Chapter One

Introduction and Cosmological Background

1.1 Preliminary Remarks

On large scales, the Universe is observed to be expanding. As it expands, galaxies separate from one another, and the density of matter (averaged over a large volume of space) decreases. If we imagine playing the cosmic movie in reverse and tracing this evolution backward in time, we would infer that there must have been an instant when the density of matter was infinite. This moment in time is the Big Bang, before which we cannot reliably extrapolate our history. But even before we get all the way back to the Big Bang, there must have been a time when stars like our Sun and galaxies like our Milky Way did not exist, because the Universe was denser than they are. If so, *how and when did the first stars and galaxies form?*

Primitive versions of this question were considered by humans for thousands of years, long before it was realized that the Universe is expanding. Religious and philosophical texts attempted to provide a sketch of the big picture from which people could derive the answer. In retrospect, these attempts appear heroic in view of the scarcity of scientific data about the Universe prior to the twentieth century. To appreciate the progress made over the past century, consider, for example, the biblical story of Genesis. The opening chapter of the Bible asserts the following sequence of events: first, the Universe was created, then light was separated from darkness, water was separated from the sky, continents were separated from water, vegetation appeared spontaneously, stars formed, life emerged, and finally humans appeared on the scene. Instead, the modern scientific order of events begins with the Big Bang, followed by an early period in which light (radiation) dominated and then a longer period in which matter was preeminent and led to the appearance of stars, planets, life on Earth, and eventually humans. Interestingly, the starting and end points of both versions are the same.

Cosmology is by now a mature empirical science. We are privileged to live in a time when the story of genesis (how the Universe started and developed) can be critically explored by direct observations. Because light takes a finite time to travel to us from distant sources, we can see images of the Universe when it was younger by looking deep into space through powerful telescopes.

Existing data sets include an image of the Universe when it was 400,000 years old (in the form of the cosmic microwave background in Figure 1.1), as well as images of individual galaxies when the Universe was older than a billion years. But there is a serious challenge: between these two epochs was a period when

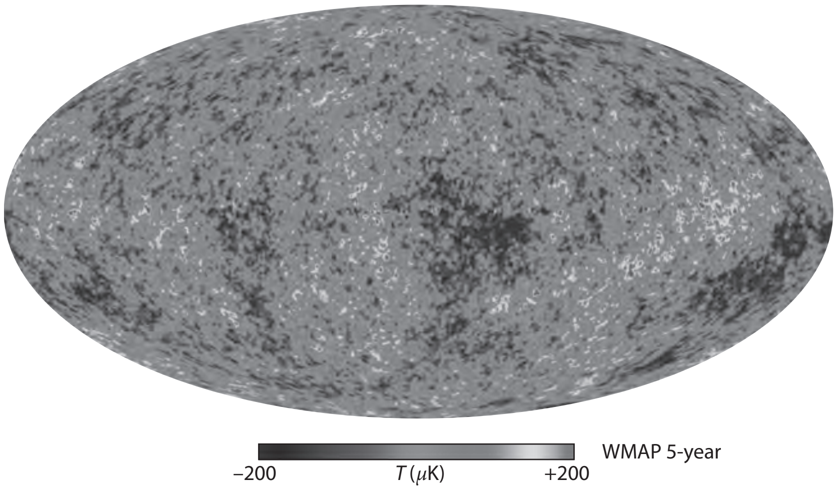

Figure 1.1 Image of the Universe when it first became transparent, 400,000 years after the Big Bang, taken over 5 years by the *Wilkinson Microwave Anisotropy Probe* (WMAP) satellite (see *Color Plate 1* for a color version of this figure). Slight density inhomogeneities at the level of one part in $\sim 10^5$ in the otherwise uniform early Universe imprinted hot and cold spots in the temperature map of the cosmic microwave background on the sky. The fluctuations are shown in units of microkelvins, and the unperturbed temperature is 2.725 K. The same primordial inhomogeneities seeded the large-scale structure in the present-day Universe. The existence of background anisotropies was predicted in a number of theoretical papers three decades before the technology for taking this image became available. Courtesy of NASA and the WMAP Science Team.

the Universe was dark, stars had not yet formed, and the cosmic microwave background no longer traced the distribution of matter. And this is precisely the most interesting period, when the primordial soup evolved into the rich zoo of objects we now see. *How can astronomers see this dark yet crucial time?*

The situation is similar to having a photo album of a person that begins with the first ultrasound image of him or her as an unborn baby and then skips to some additional photos of his or her years as teenager and adult. The later photos do not simply show a scaled-up version of the first image. We are currently searching for the missing pages of the cosmic photo album that will tell us how the Universe evolved during its infancy to eventually make galaxies like our own Milky Way.

Observers are moving ahead along several fronts. The first involves the construction of large infrared telescopes on the ground and in space that will provide us with new (although rather expensive!) photos of galaxies in the Universe at intermediate ages. Current plans include ground-based telescopes 24–42 m in diameter and NASA's successor to the Hubble Space Telescope, the James Webb Space Telescope (JWST). In addition, several observational groups around the globe are constructing radio arrays that will be capable of mapping the three-dimensional distribution of cosmic hydrogen left over from the Big Bang in the infant Universe. These arrays are aiming to detect the long-wavelength (redshifted 21-cm) radio emission from hydrogen atoms. Coincidentally, this long wavelength (or low frequency) overlaps the band used for radio and television broadcasting, and so these telescopes include arrays of regular radio antennas that one can find in electronics stores. These antennas will reveal how the clumpy distribution of neutral hydrogen evolved with cosmic time. By the time the Universe was a few hundreds of millions of years old, the hydrogen distribution had been punched with holes and resembled Swiss cheese. These holes were created by the ultraviolet radiation from the first galaxies and black holes, which ionized the cosmic hydrogen in their vicinity.

Theoretical research has focused in recent years on predicting the signals expected from the telescopes described and on providing motivation for these ambitious observational projects.

All these predictions are generated in the context of the modern cosmological paradigm, which turns the Big Bang model into a quantitative tool for understanding our Universe. In the remainder of this chapter, we briefly describe the essential aspects of this paradigm for understanding the formation of the first galaxies in the Universe.

1.2 Standard Cosmological Model

1.2.1 Cosmic Perspective

In 1915 Einstein formulated the general theory of relativity. He was inspired by the fact that all objects follow the same trajectories under the influence of gravity (the so-called equivalence principle, which by now has been tested to better than one part in a trillion), and realized that this would be a natural result if space–time is curved under the influence of matter. He wrote an equation describing how the distribution of matter (on one side of his equation) determines the curvature of space–time (on the other side of his equation). Einstein then applied his equation to describe the global dynamics of the Universe.

There were no computers available in 1915, and Einstein's equations for the Universe were particularly difficult to solve in the most general case. To get around this obstacle Einstein considered the simplest possible Universe, one that is homogeneous and isotropic. Homogeneity means uniform conditions everywhere (at any given time), and isotropy means the same conditions in all directions seen from one vantage point. The combination of these two simplifying assumptions is known as the *cosmological principle*.

The Universe can be homogeneous but not isotropic: for example, the expansion rate could vary with direction. It can also be isotropic and not homogeneous: for example, we could be at the center of a spherically symmetric mass distribution. But if it is isotropic around *every* point, then it must also be homogeneous.

