.

Perhaps the most surprising result from the high-precision measurements by WMAP and the even higher-precision measurements from Planck is that there were no surprises. The model of cosmology with ordinary matter at about 5%, dark matter at about 25%, and dark energy about 70% has survived since the late 1990s when cosmologists were forced in that direction by the supernovae data. In other words, the very strange universe that we have been describing, with only about 5% of its contents being made up of the kinds of matter we are familiar with here on Earth, really seems to be the universe we live in.

After the CMB was emitted, the universe continued to expand and cool off. By 400 to 500 million years after the Big Bang, the very first stars and galaxies had already formed. Deep in the interiors of stars, matter was reheated, nuclear reactions were ignited, and the more gradual synthesis of the heavier elements that we have discussed throughout this book began.

We conclude this quick tour of our model of the early universe with a reminder. You must not think of the Big Bang as a *localized* explosion *in space*, like an exploding superstar. There were no boundaries and there was no single site where the explosion happened. It was an explosion *of space* (and time and matter and energy) that happened everywhere in the universe. All matter and energy that exist today, including the particles of which you are made, came from the Big Bang. We were, and still are, in the midst of a Big Bang; it is all around us.

29.5 What Is the Universe Really Made Of?

Learning Objectives

By the end of this section, you will be able to:

- Specify what fraction of the density of the universe is contributed by stars and galaxies and how much ordinary matter (such as hydrogen, helium, and other elements we are familiar with here on Earth) makes up the overall density
- Describe how ideas about the contents of the universe have changed over the last 50 years
- Explain why it is so difficult to determine what dark matter really is
- Explain why dark matter helped galaxies form quickly in the early universe
- Summarize the evolution of the universe from the time the CMB was emitted to the present day

The model of the universe we described in the previous section is the simplest model that explains the observations. It assumes that general relativity is the correct theory of gravity throughout the universe. With this assumption, the model then accounts for the existence and structure of the CMB; the abundances of the light elements deuterium, helium, and lithium; and the acceleration of the expansion of the universe. All of the observations to date support the validity of the model, which is referred to as the standard (or concordance) model of cosmology.

Figure 29.21 and Table 29.2 summarize the current best estimates of the contents of the universe. Luminous matter in stars and galaxies and neutrinos contributes about 1% of the mass required to reach critical density. Another 4% is mainly in the form of hydrogen and helium in the space between stars and in intergalactic space. Dark matter accounts for about an additional 27% of the critical density. The mass equivalent of dark energy (according to $E = mc^2$) then supplies the remaining 68% of the critical density.

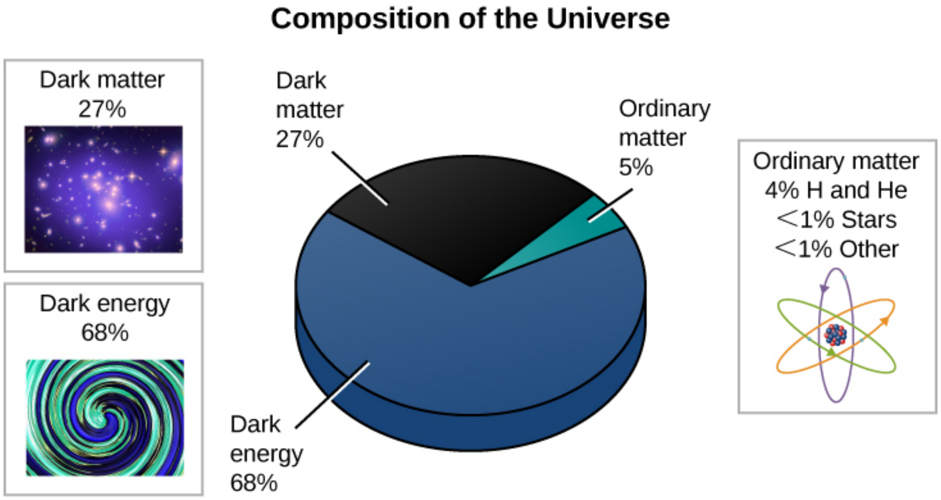

Figure 29.21 Composition of the Universe. Only about 5% of all the mass and energy in the universe is matter with which we are familiar here on Earth. Most ordinary matter consists of hydrogen and helium located in interstellar and intergalactic space. Only about one-half of 1% of the critical density of the universe is found in stars. Dark matter and dark energy, which have not yet been detected in earthbound laboratories, account for 95% of the contents of the universe.

What Different Kinds of Objects Contribute to the Density of the Universe

Object	Density as a Percent of Critical Density
Luminous matter (stars, etc.)	<1
Hydrogen and helium in interstellar and intergalactic space	4
Dark matter	27
Equivalent mass density of the dark energy	68

Table 29.2

This table should shock you. What we are saying is that 95% of the stuff of the universe is either dark matter or dark energy—neither of which has ever been detected in a laboratory here on Earth. This whole textbook, which has focused on objects that emit electromagnetic radiation, has generally been ignoring 95% of what is out there. Who says there aren’t big mysteries yet to solve in science!

Figure 29.22 shows how our ideas of the composition of the universe have changed over just the past three decades. The fraction of the universe that we think is made of the same particles as astronomy students has been decreasing steadily.

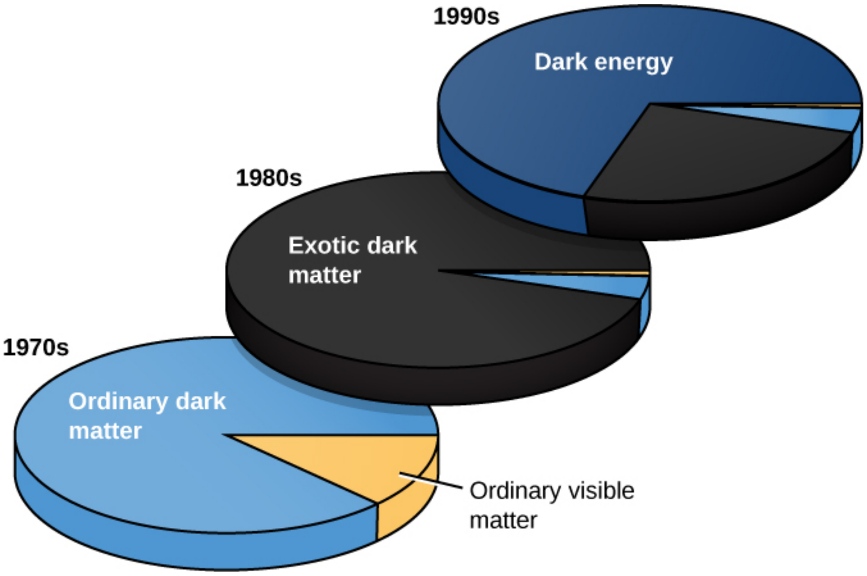

Figure 29.22 Changing Estimates of the Content of the Universe. This diagram shows the changes in our understanding of the contents of the universe over the past three decades. In the 1970s, we suspected that most of the matter in the universe was invisible, but we thought that this matter might be ordinary matter (protons, neutrons, etc.) that was simply not producing electromagnetic radiation. By the 1980s, it was becoming likely that most of the dark matter was made of something we had not yet detected on Earth. By the late 1990s, a variety of experiments had shown that we live in a critical -density universe and that dark energy contributes about 70% of what is required to reach critical density. Note how the estimate of the relative importance of ordinary luminous matter (shown in yellow) has diminished over time.

What Is Dark Matter?