Under the simplifying assumptions associated with the cosmological principle, Einstein and his contemporaries were able to solve the equations. They were looking for their "lost keys" (solutions) under a convenient "lamppost" (simplifying assumptions), but the real Universe is not bound by any contract to be the simplest that we can imagine. In fact, it is truly remarkable in the first place that we dare describe the conditions across vast regions of space based on the blueprint of the laws of physics that describe the conditions here on Earth. Our daily life teaches us too often that we fail to appreciate complexity, and that an elegant model for reality is often too idealized for describing the truth (along the lines of approximating a cow as a spherical object).

In 1915 Einstein had the wrong notion of the Universe; at the time people associated the Universe with the Milky Way galaxy and regarded all the "spiral nebulae," which we now know are distant galaxies, as constituents of our own Milky Way galaxy. Because the Milky Way is not expanding, Einstein attempted to reproduce a static universe with his equations. This turned out to be possible only after he added a cosmological constant, whose negative gravity would exactly counteract that of matter. However, Einstein later realized that this solution is unstable: a slight enhancement in density would make the density grow even further. As it turns out, there are no stable static solutions to Einstein's equations for a homogeneous and isotropic Universe. The Universe must be either expanding or contracting. Less than a decade later, Edwin Hubble discovered that the nebulae previously considered to be constituents of the Milky Way galaxy are receding from us at a speed v that is proportional to their distance r, namely, $v = H_0 r$, where H_0 is a spatial constant (which can evolve with time), commonly termed the *Hubble constant*.[i] Hubble's data indicated that the Universe is expanding. (Hubble also resolved individual bright stars in these nebulae, unambiguously determining their nature and their vast distances from the Milky Way.)

Einstein was remarkably successful in asserting the cosmological principle. As it turns out, our latest data indicate that the real Universe is homogeneous and isotropic on the largest observable scales to within one part in 10^5. In particular, isotropy is well established for the distribution of faint radio sources, optical galaxies, the X-ray background, and most important, the cosmic microwave background (CMB). The constraints on homogeneity are less strict, but a cosmological model in which the Universe is isotropic and significantly inhomogeneous in spherical shells around our special location is also excluded based on surveys of galaxies and quasars. Fortuitously, Einstein's simplifying assumptions turned out to be extremely accurate in describing reality: *the keys were indeed lying next to the lamppost*. Our Universe happens to be the simplest we could have imagined, for which Einstein's equations can easily be solved.

[i]The redshift data examined by Hubble was mostly collected by Vesto Slipher a decade earlier and only partly by Hubble's assistant, Milton L. Humason. The linear local relation between redshift and distance (based on Hubble and Humason's data) was first formulated by Georges Lemaître in 1927, 2 years prior to the observational paper written by Hubble and Humason.

Why was the Universe prepared to be in this special state? Cosmologists were able to go one step further and demonstrated that an early phase transition, called *cosmic inflation*—during which the expansion of the Universe accelerated exponentially—could have naturally produced the conditions postulated by the cosmological principle (although other explanations also may create such conditions). One is left to wonder whether the existence of inflation is just a fortunate consequence of the fundamental laws of nature, or whether perhaps the special conditions of the specific region of space–time we inhabit were selected out of many random possibilities elsewhere by the prerequisite that they allow our existence. The opinions of cosmologists on this question are split.

1.2.2 Origin of Structure

Hubble's discovery of the expansion of the Universe has immediate implications for the past and future of the Universe. If we reverse in our mind the expansion history back in time, we realize that the Universe must have been denser in its past. In fact, there must have been a point in time where the matter density was infinite, at the moment of the so-called Big Bang. Indeed, we do detect relics from a hotter, denser phase of the Universe in the form of light elements (such as deuterium, helium, and lithium) as well as the CMB. At early times, this radiation coupled extremely well to the cosmic gas and produced a spectrum known as a *blackbody*, a form predicted a century ago to characterize matter and radiation in equilibrium. The CMB provides the best example of a blackbody spectrum we have.

To get a rough estimate of when the Big Bang occurred, we may simply divide the distance of all galaxies by their recession velocity. This calculation gives a unique answer, $\sim r/v \sim 1/H_0$, that is independent of distance.[ii] The latest measurements of the Hubble constant give a value of $H_0 \approx 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, which implies a current age for the Universe $1/H_0$ of 14 billion years (or 5×10^{17} seconds).

The second implication concerns our future. A fortunate feature of a spherically symmetric Universe is that when considering a sphere of matter in it, we are allowed to ignore the gravitational influence of everything outside this sphere. If we empty the sphere and consider a test particle on the boundary of an empty void embedded in a uniform Universe, the particle will experience no net gravitational acceleration. This result, known as *Birkhoff's theorem*, is reminiscent of Newton's "iron sphere theorem." It allows us to solve the equations of motion for matter on the boundary of the sphere through a local analysis without worrying about the rest of the Universe. Therefore, if the sphere has exactly the same conditions as the rest of the Universe, we may deduce the global expansion history of the Universe by examining its behavior. If the sphere is

[ii] Although this is an approximate estimate, it turns out to be within a few percent of the true age of our Universe owing to a coincidence. The cosmic expansion at first decelerated and then accelerated, with the two almost canceling each other at the present time, giving the same age as if the expansion were at a constant speed (as would be strictly true only in an empty Universe).

slightly denser than the mean, we will infer how its density contrast will evolve relative to the background Universe.

For the moment, let us ignore the energy density of the vacuum (which is always a good approximation at sufficiently early cosmic times, when matter was denser). Then, the equation describing the motion of a spherical shell of matter is identical with the equation of motion of a rocket launched from the surface of the earth. The rocket will escape to infinity if its kinetic energy exceeds its gravitational binding energy, making its total energy positive. However, if its total energy is negative, the rocket will reach a maximum height and then fall back. To deduce the future evolution of the Universe, we need to examine the energy of a spherical shell of matter relative to the origin. With a uniform density ρ, a spherical shell of radius r has a total mass $M = \rho \times (4\pi r^3/3)$ enclosed within it. Its energy per unit mass is the sum of the kinetic energy due to its expansion speed $v = Hr$, $(1/2)v^2$, and its potential gravitational energy, $-GM/r$ (where G is Newton's constant), namely, $E = v^2/2 - GM/r$. By substituting the preceding relations for v and M, we can easily show that $E = (1/2)v^2(1 - \Omega)$, where $\Omega = \rho/\rho_c$, and $\rho_c = 3H^2/8\pi G$ is defined as the *critical density*. We therefore find that there are three possible scenarios for the cosmic expansion. The Universe has either *(i)* $\Omega > 1$, making it gravitationally bound with $E < 0$—*such a "closed Universe" will turn around and end up collapsing toward a "big crunch"*; *(ii)* $\Omega < 1$, making it gravitationally unbound with $E > 0$—*such an "open Universe" will expand forever*; or the borderline case, *(iii)* $\Omega = 1$, making the Universe marginally bound or "flat" with $E = 0$.

Einstein's equations relate the geometry of space to its matter content through the value of Ω: an open Universe has the geometry of a saddle with a negative spatial curvature, a closed Universe has the geometry of a spherical globe with a positive curvature, and a flat Universe has a flat geometry with no curvature. Our observable section of the Universe appears to be flat.

Now we are in a position to understand how objects like the Milky Way galaxy formed out of small density inhomogeneities that are amplified by gravity.

Let us consider for simplicity the background of a marginally bound (flat) Universe dominated by matter. In such a background, only a slight enhancement in density is required to exceed the critical density ρ_c. Because of Birkhoff's theorem, a spherical region denser than the mean will behave as if it is part of a closed Universe and will increase its density contrast with time, while an underdense spherical region will behave as if it is part of an open Universe and will appear more vacant with time relative to the background, as illustrated in Figure 1.2. Starting with slight density enhancements that bring them above the critical value, ρ_c, the overdense regions will initially expand, reach a maximum radius, and then collapse on themselves (like the trajectory of a rocket launched straight up, away from the center of the earth). An initially slightly inhomogeneous Universe will end up clumpy, with collapsed objects forming out of overdense regions. The material to make the objects is drained out of the intervening underdense regions, which end up as voids.

The Universe we live in started with primordial density perturbations of a fractional amplitude $\sim 10^{-5}$ when the cosmic microwave background last scattered. The overdensities were amplified at late times (once matter dominated

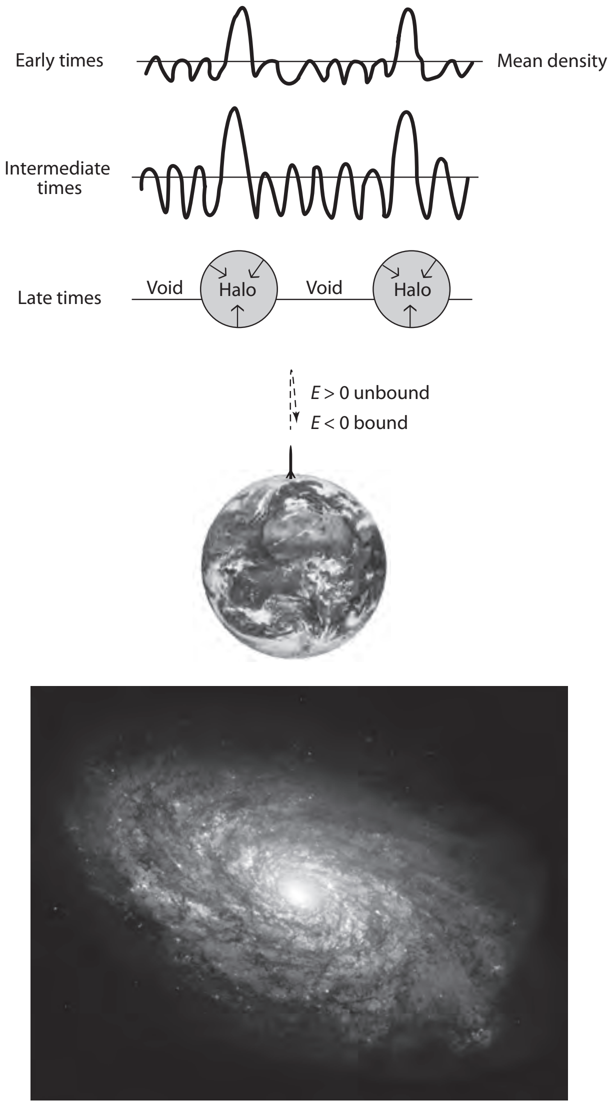

Figure 1.2 *Top:* Schematic illustration of the growth of perturbations to collapsed halos through gravitational instability. The overdense regions initially expand, reach a maximum size, and then turn around and collapse to form gravitationally bound halos if their density exceeds a critical threshold (see §3.1). The material that makes the halos originated in the voids that separate them. *Middle:* A simple model for the collapse of a spherical region. The dynamical fate of a rocket launched from the surface of the earth depends on the sign of its energy per unit mass, $E = (1/2)v^2 - GM_{\oplus}/r$. The behavior of a spherical shell of matter on the boundary of an overdense region (embedded in a homogeneous and isotropic Universe) can be analyzed in a similar fashion. *Bottom:* A collapsing region may end up as a galaxy, like NGC 4414, shown here. (Courtesy of NASA and ESA.) The halo gas cools and condenses to a compact disk surrounded by an extended dark matter halo.

the cosmic mass budget) up to values close to unity and collapsed to make objects, first on small scales. We have not yet seen the first small galaxies that started the process that eventually led to the formation of big galaxies like the Milky Way. The search for the first galaxies is a search for our origins and the main subject of this book.

Beyond being uniform, the early Universe was simple in one additional way: the process of *Big Bang nucleosynthesis* produced the first atomic nuclei, but these were mostly hydrogen and helium (with $\sim$93% of those atoms in the form of hydrogen). However, life as we know it on planet Earth requires water. The water molecule includes oxygen, an element that was not made in the Big Bang and did not exist until the first stars had formed. Therefore, our form of life could not have existed in the first hundred million years after the Big Bang, before the first stars had formed. There is also no guarantee that life will persist in the distant future.

1.2.3 Geometry of Space

The history and fate of our Universe are thus tied inexorably to its contents—be it matter, dark energy, or something even more exotic. However, measuring the average density of the Universe is extraordinarily difficult. Fortunately, Einstein's equations show that the contents of the Universe are also tied to its geometry—so measuring the latter would indirectly constrain its components.

How can we tell the difference between the flat surface of a book and the curved surface of a balloon? A simple way is to draw a triangle of straight lines between three points on those surfaces and measure the sum of the three angles of the triangle. The Greek mathematician Euclid demonstrated that the sum of these angles must be 180° (or π radians) on a flat surface. Twenty-one centuries later, the German mathematician Bernhard Riemann extended the field of geometry to curved spaces, which played an important role in the development of Einstein's general theory of relativity. For a triangle drawn on a positively curved surface, like that of a balloon, the sum of the angles is larger than 180°. (This can easily be figured out by examining a globe and noticing that any line connecting one of the poles to the equator opens an angle of 90° relative to the equator. Adding the third angle in any triangle stretched between the pole and the equator would surely result in a total of more than 180°.) According to Einstein's equations, the geometry of the Universe is dictated by its matter content; in particular, the Universe is flat only if the total Ω equals unity. *Is it possible to draw a triangle across the entire Universe and measure its geometry?*

Remarkably, the answer is yes. At the end of the twentieth century cosmologists were able to perform this experiment by adopting a simple yardstick provided by the early Universe. The familiar experience of dropping a stone in the middle of a pond results in a circular wave crest that propagates outward. Similarly, perturbing the smooth Universe at a single point at the Big Bang would have resulted in the propogation of a spherical sound wave outward from that point. The wave would have traveled at the speed of sound, which was of the order of the speed of light c (or, more precisely, $c/\sqrt{3}$) early on when radiation dominated the cosmic mass budget. At any given time, all the points extending to the distance traveled by the wave are affected by the original pointlike perturbation. The conditions outside this "sound horizon" will remain uncorrelated with the central point, because acoustic information has not yet been able to reach them. The temperature fluctuations of the CMB trace the simple sum of many such pointlike perturbations that were generated in the Big Bang. The patterns they delineate will therefore show a characteristic correlation scale, corresponding to the sound horizon at the time when the CMB was produced, 400,000 years after the Big Bang. By measuring the apparent angular scale of this "standard ruler" on the sky, known as the acoustic peak in the CMB, and comparing it with theory, experimental cosmologists inferred from the simple geometry of triangles that the Universe is flat (or at least very close to it).

The inferred flatness may be a natural consequence of the early period of cosmic inflation during which any initial curvature was flattened. Indeed, a small patch of a fixed size (representing our current observable region in the cosmological context) on the surface of a vastly inflated balloon would appear nearly flat. The sum of the angles on a nonexpanding triangle placed on this patch would get arbitrarily close to 180° as the balloon inflated.

Even though we now know that our Universe is very close to being flat, this flatness only constrains the cumulative energy density in the Universe; it tells us very little about how that energy is distributed among the different components, such as baryons, other forms of matter, and dark energy. We must probe our Universe in other ways to learn about this distribution.