Many astronomers find the situation we have described very satisfying. Several independent experiments now agree on the type of universe we live in and on the inventory of what it contains. We seem to be very close to having a cosmological model that explains nearly everything. Others are not yet ready to jump on the bandwagon. They say, “show me the 96% of the universe we can’t detect directly—for example, find me some dark matter!”

At first, astronomers thought that **dark matter** might be hidden in objects that appear dark because they emit no light (e.g., black holes) or that are too faint to be observed at large distances (e.g., planets or white dwarfs). However, these objects would be made of ordinary matter, and the deuterium abundance tells us that no more than 5% of the critical density consists of ordinary matter.

Another possible form that dark matter can take is some type of elementary particle that we have not yet detected here on Earth—a particle that has mass and exists in sufficient abundance to contribute 23% of the critical density. Some physics theories predict the existence of such particles. One class of these particles has been given the name WIMPs, which stands for **weakly interacting massive particles**. Since these particles do not participate in nuclear reactions leading to the production of deuterium, the deuterium abundance puts no limits on how many WIMPs might be in the universe. (A number of other exotic particles have also been suggested as prime constituents of dark matter, but we will confine our discussion to WIMPs as a useful example.)

If large numbers of WIMPs do exist, then some of them should be passing through our physics laboratories right now. The trick is to catch them. Since by definition they interact only weakly (infrequently) with other matter, the chances that they will have a measurable effect are small. We don’t know the mass of these particles, but various theories suggest that it might be a few to a few hundred times the mass of a proton. If WIMPs are 60 times the mass of a proton, there would be about 10 million of them passing through your outstretched hand every second—with absolutely no effect on you. If that seems too mind-boggling, bear in mind that neutrinos interact weakly with ordinary matter, and yet we were able to “catch” them eventually.

Despite the challenges, more than 30 experiments designed to detect WIMPS are in operation or in the planning stages. Predictions of how many times WIMPs might actually collide with the nucleus of an atom in the instrument designed to detect them are in the range of 1 event per year to 1 event per 1000 years per kilogram of detector. The detector must therefore be large. It must be shielded from radioactivity or other types of particles, such as neutrons, passing through it, and hence these detectors are placed in deep mines. The energy imparted to an atomic nucleus in the detector by collision with a WIMP will be small, and so the detector must be cooled to a very low temperature.

The WIMP detectors are made out of crystals of germanium, silicon, or xenon. The detectors are cooled to a few thousandths of a degree—very close to absolute zero. That means that the atoms in the detector are so cold that they are scarcely vibrating at all. If a dark matter particle collides with one of the atoms, it will cause the whole crystal to vibrate and the temperature therefore to increase ever so slightly. Some other interactions may generate a detectable flash of light.

A different kind of search for WIMPs is being conducted at the Large Hadron Collider (LHC) at CERN, Europe’s particle physics lab near Geneva, Switzerland. In this experiment, protons collide with enough energy potentially to produce WIMPs. The LHC detectors cannot detect the WIMPs directly, but if WIMPs are produced, they will pass through the detectors, carrying energy away with them. Experimenters will then add up all the energy that they detect as a result of the collisions of protons to determine if any energy is missing.

So far, none of these experiments has detected WIMPs. Will the newer experiments pay off? Or will scientists have to search for some other explanation for dark matter? Only time will tell (Figure 29.23).

Figure 29.23 Dark Matter. This cartoon from NASA takes a humorous look at how little we yet understand about dark matter. (credit: NASA)

Dark Matter and the Formation of Galaxies

As elusive as dark matter may be in the current-day universe, galaxies could not have formed quickly without it. Galaxies grew from density fluctuations in the early universe, and some had already formed only about 400–500 million years after the Big Bang. The observations with WMAP, Planck, and other experiments give us information on the size of those density fluctuations. It turns out that the density variations we observe are too small to have formed galaxies so soon after the Big Bang. In the hot, early universe, energetic photons collided with hydrogen and helium, and kept them moving so rapidly that gravity was still not strong enough to cause the atoms to come together to form galaxies. How can we reconcile this with the fact that galaxies *did* form and are all around us?

Our instruments that measure the CMB give us information about density fluctuations only for *ordinary matter*, which interacts with radiation. Dark matter, as its name indicates, does not interact with photons at all. Dark matter could have had much greater variations in density and been able to come together to form gravitational “traps” that could then have begun to attract ordinary matter immediately after the universe became transparent. As ordinary matter became increasingly concentrated, it could have turned into galaxies quickly thanks to these dark matter traps.

For an analogy, imagine a boulevard with traffic lights every half mile or so. Suppose you are part of a motorcade of cars accompanied by police who lead you past each light, even if it is red. So, too, when the early universe was opaque, radiation interacted with ordinary matter, imparting energy to it and carrying it along, sweeping past the concentrations of dark matter. Now suppose the police leave the motorcade, which then encounters some red lights. The lights act as traffic traps; approaching cars now have to stop, and so they bunch up. Likewise, after the early universe became transparent, ordinary matter interacted with radiation only occasionally and so could fall into the dark matter traps.

The Universe in a Nutshell

In the previous sections of this chapter, we traced the evolution of the universe progressively further back in time. Astronomical discovery has followed this path historically, as new instruments and new techniques have allowed us to probe ever closer to the beginning of time. The rate of expansion of the universe was determined from measurements of nearby galaxies. Determinations of the abundances of deuterium, helium, and lithium based on nearby stars and galaxies were used to put limits on how much ordinary matter is in the universe. The motions of stars in galaxies and of galaxies within clusters of galaxies could only be explained if there were large quantities of dark matter. Measurements of supernovae that exploded when the universe was about half as old as it is now indicated that the rate of expansion of the universe has sped up since those explosions occurred. Observations of extremely faint galaxies show that galaxies had begun to form when the universe was only 400–500 million years old. And observations of the CMB confirmed early theories that the universe was initially very hot.

But all this moving further and further backward in time might have left you a bit dizzy. So now let’s instead show how the universe evolves as time moves forward.

Figure 29.24 summarizes the entire history of the observable universe from the beginning in a single diagram. The universe was very hot when it began to expand. We have fossil remnants of the very early universe in the form of neutrons, protons, electrons, and neutrinos, and the atomic nuclei that formed when the universe was 3–4 minutes old: deuterium, helium, and a small amount of lithium. Dark matter also remains, but we do not yet know what form it is in.