1.2.4 Observing Our Past: Cosmic Archaeology

Our Universe is the simplest possible on two counts: it satisfies the cosmological principle, and it has a flat geometry. The mathematical description of an expanding, homogeneous, and isotropic Universe with a flat geometry is straightforward. We can imagine filling up space with clocks that are all synchronized. At any given snapshot in time the physical conditions (density, temperature) are the same everywhere. But as time goes on, the spatial separation between the clocks will increase. The stretching of space can be described by a time-dependent scale factor, $a(t)$. A separation measured at time t_1 as $r(t_1)$ will appear at time t_2 to have a length $r(t_2) = r(t_1)[a(t_2)/a(t_1)]$.

A natural question to ask is whether our human bodies or even the solar system is also expanding as the Universe expands. The answer is no, because these systems are held together by forces whose strength far exceeds the cosmic force. The mean density of the Universe today, $\bar{\rho}$, is 29 orders of magnitude smaller than the density of our body. Not only are the electromagnetic forces that keep the atoms in our body together far greater than the force of gravity, but even the gravitational self-force of our body on itself overwhelms the cosmic influence. Only on very large scales does the cosmic gravitational force dominate the scene. This also implies that we cannot observe the cosmic expansion with a local laboratory experiment; to notice the expansion we need to observe sources spread over the vast scales of millions of light-years.

The space–time of an expanding homogeneous and isotropic flat Universe can be described very simply. Because of the cosmological principle, we can establish a unique time coordinate throughout space by distributing clocks that are all synchronized throughout the Universe, so that each clock will measure the same time t since the Big Bang. The space–time (four–dimensional) line element ds, commonly defined to vanish for a photon, is described by the Friedmann-Robertson-Walker (FRW) metric,

$$ds^2 = c^2 dt^2 - d\ell^2, \tag{1.1}$$

where c is the speed of light and $d\ell$ is the spatial line element. The cosmic expansion can be incorporated through a scale factor $a(t)$ that multiplies the fixed (x, y, z) coordinates tagging the clocks, which are themselves "comoving" with the cosmic expansion. For a flat space,

$$d\ell^2 = a(t)^2(dx^2 + dy^2 + dz^2) = a^2(t)(dR^2 + R^2 d\Omega), \tag{1.2}$$

where $d\Omega = d\theta^2 + \sin^2\theta\, d\phi^2$, (R, θ, ϕ) are the comoving spherical coordinates centered on the observer, and $(x, y, z) = (R\cos\theta, R\sin\theta\cos\phi, R\sin\theta\sin\phi)$. Throughout this book, we quote distances in these comoving units—as opposed to their time-varying *proper* equivalents—unless otherwise specified.

A source located at a separation $r = a(t)R$ from us will move at a velocity $v = dr/dt = \dot{a}R = (\dot{a}/a)r$, where $\dot{a} = da/dt$. Here r is a time-independent tag denoting the present-day distance of the source (when $a(t) \equiv 1$). Defining $H = \dot{a}/a$, which is constant in space, we recover the Hubble expansion law $v = Hr$.

Edwin Hubble measured the expansion of the Universe using the Doppler effect. We are all familiar with the same effect for sound waves: when a moving car sounds its horn, the pitch (frequency) we hear is different when the car is approaching us than when it is receding from us. Similarly, the wavelength of light depends on the velocity of the source relative to us. As the Universe expands, a light source will move away from us, and its Doppler effect will change with time. The Doppler formula for a nearby source of light (with a recession speed much smaller than the speed of light) gives

$$\frac{\Delta\nu}{\nu} \approx -\frac{\Delta v}{c} = -\left(\frac{\dot{a}}{a}\right)\left(\frac{r}{c}\right) = -\frac{(\dot{a}\Delta t)}{a} = -\frac{\Delta a}{a}, \tag{1.3}$$

with the solution $\nu \propto a^{-1}$. Correspondingly, the wavelength scales as $\lambda = (c/\nu) \propto a$. We could have anticipated this outcome, since a wavelength can be used as a measure of distance and should therefore be stretched as the Universe expands. This relation holds also for the de Broglie wavelength, $\lambda_{\rm dB} = (h/p) \propto a$, characterizing the quantum-mechanical wavefunction of a massive particle with momentum p (where h is Planck's constant). Consequently, the kinetic energy of a nonrelativistic particle scales as $(p^2/2m_p) \propto a^{-2}$. Thus, in the absence of heat exchange with other systems, the temperature of a gas of nonrelativistic protons and electrons will cool faster ($\propto a^{-2}$) than the CMB temperature ($h\nu \propto a^{-1}$) as the Universe expands and a increases. The redshift z is defined through the factor $(1+z)$ by which the photon wavelength was stretched (or its frequency reduced) between its emission and observation times. If we define $a = 1$ today, then $a = 1/(1 + z)$ at earlier times. Higher redshifts correspond to a higher recession speed of the source relative to us (that ultimately approaches the speed of light when the redshift goes to infinity), which in turn implies a larger distance (that ultimately approaces our horizon, which is the distance traveled by light since the Big Bang) and an earlier emission time of the source for the photons to reach us today.

We see high-redshift sources as they looked at early cosmic times. Observational cosmology is like archaeology—the deeper we look into space, the more ancient the clues about our history are (see Figure 1.3).[iii] But there is a limit to how far back we can see: we can image the Universe only if it is transparent. Earlier than 400,000 years after the Big Bang, the cosmic gas was sufficiently hot to be fully ionized, and the Universe was opaque owing to scattering by the dense fog of free electrons that filled it. Thus, telescopes cannot be used to image the infant Universe at earlier times (at redshifts $> 10^3$). The earliest possible image of the Universe can be seen in the cosmic microwave background, the thermal radiation left over from the transition to transparency (Figure 1.1). The first galaxies are believed to have formed long after that.

The expansion history of the Universe is captured by the scale factor $a(t)$. We can write a simple equation for the evolution of $a(t)$ based on the behavior of a small region of space. For that purpose we need to incorporate the fact that in Einstein's theory of gravity, not only does mass density ρ gravitate but pressure p does as well. In a homogeneous and isotropic Universe, the quantity $\rho_{\rm grav} = (\rho + 3p/c^2)$ plays the role of the gravitating mass density ρ of Newtonian gravity. There are several examples to consider. For a radiation fluid,[iv] $p_{\rm rad}/c^2 = (1/3)\rho_{\rm rad}$, which implies that $\rho_{\rm grav} = 2\rho_{\rm rad}$.

However, if the vacuum has a nonzero energy density that is constant in space and time, the *cosmological constant*, then the pressure of the vacuum is negative, because by opening up a new volume increment ΔV one *gains* an energy $\rho_{\rm vac}c^2\Delta V$ instead of losing it, as is the case for normal fluids that expand into more space. In thermodynamics, pressure is derived from the deficit in energy per unit of new volume, which in this case gives $p_{\rm vac}/c^2 = -\rho_{\rm vac}$. This relation in turn leads to another reversal of signs, $\rho_{\rm grav} = (\rho_{\rm vac} + 3p_{\rm vac}/c^2) = -2\rho_{\rm vac}$, which may be interpreted as repulsive gravity! This surprising result gives rise to the phenomenon of accelerated cosmic expansion, which characterized the early period of cosmic inflation as well as the latest 6 billion years of cosmic history.