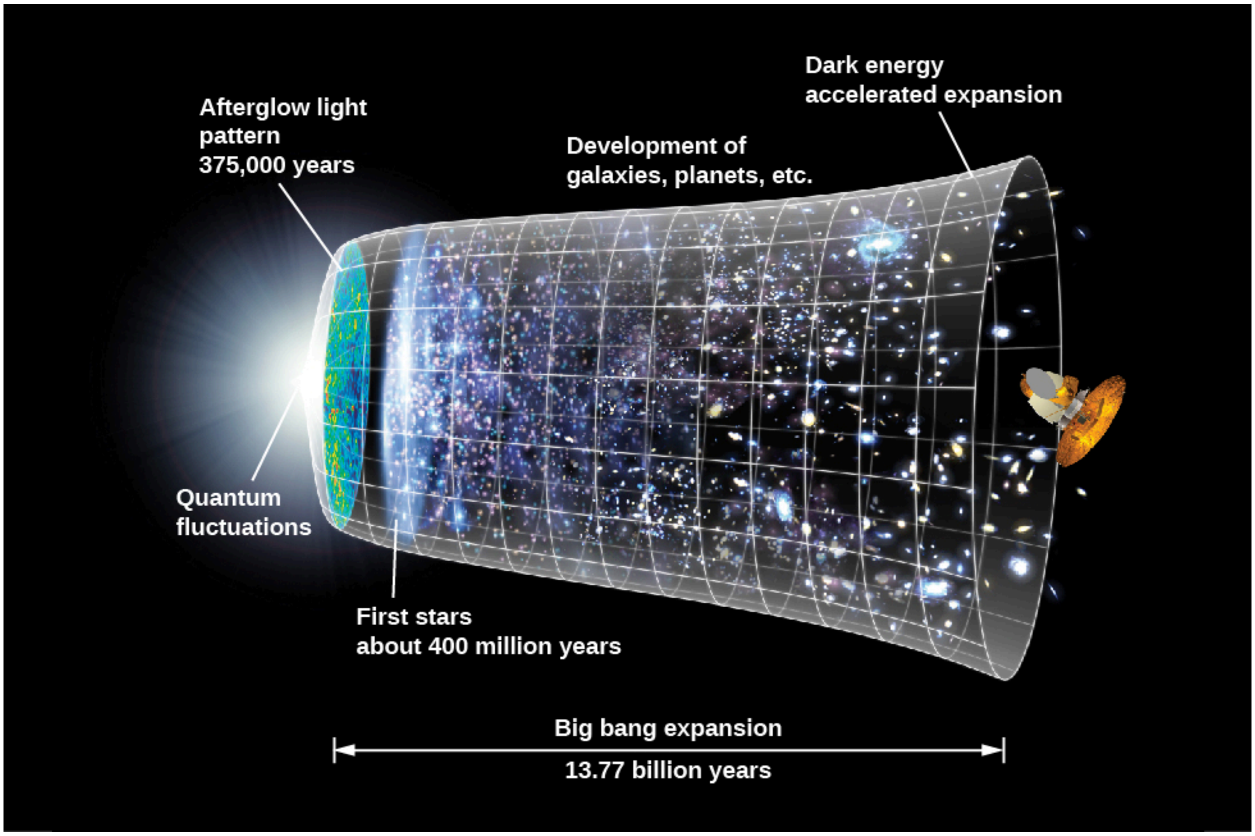

Figure 29.24 History of the Universe. This image summarizes the changes that have occurred in the universe during the last 13.8 billion years. Protons, deuterium, helium, and some lithium were produced in the initial fireball. About 380,000 years after the Big Bang, the universe became transparent to electromagnetic radiation for the first time. COBE, WMAP, Planck, and other instruments have been used to study the radiation that was emitted at that time and that is still visible today (the CMB). The universe was then dark (except for this background radiation) until the first stars and galaxies began to form only a few hundred million years after the Big Bang. Existing space and ground-based telescopes have made substantial progress in studying the subsequent evolution of galaxies. (credit: modification of work by NASA/WMAP Science Team)

The universe gradually cooled; when it was about 380,000 years old, and at a temperature of about 3000 K, electrons combined with protons to form hydrogen atoms. At this point, as we saw, the universe became transparent to light, and astronomers have detected the CMB emitted at this time. The universe still contained no stars or galaxies, and so it entered what astronomers call “the dark ages” (since stars were not lighting up the darkness). During the next several hundred million years, small fluctuations in the density of the dark matter grew, forming gravitational traps that concentrated the ordinary matter, which began to form galaxies about 400–500 million years after the Big Bang.

By the time the universe was about a billion years old, it had entered its own renaissance: it was again blazing with radiation, but this time from newly formed stars, star clusters, and small galaxies. Over the next several billion years, small galaxies merged to form the giants we see today. Clusters and superclusters of galaxies began to grow, and the universe eventually began to resemble what we see nearby.

During the next 20 years, astronomers plan to build giant new telescopes both in space and on the ground to explore even further back in time. In 2021, the James Webb Space Telescope, a 6.5-meter telescope that is the successor to the Hubble Space Telescope, will be launched and assembled in space. The predictions are that with this powerful instrument (see Figure 29.1) we should be able to look back far enough to analyze in detail the formation of the first galaxies.

29.6 The Inflationary Universe

Learning Objectives

By the end of this section, you will be able to:

- Describe two important properties of the universe that the simple Big Bang model cannot explain
- Explain why these two characteristics of the universe can be accounted for if there was a period of rapid expansion (inflation) of the universe just after the Big Bang
- Name the four forces that control all physical processes in the universe

The hot Big Bang model that we have been describing is remarkably successful. It accounts for the expansion of the universe, explains the observations of the CMB, and correctly predicts the abundances of the light elements. As it turns out, this model also predicts that there should be exactly three types of neutrinos in nature, and this prediction has been confirmed by experiments with high-energy accelerators. We can’t relax just yet, however. This standard model of the universe doesn’t explain *all* the observations we have made about the universe as a whole.

Problems with the Standard Big Bang Model

There are a number of characteristics of the universe that can only be explained by considering further what might have happened before the emission of the CMB. One problem with the standard Big Bang model is that it does not explain why the density of the universe is equal to the critical density. The mass density could have been, after all, so low and the effects of dark energy so high that the expansion would have been too rapid to form any galaxies at all. Alternatively, there could have been so much matter that the universe would have already begun to contract long before now. Why is the universe balanced so precisely on the knife edge of the critical density?

Another puzzle is the remarkable *uniformity* of the universe. The temperature of the CMB is the same to about 1 part in 100,000 everywhere we look. This sameness might be expected if all the parts of the visible universe were in contact at some point in time and had the time to come to the same temperature. In the same way, if we put some ice into a glass of lukewarm water and wait a while, the ice will melt and the water will cool down until they are the same temperature.

However, if we accept the standard Big Bang model, all parts of the visible universe were *not* in contact at any time. The fastest that information can go from one point to another is the speed of light. There is a maximum distance that light can have traveled from any point since the time the universe began—that’s the distance light could have covered since then. This distance is called that point’s *horizon distance* because anything farther away is “below its horizon”—unable to make contact with it. One region of space separated by more than the horizon distance from another has been completely isolated from it through the entire history of the universe.

If we measure the CMB in two opposite directions in the sky, we are observing regions that were significantly beyond each other’s horizon distance at the time the CMB was emitted. We can see both regions, but *they* can never have seen each other. Why, then, are their temperatures so precisely the same? According to the standard Big Bang model, they have never been able to exchange information, and there is no reason they should have identical temperatures. (It’s a little like seeing the clothes that all the students wear at two schools in different parts of the world become identical, without the students ever having been in contact.) The only explanation we could suggest was simply that the universe somehow *started out* being absolutely uniform (which is like saying all students were born liking the same clothes). Scientists are always uncomfortable when they must appeal to a special set of initial conditions to account for what they see.

The Inflationary Hypothesis

Some physicists suggested that these fundamental characteristics of the cosmos—its flatness and uniformity—can be explained if shortly after the Big Bang (and before the emission of the CMB), the universe experienced a sudden increase in size. A model universe in which this rapid, early expansion occurs is called an **inflationary universe**. The inflationary universe is identical to the Big Bang universe for all time after the first 10^{-30} second. Prior to that, the model suggests that there was a brief period of extraordinarily rapid expansion or inflation, during which the scale of the universe increased by a factor of about 10^{50} times more than predicted by standard Big Bang models (Figure 29.25).