[iii]Cosmology and archaeology share another similarity: both are *observational*, rather than *experimental*, sciences. Thus, we are forced to interpret the complicated physics of actual systems rather than design elegant experiments that can answer targeted questions. Although simplified models can be built in the laboratory (or even inside computers), the primary challenge of cosmology is figuring out how to extract useful information from real and complex systems that cannot be artificially altered.

[iv]The momentum of each photon is $1/c$ of its energy. The pressure is defined as the momentum flux along one dimension out of three and is therefore given by $(1/3)\rho_{\rm rad}c^2$, where $\rho_{\rm rad}$ is the equivalent mass density of the radiation.

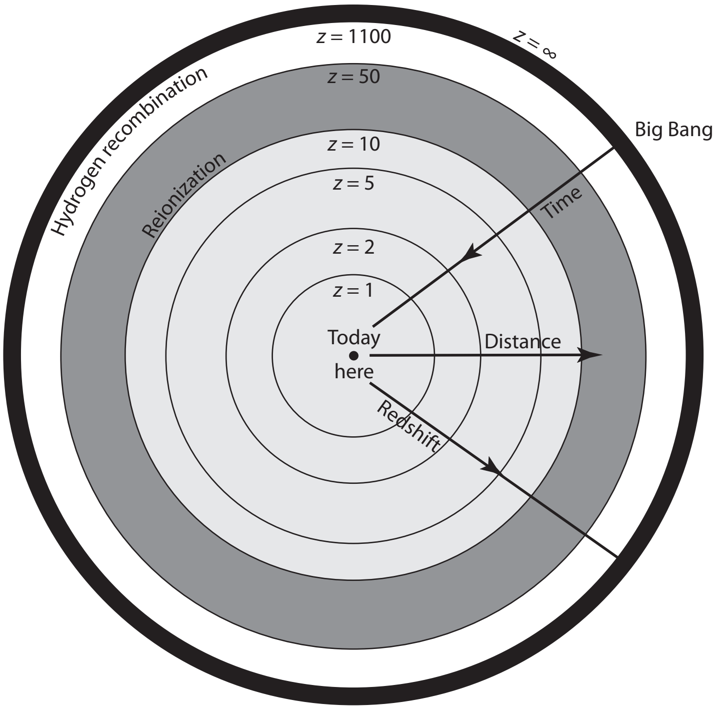

Figure 1.3 Cosmic archaeology of the observable volume of the Universe, in comoving coordinates (which factor out the cosmic expansion). The outermost observable boundary ($z = \infty$) marks the comoving distance that light has traveled since the Big Bang. Future observatories aim to map most of the observable volume of our Universe and to improve dramatically the statistical information we have about the density fluctuations within it. Existing data on the CMB probe mainly a very thin shell at the hydrogen recombination epoch ($z \sim 10^3$, beyond which the Universe is opaque), and current large-scale galaxy surveys map only a small region near us at the center of the diagram. The formation epoch of the first galaxies that culminated with hydrogen reionization at a redshift $z \sim 10$ is shaded dark gray. Note that the comoving volume out to any of these redshifts scales as the distance cubed.

As the Universe expands and the scale factor increases, the matter mass density declines inversely with volume, $\rho_{\rm matter} \propto a^{-3}$, whereas the radiation energy density (which includes the CMB and three species of relativistic neutrinos) decreases as $\rho_{\rm rad}c^2 \propto a^{-4}$, because not only is the density of photons diluted as a^{-3}, but the energy per photon $h\nu = hc/\lambda$ (where h is Planck's constant) declines as a^{-1}. Today $\rho_{\rm matter}$ is larger than $\rho_{\rm rad}$ (assuming massless neutrinos) by a factor of $\sim$3,300, but at $(1 + z) \sim 3{,}300$ the two were equal, and at even higher redshifts the radiation dominated. Since a stable vacuum does not get diluted with cosmic expansion, the present-day $\rho_{\rm vac}$ remained a constant and dominated over $\rho_{\rm matter}$ and $\rho_{\rm rad}$ only at late times (whereas the unstable "false vacuum" that dominated during inflation decayed when inflation ended).

In this book, we will primarily be concerned with the *cosmic dawn*, or the era in which the first galaxies formed at $z \sim 6$–30. At these early times, the cosmological constant was very small compared with the matter densities and can generally be ignored.

1.3 Milestones in Cosmic Evolution

The gravitating mass, $M_{\rm grav} = \rho_{\rm grav} V$, enclosed by a spherical shell of radius $r(t) = a(t)$ and volume $V = (4\pi/3)a^3$, induces an acceleration

$$\frac{d^2 a}{dt^2} = -\frac{G M_{\rm grav}}{a^2}. \tag{1.4}$$

Since $\rho_{\rm grav} = \rho + 3p/c^2$, we need to know how pressure evolves with the expansion factor $a(t)$. We obtain this information from the thermodynamic relation mentioned previously between the change in the internal energy $d(\rho c^2 V)$ and the $p\,dV$ work done by the pressure, $d(\rho c^2 V) = -p\,dV$. This relation implies $-3pa\dot{a}/c^2 = a^2\dot{\rho} + 3\rho a\dot{a}$, where an overdot denotes a time derivative. Multiplying equation (1.4) by $\dot{a}$ and making use of this relation yields our familiar result

$$E = \frac{1}{2}\dot{a}^2 - \frac{GM}{a}, \tag{1.5}$$

where E is a constant of integration, and $M \equiv \rho V$. As discussed before, the spherical shell will expand forever (being gravitationally unbound) if $E \geq 0$ but will eventually collapse (being gravitationally bound) if $E < 0$. Making use of the Hubble parameter, $H = \dot{a}/a$, we can rewrite equation (1.5) as

$$\frac{E}{\dot{a}^2/2} = 1 - \Omega, \tag{1.6}$$

where $\Omega = \rho/\rho_c$, with

$$\rho_c = \frac{3H^2}{8\pi G} = 9.2 \times 10^{-30}\frac{\rm g}{\rm cm^3}\left(\frac{H}{70\,{\rm km\,s^{-1}Mpc^{-1}}}\right)^2. \tag{1.7}$$

If we denote the present contributions to Ω from *matter* (including cold dark matter as well as a contribution Ω_b from ordinary matter of protons and neutrons, or "baryons"), *vacuum density* (cosmological constant), and *radiation*, with Ω_m, Ω_Λ, and Ω_r, respectively, a flat universe with $E = 0$ satisfies

$$\frac{H(t)}{H_0} = \left[\frac{\Omega_m}{a^3} + \Omega_\Lambda + \frac{\Omega_r}{a^4}\right]^{1/2}, \tag{1.8}$$

where we define H_0 and $\Omega_0 = (\Omega_m + \Omega_\Lambda + \Omega_r) = 1$ to be the present-day values of H and Ω, respectively.

In the particularly simple case of a flat Universe, we find that if matter dominates (i.e., $\Omega_0 = 1$), then $a \propto t^{2/3}$; if radiation dominates, then $a \propto t^{1/2}$;

and if the vacuum density dominates, then $a \propto \exp\{H_{\rm vac}t\}$, where $H_{\rm vac} = (8\pi G\rho_{\rm vac}/3)^{1/2}$ is a constant. After inflation ended, the mass density of our Universe, ρ, was at first dominated by radiation at redshifts $z > 3{,}300$, by matter at $0.3 < z < 3{,}300$, and finally by the vacuum at $z < 0.3$. The vacuum had already started to dominate $\rho_{\rm grav}$ at $z < 0.7$, or 6 billion years ago. Figure 1.6 illustrates the mass budget in the present-day Universe and during the epoch when the first galaxies formed.

The preceding results for $a(t)$ have two interesting implications. First, we can calculate the relationship between the time since the Big Bang and redshift, since $a = (1+z)^{-1}$. For example, during the matter-dominated era ($1 < z < 10^3$, with the low-z end set by the condition $[1+z] \gg [\Omega_\Lambda/\Omega_m]^{1/3}$),

$$t \approx \frac{2}{3H_0\Omega_m{}^{1/2}(1+z)^{3/2}} = \frac{0.95\times10^9\ \rm yr}{[(1+z)/7]^{3/2}}. \tag{1.9}$$

In this same regime, where $\Omega_m \approx 1$, $H \approx 2/(3t)$, and $a = (1+z)^{-1} \approx (3H_0\sqrt{\Omega_m}/2)^{2/3}t^{2/3}$.

Second, we note the remarkable exponential expansion for a vacuum-dominated phase. This accelerated expansion serves an important purpose in explaining a few puzzling features of our Universe. We have already noticed that our Universe was prepared in a very special initial state: nearly isotropic and homogeneous, with Ω close to unity and a flat geometry. In fact, it took the CMB photons nearly the entire age of the Universe to travel toward us. Therefore, it should take them twice as long to bridge their points of origin on opposite sides of the sky. *How is it possible then that the conditions of the Universe (as reflected in the nearly uniform CMB temperature) were prepared to be the same in regions that were never in causal contact before?* Such a degree of organization is highly unlikely to occur at random. If we receive our clothes ironed and folded neatly, we know that there must have a been a process that caused this to happen. Cosmologists have identified an analogous "ironing process" in the form of cosmic inflation. This process is associated with an early period during which the Universe was dominated temporarily by the mass density of an elevated vacuum state and experienced exponential expansion by at least ~60 *e*-folds. This vast expansion "ironed out" any initial curvature of our environment and generated a flat geometry and nearly uniform conditions across a region far greater than our current horizon. After the elevated vacuum state decayed, the Universe became dominated by radiation.

The early epoch of inflation was important not just in producing the global properties of the Universe but also in generating the inhomogeneities that seeded the formation of galaxies within it. The vacuum energy density that had driven inflation encountered quantum-mechanical fluctuations. After the perturbations were stretched beyond the horizon of the infant Universe (which today would have occupied a size no bigger than a human hand), they materialized as perturbations in the mass density of radiation and matter. The last perturbations to leave the horizon during inflation eventually reentered after inflation ended (when the scale factor grew more slowly than ct). It is tantalizing to contemplate the notion that galaxies, which represent massive classical objects with $\sim 10^{67}$ atoms in today's Universe, might have originated from subatomic quantum-mechanical fluctuations at early times.

After inflation, an unknown process, called "baryogenesis" or "leptogenesis," generated an excess of particles (baryons and leptons) over antiparticles.[v] As the Universe cooled to a temperature of hundreds of millions of electron-volts (where $1\,\text{MeV}/k_B = 1.1604 \times 10^{10}\,\text{K}$), protons and neutrons condensed out of the primordial quark–gluon plasma through the so-called quantum chromodynamics (QCD) phase transition. At about one second after the Big Bang, the temperature declined to $\sim 1\,\text{MeV}$, and the weakly interacting neutrinos decoupled. Shortly afterward, the abundance of neutrons relative to protons froze, and electrons and positrons annihilated one another. In the next few minutes, nuclear fusion reactions produced light elements more massive than hydrogen, such as deuterium, helium, and lithium, in abundances that match those observed today in regions where gas has not been processed subsequently through stellar interiors. Although the transition to matter domination occurred at a redshift $z \sim 3{,}300$, the Universe remained hot enough for the gas to be ionized, and electron–photon scattering effectively coupled ordinary matter and radiation. At $z \sim 1{,}100$ the temperature dipped below $\sim 3{,}000\,\text{K}$, and free electrons recombined with protons to form neutral hydrogen atoms. As soon as the dense fog of free electrons was depleted, the Universe became transparent to the relic radiation, which is observed at present as the CMB. These milestones of the thermal history are depicted in Figure 1.4.

The Big Bang is the only known event in our past history in which particles interacted with center-of-mass energies approaching the so-called Planck scale[vi] $[(hc^5/G)^{1/2} \sim 10^{19}\,\text{GeV}]$, at which quantum mechanics and gravity are expected to be unified. Unfortunately, the exponential expansion of the Universe during inflation erased memory of earlier cosmic epochs, such as the Planck time.

1.3.1 Luminosity and Angular-Diameter Distances

When we look at our image reflected off a mirror at a distance of 1 m, we see the way we looked 6 nanoseconds ago, the time it took light to travel to the mirror and back. If the mirror is spaced $10^{19}\,\text{cm} = 3\,\text{pc}$ away, we will see the way we looked 21 years ago. Light propagates at a finite speed, so by observing distant regions, we are able to see what the Universe looked like in the past, a light-travel time ago (see Figure 1.3). The statistical homogeneity of the Universe on large scales guarantees that what we see far away is a fair statistical representation of the conditions that were present in our region of the Universe a long time ago.

This fortunate situation makes cosmology an empirical science. We do not need to guess how the Universe evolved. By using telescopes we can simply see

[v] The origin of the asymmetry in the cosmic abundance of matter over antimatter is still an unresolved puzzle.

[vi] The Planck energy scale is obtained by equating the quantum-mechanical wavelength of a relativistic particle with energy E, namely, hc/E, to its "black hole" radius, $\sim GE/c^4$, and solving for E.

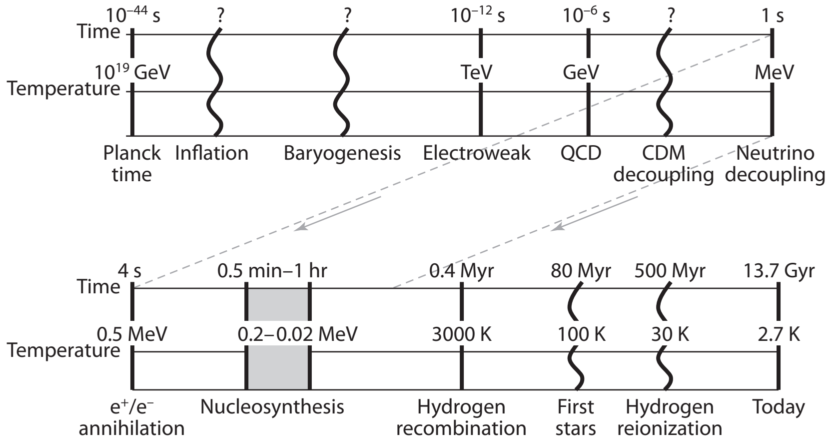

Figure 1.4 Following inflation, the Universe went through several other milestones that left a detectable record. These include baryogenesis (which resulted in the observed asymmetry between matter and antimatter), the electroweak phase transition (during which the symmetry between electromagnetic and weak interactions was broken), the QCD phase transition (during which protons and neutrons nucleated out of a soup of quarks and gluons), the dark matter decoupling epoch (during which the dark matter decoupled thermally from the cosmic plasma), neutrino decoupling, electron–positron annihilation, light-element nucleosynthesis (during which helium, deuterium, and lithium were synthesized), and hydrogen recombination. The cosmic time and CMB temperature of the various milestones are marked. Wavy lines and question marks indicate milestones with uncertain properties. The signatures that the same milestones left in the Universe are used to constrain its parameters.

the way distant regions appeared at earlier cosmic times. Since a greater distance means a fainter flux from a source of a fixed luminosity, the observation of the earliest sources of light requires the development of sensitive instruments and poses technological challenges to observers.