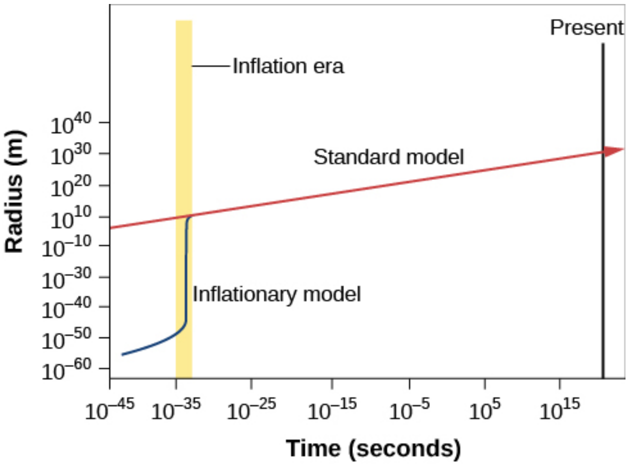

Figure 29.25 Expansion of the Universe. This graph shows how the scale factor of the observable universe changes with time for the standard Big Bang model (red line) and for the inflationary model (blue line). (Note that the time scale at the bottom is extremely compressed.) During inflation, regions that were very small and in contact with each other are suddenly blown up to be much larger and outside each other’s horizon distance. The two models are the same for all times after 10^{-30} second.

Prior to (and during) inflation, all the parts of the universe that we can now see were so small and close to each other that they *could* exchange information, that is, the horizon distance included all of the universe that we can now observe. Before (and during) inflation, there was adequate time for the observable universe to homogenize itself and come to the same temperature. Then, inflation expanded those regions tremendously, so that many parts of the universe are now beyond each other’s horizon.

Another appeal of the inflationary model is its prediction that the density of the universe should be exactly equal to the critical density. To see why this is so, remember that curvature of spacetime is intimately linked to the density of matter. If the universe began with some curvature of its spacetime, one analogy for it might be the skin of a balloon. The period of inflation was equivalent to blowing up the balloon to a tremendous size. The universe became so big that from our vantage point, no curvature should be visible (Figure 29.26). In the same way, Earth’s surface is so big that it looks flat to us no matter where we are. Calculations show that a universe with no curvature is one that is at critical density. Universes with densities either higher or lower than the critical density would show marked curvature. But we saw that the observations of the CMB in Figure 29.18, which show that the universe has critical density, rule out the possibility that space is significantly curved.

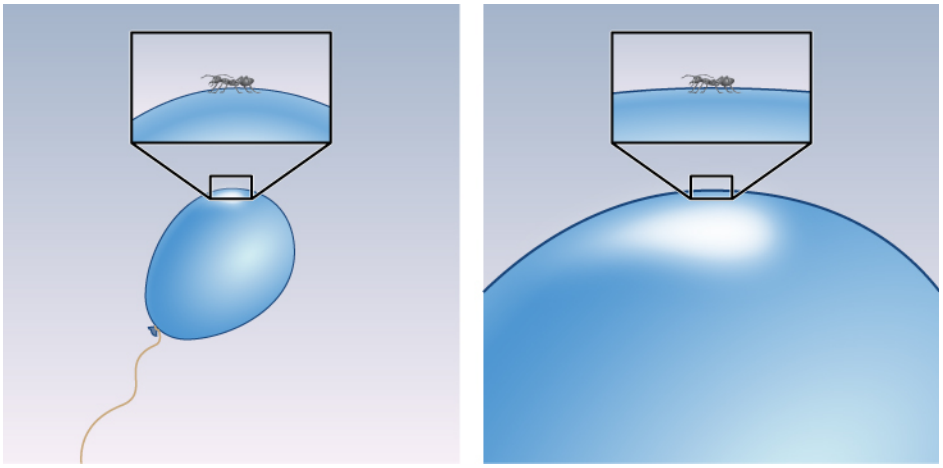

Figure 29.26 Analogy for Inflation. During a period of rapid inflation, a curved balloon grows so large that to any local observer it looks flat. The inset shows the geometry from the ant's point of view.

Grand Unified Theories

While inflation is an intriguing idea and widely accepted by researchers, we cannot directly observe events so early in the universe. The conditions at the time of inflation were so extreme that we cannot reproduce them in our laboratories or high-energy accelerators, but scientists have some ideas about what the universe might have been like. These ideas are called **grand unified theories** or GUTs.

In GUT models, the forces that we are familiar with here on Earth, including gravity and electromagnetism, behaved very differently in the extreme conditions of the early universe than they do today. In physical science, the term *force* is used to describe anything that can change the motion of a particle or body. One of the remarkable discoveries of modern science is that all known physical processes can be described through the action of just four forces: gravity, electromagnetism, the strong nuclear force, and the weak nuclear force (Table 29.3).

The Forces of Nature

Force	Relative Strength Today	Range of Action	Important Applications
Gravity	1	Whole universe	Motions of planets, stars, galaxies
Electromagnetism	10^{36}	Whole universe	Atoms, molecules, electricity, magnetic fields
Weak nuclear force	10^{33}	10^{-17} meters	Radioactive decay
Strong nuclear force	10^{38}	10^{-15} meters	The existence of atomic nuclei

Table 29.3

Gravity is perhaps the most familiar force, and certainly appears strong if you jump off a tall building. However, the force of gravity between two elementary particles—say two protons—is by far the weakest of the four forces. Electromagnetism—which includes both magnetic and electrical forces, holds atoms together, and produces the electromagnetic radiation that we use to study the universe—is much stronger, as you can see in Table 29.3. The weak nuclear force is only weak in comparison to its strong "cousin," but it is in fact much stronger than gravity.

Both the weak and strong nuclear forces differ from the first two forces in that they act only over very small distances—those comparable to the size of an atomic nucleus or less. The weak force is involved in radioactive decay and in reactions that result in the production of neutrinos. The strong force holds protons and neutrons together in an atomic nucleus.

Physicists have wondered why there are four forces in the universe—why not 300 or, preferably, just one? An important hint comes from the name *electromagnetic force*. For a long time, scientists thought that the forces of electricity and magnetism were separate, but James Clerk Maxwell (see the chapter on Radiation and Spectra) was able to *unify* these forces—to show that they are aspects of the same phenomenon. In the same way, many scientists (including Einstein) have wondered if the four forces we now know could also be unified. Physicists have actually developed GUTs that unify three of the four forces (but not gravity).

In these theories, the strong, weak, and electromagnetic forces are not three independent forces but instead are different manifestations or aspects of what is, in fact, a single force. The theories predict that at high enough temperatures, there would be only one force. At lower temperatures (like the ones in the universe today), however, this single force has changed into three different forces (Figure 29.27). Just as different gases or liquids freeze at different temperatures, we can say that the different forces "froze out" of the unified force at different temperatures. Unfortunately, the temperatures at which the three forces acted as one force are so high that they cannot be reached in any laboratory on Earth. Only the early universe, at times prior to 10^{-35} second, was hot enough to unify these forces.

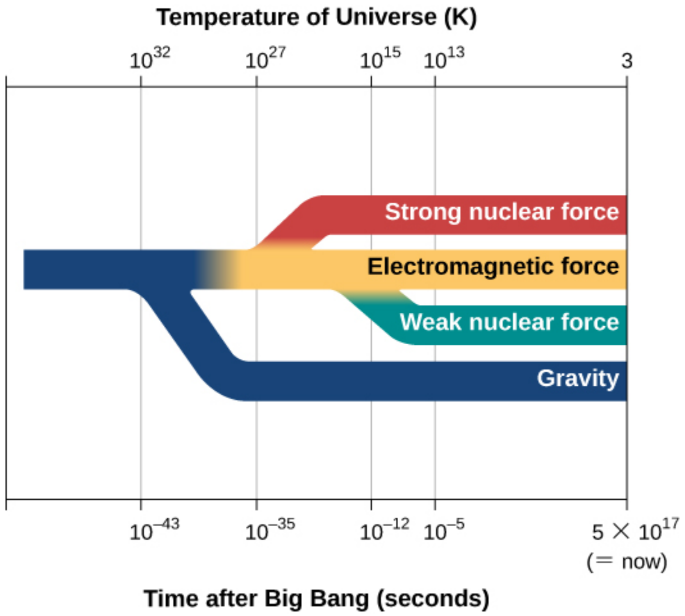

Figure 29.27 Four Forces That Govern the Universe. The behavior of the four forces depends on the temperature of the universe. This diagram (inspired by some grand unified theories) shows that at very early times when the temperature of the universe was very high, all four forces resembled one another and were indistinguishable. As the universe cooled, the forces took on separate and distinctive characteristics.

Many physicists think that gravity was also unified with the three other forces at still higher temperatures, and scientists have tried to develop a theory that combines all four forces. For example, in string theory, the point-like particles of matter that we have discussed in this book are replaced by one-dimensional objects called strings. In this theory, infinitesimal strings, which have length but not height or width, are the building blocks used to construct all the forms of matter and energy in the universe. These strings exist in 11-dimensional space (not the 4-dimensional spacetime with which we are familiar). The strings vibrate in the various dimensions, and depending on how they vibrate, they are seen in our world as matter or gravity or light. As you can imagine, the mathematics of string theory is very complex, and the theory remains untested by experiments. Even the largest particle accelerators on Earth do not achieve high enough energy to show whether string theory applies to the real world.

String theory is interesting to scientists because it is currently the only approach that seems to have the potential of combining all four forces to produce what physicists have termed the Theory of Everything.[3] Theories of the earliest phases of the universe must take both quantum mechanics and gravity into account, but at the simplest level, gravity and quantum mechanics are incompatible. General relativity, our best theory of gravity, says that the motions of objects can be predicted exactly. Quantum mechanics says you can only calculate the probability (chance) that an object will do something. String theory is an attempt to resolve this paradox. The mathematics that underpins string theory is elegant and beautiful, but it remains to be seen whether it will make predictions that can be tested by observations in yet-to-be-developed, high-energy accelerators on Earth or by observations of the early universe.

The earliest period in the history of the universe from time zero to 10^{-43} second is called the Planck time. The universe was unimaginably hot and dense, and theorists believe that at this time, quantum effects of gravity dominated physical interactions—and, as we have just discussed, we have no tested theory of quantum gravity. Inflation is hypothesized to have occurred somewhat later, when the universe was between perhaps 10^{-35} and 10^{-33} second old and the temperature was 10^{27} to 10^{28} K. This rapid expansion took place when three forces (electromagnetic, strong, and weak) are thought to have been unified, and this is when GUTs are applicable.

After inflation, the universe continued to expand (but more slowly) and to cool. An important milestone was reached when the temperature was down to 10^{15} K and the universe was 10^{-10} second old. Under these conditions, all four forces were separate and distinct. High-energy particle accelerators can achieve similar conditions, and so theories of the history of the universe from this point on have a sound basis in experiments.

As yet, we have no direct evidence of what the conditions were during the inflationary epoch, and the ideas presented here are speculative. Researchers are trying to devise some experimental tests. For example, the quantum fluctuations in the very early universe would have caused variations in density and produced gravitational waves that may have left a detectable imprint on the CMB. Detection of such an imprint will require observations with equipment whose sensitivity is improved from what we have today. Ultimately, however, it may provide confirmation that we live in a universe that once experienced an epoch of rapid inflation.

If you are typical of the students who read this book, you may have found this brief discussion of dark matter, inflation, and cosmology a bit frustrating. We have offered glimpses of theories and observations, but have raised more questions than we have answered. What is dark matter? What is dark energy? Inflation explains the observations of flatness and uniformity of the university, but did it actually happen? These ideas are at the forefront of modern science, where progress almost always leads to new puzzles, and much more work is needed before we can see clearly. Bear in mind that less than a century has passed since Hubble demonstrated the existence of other galaxies. The quest to understand just how the universe of galaxies came to be will keep astronomers busy for a long time to come.

29.7 The Anthropic Principle

Learning Objectives

By the end of this section, you will be able to:

- Name some properties of the universe that, if different, would have precluded the development of humans

---

[3] This name became the title of a film about physicist Stephen Hawking in 2014.

Despite our uncertainties, we must admit that the picture we have developed about the evolution of our universe is a remarkable one. With new telescopes, we have begun to collect enough observational evidence that we can describe how the universe evolved from a mere fraction of a second after the expansion began. Although this is an impressive achievement, there are still some characteristics of the universe that we cannot explain. And yet, it turns out that if these characteristics were any different, we would not be here to ask about them. Let's look at some of these "lucky accidents," beginning with the observations of the cosmic microwave background (CMB).

Lucky Accidents

As we described in this chapter, the CMB is radiation that was emitted when the universe was a few hundred thousand years old. Observations show that the temperature of the radiation varies from one region to another, typically by about 10 millionths of a degree, and these temperature differences signal small differences in density. But suppose the tiny, early fluctuations in density had been much smaller. Then calculations show that the pull of gravity near them would have been so small that no galaxies would ever have formed.

What if the fluctuations in density had been much larger? Then it is possible that very dense regions would have condensed, and these would simply have collapsed directly to black holes without ever forming galaxies and stars. Even if galaxies had been able to form in such a universe, space would have been filled with intense X-rays and gamma rays, and it would have been difficult for life forms to develop and survive. The density of stars within galaxies would be so high that interactions and collisions among them would be frequent. In such a universe, any planetary systems could rarely survive long enough for life to develop.

So for us to be here, the density fluctuations need to be "just right"—not too big and not too small.

Another lucky accident is that the universe is finely balanced between expansion and contraction. It is expanding, but very slowly. If the expansion had been at a much higher rate, all of the matter would have thinned out before galaxies could form. If everything were expanding at a much slower rate, then gravity would have "won." The expansion would have reversed and all of the matter would have recollapsed, probably into a black hole—again, no stars, no planets, no life.

The development of life on Earth depends on still-luckier coincidences. Had matter and antimatter been present initially in exactly equal proportions, then all matter would have been annihilated and turned into pure energy. We owe our existence to the fact that there was slightly more matter than antimatter. (After most of the matter made contact with an equal amount of antimatter, turning into energy, a small amount of additional matter must have been present. We are all descendants of that bit of "unbalanced" matter.)

If nuclear fusion reactions occurred at a somewhat faster rate than they actually do, then at the time of the initial fireball, all of the matter would have been converted from hydrogen into helium into carbon and all the way into iron (the most stable nucleus). That would mean that no stars would have formed, since the existence of stars depends on there being light elements that can undergo fusion in the main-sequence stage and make the stars shine. In addition, the structure of atomic nuclei had to be just right to make it possible for three helium atoms to come together easily to fuse carbon, which is the basis of life. If the *triple-alpha process* we discussed in the chapter on Stars from Adolescence to Old Age were too unlikely, not enough carbon would have formed to lead to biology as we know it. At the same time, it had to be hard enough to fuse carbon into oxygen that a large amount of carbon survived for billions of years.

There are additional factors that have contributed to life like us being possible. Neutrinos have to interact with matter at just the right, albeit infrequent, rate. Supernova explosions occur when neutrinos escape from the cores of collapsing stars, deposit some of their energy in the surrounding stellar envelope, and cause it to blow out and away into space. The heavy elements that are ejected in such explosions are essential ingredients of life here on Earth. If neutrinos did not interact with matter at all, they would escape from the cores of collapsing stars without causing the explosion. If neutrinos interacted strongly with matter, they would remain trapped in the stellar core. In either case, the heavy elements would remain locked up inside the collapsing star.