How faint will the earliest galaxies appear to our telescopes? In an expanding Universe there is some ambiguity as to which “distance” is most relevant. For example, the framework we described earlier—in which the clocks are synchronized relative to the Big Bang—is not appropriate for observations, because light has a finite speed, so that a signal emitted from one clock at time t_A will be observed by another clock at a time $t_B > t_A$. Which of these times should we use to compute the scale factor in a distance formula? Moreover, the *method* of observation influences the choice of the relevant distance, because the photons themselves evolve as they travel.

To answer these questions, we can easily express the flux observed from a galaxy of luminosity L at a redshift z. The observed flux (energy per unit time per unit telescope area) is obtained by spreading the energy emitted from the source per unit time, L, over the surface area of a sphere whose radius equals the effective distance of the source,

$$f = \frac{L}{4\pi d_{\rm L}^2}, \tag{1.10}$$

where $d_{\rm L}$ is defined as the *luminosity distance* in cosmology. For a flat Universe, the comoving distance of a galaxy that emitted its photons at a time $t_{\rm em}$ and is observed at time $t_{\rm obs}$ is obtained by summing over infinitesimal distance elements along the path length of a photon, $c\,dt$, each expanded by a factor $(1+z)$ to the present time (corresponding to setting $ds^2 = 0$ in equation 1.1 for a photon trajectory):

$$R_{\rm em} = \int_{t_{\rm em}}^{t_{\rm obs}} \frac{c\,dt}{a(t)} = \frac{c}{H_0}\int_0^z \frac{dz'}{\sqrt{\Omega_m(1+z')^3+\Omega_\Lambda}}, \tag{1.11}$$

where $a = (1+z)^{-1}$. The *angular-diameter distance* $d_{\rm A}$, corresponding to the angular diameter $\theta = D/d_{\rm A}$ occupied by a galaxy of size D, must take into account the fact that we were closer to that galaxy[vii] by a factor $(1+z)$ when the photons started their journey at a redshift z, so it is simply given by $d_{\rm A} = R_{\rm em}/(1+z)$. But to find $d_{\rm L}$ we must take account of additional redshift factors.

If a galaxy has an intrinsic luminosity L, then it will emit an energy $L\,dt_{\rm em}$ over a time interval $dt_{\rm em}$. This energy is redshifted by a factor of $(1+z)$ and is observed over a longer time interval $dt_{\rm obs} = dt_{\rm em}(1+z)$ after being spread over a sphere of surface area $4\pi R_{\rm em}^2$. Thus, the observed flux will be

$$f = \frac{L\,dt_{\rm em}/(1+z)}{4\pi R_{\rm em}^2 dt_{\rm obs}} = \frac{L}{4\pi R_{\rm em}^2(1+z)^2}, \tag{1.12}$$

which implies that

$$d_{\rm L} = R_{\rm em}(1+z) = d_{\rm A}(1+z)^2. \tag{1.13}$$

Unfortunately, for a flat universe with a cosmological constant, these distance integrals cannot be expressed analytically. However, a convenient numerical approximation, valid to 0.4% relative error in the range $0.2 \le \Omega_m \le 1$ (where Ω_m is the total matter density) is[1]

$$d_L = \frac{c}{H_0}a^{-1}\left[\eta(1,\Omega_m) - \eta(a,\Omega_m)\right], \tag{1.14}$$

where

$$\eta(a,\Omega_m) = 2\sqrt{s^3+1}\left[\frac{1}{a^4} - 0.1540\frac{s}{a^3} + 0.4304\frac{s^2}{a^2} + 0.19097\frac{s^3}{a} + 0.066941 s^4\right]^{-1/8}, \tag{1.15}$$

and $s^3 = 1/\Omega_m - 1$.

[vii] In a flat Universe, photons travel along straight lines. The angle at which a photon is seen is not modified by the cosmic expansion, since the Universe expands at the same rate both parallel and perpendicular to the line of sight.

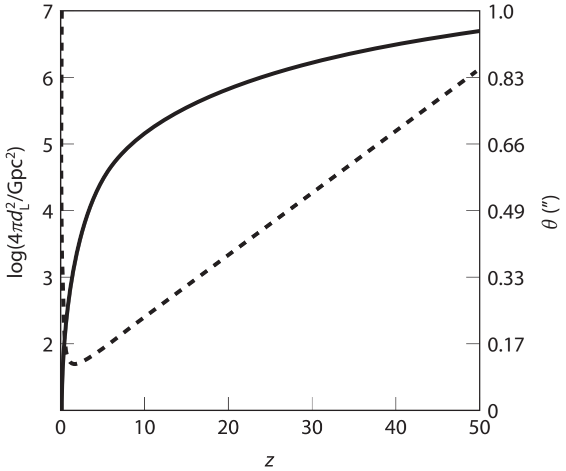

Figure 1.5 The solid line (corresponding to the label on the left-hand side) shows $\log_{10}$ of the conversion factor between the luminosity of a source and its observed flux, $4\pi d_L^2$ (in Gpc^2), as a function of redshift, z. The dashed line (labeled on the right) gives the angle θ (in arcseconds) occupied by a galaxy of 1 kpc diameter as a function of redshift.

The area dilution factor $4\pi d_L^2$ is plotted as a function of redshift in the solid curve of Figure 1.5. If the observed flux is measured over only a narrow band of frequencies, one needs to take account of the additional conversion factor $(1+z) = (d\nu_{\rm em}/d\nu_{\rm obs})$ between the emitted frequency interval $d\nu_{\rm em}$ and its observed value $d\nu_{\rm obs}$. This correction yields the relation $(df/d\nu_{\rm obs}) = (1+z) \times (dL/d\nu_{\rm em})/(4\pi d_L^2)$.

In practice, observed brightnesses are often expressed using the *AB magnitude* system. The conversion from flux density to AB magnitude is

$$\mathrm{AB} = -2.5 \log\left[\frac{df}{d\nu_{\rm obs}}\right] - 48.6, \tag{1.16}$$

where the flux density is expressed in units of erg $\mathrm{s}^{-1}\ \mathrm{cm}^{-2}\ \mathrm{Hz}^{-1}$.

1.4 Most Matter Is Dark

Surprisingly, most of the matter in the Universe is not the same ordinary matter of which we are made (see Figure 1.6). If it were ordinary matter (which also makes stars and diffuse gas), it would have interacted with light, thereby revealing its existence to observations through telescopes. Instead, observations of many different astrophysical environments require the existence of some mysterious dark component of matter that reveals itself only through its gravitational influence and leaves no other clue about its nature. Cosmologists are like detectives who find evidence for some unknown criminal at a crime scene

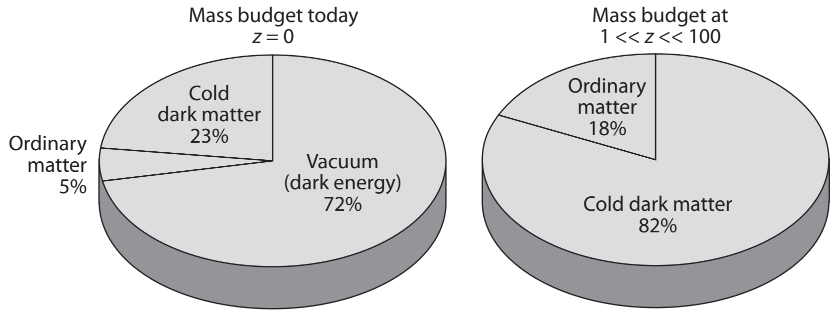

Figure 1.6 Mass budgets of different components in the present-day Universe and in the infant Universe when the first galaxies formed (redshifts $z = 10$–50). The CMB radiation (not shown) makes up a fraction (∼0.03%) of the budget today but was dominant at redshifts $z > 3{,}300$. The cosmological constant (vacuum) contribution was negligible at high redshifts ($z \gg 1$).

and are anxious to find his or her identity. The evidence for dark matter is clear and indisputable, assuming that the laws of gravity are not modified (although a small minority of scientists are exploring this alternative).