If gravity were a much stronger force than it is, stars could form with much smaller masses, and their lifetimes would be measured in years rather than billions of years. Chemical processes, on the other hand, would not be sped up if gravity were a stronger force, and so there would be no time for life to develop while stars were so short-lived. Even if life did develop in a stronger-gravity universe, life forms would have to be tiny or they could not stand up or move around.

What Had to Be, Had to Be

In summary, we see that a specific set of rules and conditions in the universe has allowed complexity and life on Earth to develop. As yet, we have no theory to explain why this “right” set of conditions occurred. For this reason, many scientists are beginning to accept an idea we call the **anthropic principle**—namely, that the physical laws we observe must be what they are precisely because these are the only laws that allow for the existence of humans.

Some scientists speculate that our universe is but one of countless universes, each with a different set of physical laws—an idea that is sometimes referred to as the **multiverse**. Some of those universes might be stillborn, collapsing before any structure forms. Others may expand so quickly that they remain essentially featureless with no stars and galaxies. In other words, there may be a much larger multiverse that contains our own universe and many others. This multiverse (existing perhaps in more dimensions that we can become aware of) is infinite and eternal; it generates many, many inflating regions, each of which evolves into a separate universe, which may be completely unlike any of the other separate universes. Our universe is then the way it is because it is the only way it could be and have humans like ourselves in it to discover its properties and ask such questions.

LINK TO LEARNING

View the 2011 introductory talk (https://openstax.org/l/30mulcosinfla) on the Multiverse and Cosmic Inflation by Dr. Anthony Aguirre of the University of California, part of the Silicon Valley Astronomy Lecture Series.

It is difficult to know how to test these ideas since we can never make contact with any other universe. For most scientists, our discussion in this section borders on the philosophical and metaphysical. Perhaps in the future our understanding of physics will develop to the point that we can know why the gravitational constant is as strong as it is, why the universe is expanding at exactly the rate it is, and why all of the other “lucky accidents” happened—why they were inevitable and could be no other way. Then this anthropic idea would no longer be necessary. No one knows, however, whether we will ever have an explanation for why this universe works the way it does.

We have come a long way in our voyage through the universe. We have learned a remarkable amount about *how* and *when* the cosmos came to be, but the question of *why* the universe is the way it is remains as elusive as ever.

Key Terms

anthropic principle idea that physical laws must be the way they are because otherwise we could not be here to measure them

Big Bang the theory of cosmology in which the expansion of the universe began with a primeval explosion (of space, time, matter, and energy)

closed universe a model in which the universe expands from a Big Bang, stops, and then contracts to a big crunch

cosmic microwave background (CMB) microwave radiation coming from all directions that is the redshifted afterglow of the Big Bang

cosmological constant the term in the equations of general relativity that represents a repulsive force in the universe

cosmology the study of the organization and evolution of the universe

critical density in cosmology, the density that is just sufficient to bring the expansion of the universe to a stop after infinite time

dark energy the energy that is causing the expansion of the universe to accelerate; its existence is inferred from observations of distant supernovae

dark matter nonluminous material, whose nature we don’t yet understand, but whose presence can be inferred because of its gravitational influence on luminous matter

deuterium a form of hydrogen in which the nucleus of each atom consists of one proton and one neutron

flat universe a model of the universe that has a critical density and in which the geometry of the universe is flat, like a sheet of paper

fusion the building of heavier atomic nuclei from lighter ones

grand unified theories (GUTs) physical theories that attempt to describe the four forces of nature as different manifestations of a single force

inflationary universe a theory of cosmology in which the universe is assumed to have undergone a phase of very rapid expansion when the universe was about 10^{-35} second old; after this period of rapid expansion, the standard Big Bang and inflationary models are identical

lithium the third element in the periodic table; lithium nuclei with three protons and four neutrons were manufactured during the first few minutes of the expansion of the universe

multiverse the speculative idea that our universe is just one of many universes, each with its own set of physical laws

open universe a model in which the density of the universe is not high enough to bring the expansion of the universe to a halt

photon decoupling time when radiation began to stream freely through the universe without interacting with matter

weakly interacting massive particles (WIMPs) weakly interacting massive particles are one of the candidates for the composition of dark matter

Summary

29.1 The Age of the Universe

Cosmology is the study of the organization and evolution of the universe. The universe is expanding, and this is one of the key observational starting points for modern cosmological theories. Modern observations show that the rate of expansion has not been constant throughout the life of the universe. Initially, when galaxies were close together, the effects of gravity were stronger than the effects of dark energy, and the expansion rate gradually slowed. As galaxies moved farther apart, the influence of gravity on the expansion rate weakened. Measurements of distant supernovae show that when the universe was about half its current age, dark energy began to dominate the rate of expansion and caused it to speed up. In order to estimate the age of the universe, we must allow for changes in the rate of expansion. After allowing for these effects, astronomers estimate that all of the matter within the observable universe was concentrated in an extremely small volume 13.8 billion years ago, a time we call the Big Bang.

29.2 A Model of the Universe

For describing the large-scale properties of the universe, a model that is isotropic and homogeneous (same everywhere) is a pretty good approximation of reality. The universe is expanding, which means that the universe undergoes a change in scale with time; space stretches and distances grow larger by the same factor everywhere at a given time. Observations show that the mass density of the universe is less than the critical density. In other words, there is not enough matter in the universe to stop the expansion. With the discovery of dark energy, which is accelerating the rate of expansion, the observational evidence is strong that the universe will expand forever. Observations tell us that the expansion started about 13.8 billion years ago.

29.3 The Beginning of the Universe

Lemaître, Alpher, and Gamow first worked out the ideas that are today called the Big Bang theory. The universe cools as it expands. The energy of photons is determined by their temperature, and calculations show that in the hot, early universe, photons had so much energy that when they collided with one another, they could produce material particles. As the universe expanded and cooled, protons and neutrons formed first, then came electrons and positrons. Next, fusion reactions produced deuterium, helium, and lithium nuclei. Measurements of the deuterium abundance in today’s universe show that the total amount of ordinary matter in the universe is only about 5% of the critical density.

29.4 The Cosmic Microwave Background

When the universe became cool enough to form neutral hydrogen atoms, the universe became transparent to radiation. Scientists have detected the cosmic microwave background (CMB) radiation from this time during the hot, early universe. Measurements with the COBE satellite show that the CMB acts like a blackbody with a temperature of 2.73 K. Tiny fluctuations in the CMB show us the seeds of large-scale structures in the universe. Detailed measurements of these fluctuations show that we live in a critical-density universe and that the critical density is composed of 31% matter, including dark matter, and 69% dark energy. Ordinary matter—the kinds of elementary particles we find on Earth—make up only about 5% of the critical density. CMB measurements also indicate that the universe is 13.8 billion years old.

29.5 What Is the Universe Really Made Of?

Twenty-seven percent of the critical density of the universe is composed of dark matter. To explain so much dark matter, some physics theories predict that additional types of particles should exist. One type has been given the name of WIMPs (weakly interacting massive particles), and scientists are now conducting experiments to try to detect them in the laboratory. Dark matter plays an essential role in forming galaxies. Since, by definition, these particles interact only very weakly (if at all) with radiation, they could have congregated while the universe was still very hot and filled with radiation. They would thus have formed gravitational traps that quickly attracted and concentrated ordinary matter after the universe became transparent, and matter and radiation decoupled. This rapid concentration of matter enabled galaxies to form by the time the universe was only 400–500 million years old.

29.6 The Inflationary Universe

The Big Bang model does not explain why the CMB has the same temperature in all directions. Neither does it explain why the density of the universe is so close to critical density. These observations can be explained if the universe experienced a period of rapid expansion, which scientists call inflation, about 10^{-35} second after the Big Bang. New grand unified theories (GUTs) are being developed to describe physical processes in the universe before and at the time that inflation occurred.

29.7 The Anthropic Principle

Recently, many cosmologists have noted that the existence of humans depends on the fact that many properties of the universe—the size of density fluctuations in the early universe, the strength of gravity, the structure of atoms—were just right. The idea that physical laws must be the way they are because otherwise we could not be here to measure them is called the anthropic principle. Some scientists speculate that there may be a multiverse of universes, in which ours is just one.

For Further Exploration

Articles

The January 2021 issue of *Astronomy* magazine is devoted to our modern understanding of the beginning and end of the universe.

Kruesi, L. “Cosmology: 5 Things You Need to Know.” *Astronomy* (May 2007): 28. Five questions students often ask, and how modern cosmologists answer them.

Kruesi, L. “How Planck Has Redefined the Universe.” *Astronomy* (October 2013): 28. Good review of what this space mission has told us about the CMB and the universe.

Lineweaver, C. & Davis, T. “Misconceptions about the Big Bang.” *Scientific American* (March 2005): 36. Some basic ideas about modern cosmology clarified, using general relativity.

Nadis, S. “Sizing Up Inflation.” *Sky & Telescope* (November 2005): 32. Nice review of the origin and modern variants on the inflationary idea.

Nadis, S. “How We Could See Another Universe.” *Astronomy* (June 2009): 24. On modern ideas about multiverses and how such bubbles of space-time might collide.

Nadis, S. “Dark Energy’s New Face: How Exploding Stars Are Changing our View.” *Astronomy* (July 2012): 45. About our improving understanding of the complexities of type Ia supernovae.

Naze, Y. “The Priest, the Universe, and the Big Bang.” *Astronomy* (November 2007): 40. On the life and work of Georges Lemaître.

Panek, R. “Going Over to the Dark Side.” *Sky & Telescope* (February 2009): 22. A history of the observations and theories about dark energy.

Pendrick, D. “Is the Big Bang in Trouble?” *Astronomy* (April 2009): 48. This sensationally titled article is really more of a quick review of how modern ideas and observations are fleshing out the Big Bang hypothesis (and raising questions.)

Reddy, F. “How the Universe Will End.” *Astronomy* (September 2014): 38. Brief discussion of local and general future scenarios.

Riess, A. and Turner, M. “The Expanding Universe: From Slowdown to Speedup.” *Scientific American* (September 2008): 62.

Turner, M. “The Origin of the Universe.” *Scientific American* (September 2009): 36. An introduction to modern cosmology.

Websites

Cosmology Primer: https://preposterousuniverse.com/cosmologyprimer/ (https://preposterousuniverse.com/cosmologyprimer/). Caltech Astrophysicist Sean Carroll offers a non-technical site with brief overviews of many key topics in modern cosmology.

How Big Is the Universe?: http://www.pbs.org/wgbh/nova/space/how-big-universe.html (http://www.pbs.org/

wgbh/nova/space/how-big-universe.html). A clear essay by a noted astronomer Brent Tully summarizes some key ideas in cosmology and introduces the notion of the acceleration of the universe.

Universe 101: WMAP Mission Introduction to the Universe: http://map.gsfc.nasa.gov/universe/ (http://map.gsfc.nasa.gov/universe/). Concise NASA primer on cosmological ideas from the WMAP mission team.

Cosmic Times Project: http://cosmictimes.gsfc.nasa.gov/ (http://cosmictimes.gsfc.nasa.gov/). James Lochner and Barbara Mattson have compiled a rich resource of twentieth-century cosmology history in the form of news reports on key events, from NASA's Goddard Space Flight Center.

Videos

The Day We Found the Universe: https://www.youtube.com/watch?v=HV23qWIieBw (https://www.youtube.com/watch?v=HV23qWIieBw). Distinguished science writer Marcia Bartusiak discusses Hubble's work and the discovery of the expansion of the cosmos.

Images of the Infant Universe: https://www.youtube.com/watch?v=x0AqCwElyUk (https://www.youtube.com/watch?v=x0AqCwElyUk). Lloyd Knox's public talk on the latest discoveries about the CMB and what they mean for cosmology (1:16:00).

Runaway Universe: https://www.youtube.com/watch?v=kNYVFrnmcOU (https://www.youtube.com/watch?v=kNYVFrnmcOU). Roger Blandford (Stanford Linear Accelerator Center) public lecture on the discovery and meaning of cosmic acceleration and dark energy (1:08:08).

From the Big Bang to the Nobel Prize and on to the James Webb Space Telescope and the Discovery of Alien Life: http://svs.gsfc.nasa.gov/vis/a010000/a010300/a010370/index.html (http://svs.gsfc.nasa.gov/vis/a010000/a010300/a010370/index.html). John Mather, NASA Goddard (1:01:02). His Nobel Prize talk from Dec. 8, 2006 can be found at http://www.nobelprize.org/mediaplayer/index.php?id=74&view=1 (http://www.nobelprize.org/mediaplayer/index.php?id=74&view=1).

Supernovae and the Discovery of the Accelerating Universe: https://www.youtube.com/watch?v=_D6cwrl0CxA (https://www.youtube.com/watch?v=_D6cwrl0CxA). A public lecture by Nobel Laureate Adam Riess in 2019 (58 min).

Collaborative Group Activities

A. This chapter deals with some pretty big questions and ideas. Some belief systems teach us that there are questions to which "we were not meant to know" the answers. Other people feel that if our minds and instruments are capable of exploring a question, then it becomes part of our birthright as thinking human beings. Have your group discuss your personal reactions to discussing questions like the beginning of time and space, and the ultimate fate of the universe. Does it make you nervous to hear about scientists discussing these issues? Or is it exciting to know that we can now gather scientific evidence about the origin and fate of the cosmos? (In discussing this, you may find that members of your group strongly disagree; try to be respectful of others' points of view.)

B. A popular model of the universe in the 1950s and 1960s was the so-called steady-state cosmology. In this model, the universe was not only the same everywhere and in all directions (homogeneous and isotropic), but also the same *at all times*. We know the universe is expanding and the galaxies are thinning out, and so this model hypothesized that new matter was continually coming into existence to fill in the space between galaxies as they moved farther apart. If so, the infinite universe did not have to have a sudden beginning, but could simply exist forever in a steady state. Have your group discuss your reaction to this model. Do you find it more appealing philosophically than the Big Bang model? Can you cite some evidence that indicates that the universe was not the same billions of years ago as it is now—that it is not in a steady state?

C. One of the lucky accidents that characterizes our universe is the fact that the time scale for the development of intelligent life on Earth and the lifetime of the Sun are comparable. Have your group discuss what would happen if the two time scales were very different. Suppose, for example, that the time for intelligent life to evolve was 10 times greater than the main-sequence lifetime of the Sun. Would our civilization have ever developed? Now suppose the time for intelligent life to evolve is ten times shorter than the main-sequence lifetime of the Sun. Would we be around? (This latter discussion requires considerable thought, including such ideas as what the early stages in the Sun's life were like and how much the early Earth was bombarded by asteroids and comets.)

D. The grand ideas discussed in this chapter have a powerful effect on the human imagination, not just for scientists, but also for artists, composers, dramatists, and writers. Here we list just a few of these responses to cosmology. Each member of your group can select one of these, learn more about it, and then report back, either to the group or to the whole class.