Without dark matter we would never have existed by now, because ordinary matter is coupled to the CMB radiation that filled the early Universe. The diffusion of photons on small scales smoothed out perturbations in this primordial radiation fluid. The smoothing length was stretched to a scale as large as hundreds of millions of light-years in the present-day Universe. This is a huge scale by local standards, since galaxies—like the Milky Way—were assembled out of matter in regions a hundred times smaller than that. Because ordinary matter was coupled strongly to the radiation in the early dense phase of the Universe, it also was smoothed on small scales. If there were nothing else in addition to the radiation and ordinary matter, then this smoothing process would have had a devastating effect on the prospects for life in our Universe. Galaxies like the Milky Way would never have formed by the present time, since there would have been no density perturbations on the relevant small scales to seed their formation. The existence of dark matter not coupled to the radiation came to the rescue by remembering the initial seeds of density perturbations on small scales. In our neighborhood, these seed perturbations led eventually to the formation of the Milky Way galaxy inside which the Sun was made as one out of tens of billions of stars, and Earth was born out of the debris left over from the formation process of the Sun. This sequence of events would never have occurred without the dark matter.

We do not know what constitutes the dark matter, but from the good match obtained between observations of large-scale structure and the equations describing a pressureless fluid (see equations 2.3–2.4), we infer that it is likely made of particles with small random velocities. It is therefore called "cold dark matter" (CDM). The popular view is that CDM is composed of particles that weakly interact with ordinary matter, much like the elusive neutrinos we know to exist. The abundance of such particles would naturally "freeze out" at a temperature $T > 1\,\mathrm{MeV}$, at which the Hubble expansion rate is comparable to the annihilation rate of the CDM particles. Interestingly, such a decoupling temperature, together with a weak interaction cross section and particle masses of $mc^2 > 100\,\mathrm{GeV}$ (as expected for the lightest, and hence stable, supersymmetric particle in simple extensions of the standard model of particle physics), naturally leads through a Boltzmann suppression factor $\sim \exp(-mc^2/k_BT)$ to $\Omega_m \sim 1$. The hope is that CDM particles, owing to their weak but nonvanishing coupling to ordinary matter, will nevertheless be produced in small quantities through collisions of energetic particles in future laboratory experiments such as the Large Hadron Collider (LHC). Other experiments are attempting to detect directly the astrophysical CDM particles in the Milky Way halo. A positive result from any of these experiments will be equivalent to our detective friend's being successful in finding a DNA sample of the previously unidentified criminal.

The most popular candidate for the CDM particle is a weakly interacting massive particle (WIMP). The lightest supersymmetric particle (LSP) could be a WIMP. The CDM particle mass depends on free parameters in the particle physics model; the LSP hypothesis will be tested at the Large Hadron Collider or in direct detection experiments. The properties of the CDM particles affect their response to the primordial inhomogeneities on small scales. The particle cross section for scattering off standard model particles sets the epoch of their thermal decoupling from the cosmic plasma.

In addition to dark matter, the observed acceleration in the current expansion rate of the Universe implies that the vacuum contributes $\sim$72% of the cosmic mass density at present. If the vacuum density will behave as a cosmological constant, it will dominate even more in the future (since $\rho_m/\rho_v \propto a^{-3}$). The exponential future expansion will carry all galaxies outside the local group out of our horizon within $\sim 10^{11}$ years,[2] and will stretch the characteristic wavelength of the cosmic microwave background to be larger than the horizon in $\sim 10^{12}$ years.[3]

The dark ingredients of the Universe can be probed only indirectly through a variety of luminous tracers. The distribution and nature of the dark matter are constrained by detailed X-ray and optical observations of galaxies and galaxy clusters. The evolution of the dark energy with cosmic time will be constrained over the coming decade by surveys of Type Ia supernovae, as well as surveys of X-ray clusters, up to a redshift of 2.

According to the standard cosmological model, the CDM behaves as a collection of collisionless particles that started out at the epoch of matter domination with negligible thermal velocities and later evolved exclusively under gravitational forces. The model explains how both individual galaxies and the large-scale patterns in their distribution originated from the small initial density fluctuations. On the largest scales, observations of the present galaxy distribution have indeed found the same statistical patterns as seen in the CMB, enhanced as expected by billions of years of gravitational evolution. On smaller scales, the model describes how regions that were denser than average

Table 1.1 Standard Set of Cosmological Parameters (defined and adopted throughout the book). Based on Komatsu, E., et al., *Astrophys. J. Suppl.* **180**, 330 (2009).

Ω_Λ	Ω_m	Ω_b	h	n_s	σ_8
0.72	0.28	0.05	0.7	1	0.82

collapsed owing to their enhanced gravity and eventually formed gravitationally bound halos, first on small spatial scales and later on larger ones. In this hierarchical model of galaxy formation, the small galaxies formed first and then merged, or accreted gas, to form larger galaxies. At each snapshot of this cosmic evolution, the abundance of collapsed halos, whose masses are dominated by dark matter, can be computed from the initial conditions. The common understanding of galaxy formation is based on the notion that stars formed out of the gas that cooled and subsequently condensed to high densities in the cores of some of these halos.

Gravity thus explains how some gas is pulled into the deep potential wells within dark matter halos and forms galaxies. One might naively expect that the gas outside halos would remain mostly undisturbed. However, observations show that it did not remain neutral (i.e., in atomic form) but was largely ionized by the UV radiation emitted by the galaxies. The diffuse gas pervading the space outside and between galaxies is referred to as the *intergalactic medium* (IGM). For the first hundreds of millions of years after cosmological recombination (when protons and electrons combined to make neutral hydrogen), the so-called cosmic dark ages, the universe was filled with diffuse atomic hydrogen. As soon as galaxies formed, they started to ionize diffuse hydrogen in their vicinity. Within less than a billion years, most of the IGM was reionized. This *reionization epoch* marks a crucial transition in the history of the Universe and is a prime foucs of both modern astrophysics research and this book.

The initial conditions of the Universe can be summarized on a single sheet of paper. The small number of parameters that provide an accurate statistical description of these initial conditions are summarized in Table 1.1 (see also Appendix B). However, thousands of books in libraries throughout the world cannot summarize the complexities of galaxies, stars, planets, life, and intelligent life in the present-day Universe. If we feed the simple initial cosmic conditions into a gigantic computer simulation incorporating the known laws of physics, we should be able to reproduce all the complexity that emerged out of the simple early Universe. Hence, all the information associated with this later complexity was encapsulated in those simple initial conditions. We will follow the process through which late-time complexity appeared and established an irreversible arrow to the flow of cosmic time.[viii]

[viii] In previous decades, astronomers used to associate the simplicity of the early Universe with the fact that the data about it were scarce. Although this was true at the infancy of observational cosmology, it is not true any more. With much richer data in our hands, the initial simplicity is now interpreted as an outcome of inflation.