- The California poet Robinson Jeffers was the brother of an astronomer who worked at the Lick Observatory. His poem "Margrave" is a meditation on cosmology and on the kidnap and murder of a child: http://www.poemhunter.com/best-poems/robinson-jeffers/margrave/.
- In the science fiction story "The Gravity Mine" by Stephen Baxter, the energy of evaporating supermassive black holes is the last hope of living beings in the far future in an ever-expanding universe. The story has poetic description of the ultimate fate of matter and life and is available online at: http://www.infinityplus.co.uk/stories/gravitymine.htm.
- The musical piece *YLEM* by Karlheinz Stockhausen takes its title from the ancient Greek term for primeval material revived by George Gamow. It tries to portray the oscillating universe in musical terms. Players actually expand through the concert hall, just as the universe does, and then return and expand again. See: http://www.karlheinzstockhausen.org/ylem_english.htm.
- The musical piece *Supernova Sonata* http://www.astro.uvic.ca/~alexhp/new/supernova_sonata.html by Alex Parker and Melissa Graham is based on the characteristics of 241 type Ia supernova explosions, the ones that have helped astronomers discover the acceleration of the expanding universe.
- Gregory Benford's short story "The Final Now" envisions the end of an accelerating open universe, and blends religious and scientific imagery in a very poetic way. Available free online at: http://www.tor.com/stories/2010/03/the-final-now.

E. When Einstein learned about Hubble's work showing that the universe of galaxies is expanding, he called his introduction of the cosmological constant into his general theory of relativity his "biggest blunder." Can your group think of other "big blunders" from the history of astronomy, where the thinking of astronomers was too conservative and the universe turned out to be more complicated or required more "outside-the-box" thinking?

Exercises

Review Questions

1. What are the basic observations about the universe that any theory of cosmology must explain?

2. Describe some possible futures for the universe that scientists have come up with. What property of the universe determines which of these possibilities is the correct one?

3. What does the term Hubble time mean in cosmology, and what is the current best calculation for the Hubble time?

4. Which formed first: hydrogen nuclei or hydrogen atoms? Explain the sequence of events that led to each.

5. Describe at least two characteristics of the universe that are explained by the standard Big Bang model.

6. Describe two properties of the universe that are not explained by the standard Big Bang model (without inflation). How does inflation explain these two properties?
7. Why do astronomers believe there must be dark matter that is not in the form of atoms with protons and neutrons?
8. What is dark energy and what evidence do astronomers have that it is an important component of the universe?
9. Thinking about the ideas of space and time in Einstein’s general theory of relativity, how do we explain the fact that all galaxies outside our Local Group show a redshift?
10. Astronomers have found that there is more helium in the universe than stars could have made in the 13.8 billion years that the universe has been in existence. How does the Big Bang scenario solve this problem?
11. Describe the anthropic principle. What are some properties of the universe that make it “ready” to have life forms like you in it?
12. Describe the evidence that the expansion of the universe is accelerating.

Thought Questions

13. What is the most useful probe of the early evolution of the universe: a giant elliptical galaxy or an irregular galaxy such as the Large Magellanic Cloud? Why?
14. What are the advantages and disadvantages of using quasars to probe the early history of the universe?
15. Would acceleration of the universe occur if it were composed entirely of matter (that is, if there were no dark energy)?
16. Suppose the universe expands forever. Describe what will become of the radiation from the primeval fireball. What will the future evolution of galaxies be like? Could life as we know it survive forever in such a universe? Why?
17. Some theorists expected that observations would show that the density of matter in the universe is just equal to the critical density. Do the current observations support this hypothesis?
18. There are a variety of ways of estimating the ages of various objects in the universe. Describe two of these ways, and indicate how well they agree with one another and with the age of the universe itself as estimated by its expansion.
19. Since the time of Copernicus, each revolution in astronomy has moved humans farther from the center of the universe. Now it appears that we may not even be made of the most common form of matter. Trace the changes in scientific thought about the central nature of Earth, the Sun, and our Galaxy on a cosmic scale. Explain how the notion that most of the universe is made of dark matter continues this “Copernican tradition.”
20. The anthropic principle suggests that in some sense we are observing a special kind of universe; if the universe were different, we could never have come to exist. Comment on how this fits with the Copernican tradition described in Exercise 29.19.
21. Penzias and Wilson’s discovery of the Cosmic Microwave Background (CMB) is a nice example of scientific *serendipity*—something that is found by chance but turns out to have a positive outcome. What were they looking for and what did they discover?
22. Construct a timeline for the universe and indicate when various significant events occurred, from the beginning of the expansion to the formation of the Sun to the appearance of humans on Earth.

Figuring for Yourself

23. Suppose the Hubble constant were not 22 but 33 km/s per million light-years. Then what would the critical density be?

24. Assume that the average galaxy contains 10^{11} M_{Sun} and that the average distance between galaxies is 10 million light-years. Calculate the average density of matter (mass per unit volume) in galaxies. What fraction is this of the critical density we calculated in the chapter?

25. The CMB contains roughly 400 million photons per m^3. The energy of each photon depends on its wavelength. Calculate the typical wavelength of a CMB photon. Hint: The CMB is blackbody radiation at a temperature of 2.73 K. According to Wien's law, the peak wave length in nanometers is given by $\lambda_{\text{max}} = \frac{3 \times 10^6}{T}$. Calculate the wavelength at which the CMB is a maximum and, to make the units consistent, convert this wavelength from nanometers to meters.

26. Following up on Exercise 29.25 calculate the energy of a typical photon. Assume for this approximate calculation that each photon has the wavelength calculated in Exercise 29.25. The energy of a photon is given by $E = \frac{hc}{\lambda}$, where h is Planck's constant and is equal to 6.626×10^{-34} J × s, c is the speed of light in m/s, and λ is the wavelength in m.

27. Continuing the thinking in Exercise 29.25 and Exercise 29.26, calculate the energy in a cubic meter of space, multiply the energy per photon calculated in Exercise 29.26 by the number of photons per cubic meter given above.

28. Continuing the thinking in the last three exercises, convert this energy to an equivalent in mass, use Einstein's equation $E = mc^2$. Hint: Divide the energy per m^3 calculated in Exercise 29.27 by the speed of light squared. Check your units; you should have an answer in kg/m^3. Now compare this answer with the critical density. Your answer should be several powers of 10 smaller than the critical density. In other words, you have found for yourself that the contribution of the CMB photons to the overall density of the universe is much, much smaller than the contribution made by stars and galaxies.

29. There is still some uncertainty in the Hubble constant. (a) Current estimates range from about 19.9 km/s per million light-years to 23 km/s per million light-years. Assume that the Hubble constant has been constant since the Big Bang. What is the possible range in the ages of the universe? Use the equation in the text, $T_0 = \frac{1}{H}$, and make sure you use consistent units. (b) Twenty years ago, estimates for the Hubble constant ranged from 50 to 100 km/s per Mps. What are the possible ages for the universe from those values? Can you rule out some of these possibilities on the basis of other evidence?

30. It is possible to derive the age of the universe given the value of the Hubble constant and the distance to a galaxy, again with the assumption that the value of the Hubble constant has not changed since the Big Bang. Consider a galaxy at a distance of 400 million light-years receding from us at a velocity, v. If the Hubble constant is 20 km/s per million light-years, what is its velocity? How long ago was that galaxy right next door to our own Galaxy if it has always been receding at its present rate? Express your answer in years. Since the universe began when all galaxies were very close together, this number is a rough estimate for the age of the universe.