16
Cosmology

Chapter Outline

One of the most exciting and rapidly developing areas in all of physics is **cosmology, the study of the origin, content, form, and time evolution of the Universe.** Initial cosmological speculations of a homogeneous, eternal, static Universe with constant average separation of clumps of matter on a very large scale were rather staid and dull. Now, however, proof has accumulated of a surprising *expanding* Universe with a distinct origin in time 14 billion years ago! In the past 50 years interesting and important refinements in the expansion rate (recession rate of all galaxies from each other) have been confirmed: a very rapid early expansion (inflation) filling the Universe with an extremely uniform distribution of radiation and matter, a decelerating period of expansion dominated by gravitational attraction in which galaxies had time to form yet still moved apart from each other, and finally, a preposterous application of the accelerator to the cosmic car about 5 billion years ago so that galaxies are currently accelerating away from each other again. In this chapter we explore the experimental evidence for these ideas. We also speculate on the causes of a slower or faster expansion rate. These include familiar forms of matter and energy as well as unusual forms like dark cold matter and dark energy, which produces gravitational repulsion.

16.1   EPOCHS OF THE BIG BANG

In this chapter we describe one of the most fascinating theories in all of science—the Big Bang theory of the creation of the Universe—and the experimental evidence that supports it. This theory of cosmology states that the Universe had a beginning and erupted from an extremely dense, pointlike singularity about 14 billion years ago.[1] Such an extreme of energy occurred in the first few instants after the Big Bang that it is believed that all four interactions of physics were unified and that all matter melted down into an undifferentiated "quark-gluon primordial soup."

Figure 16.1 shows the evolution of the four fundamental forces from the Big Bang to the present. During the first 10^{-43} s (the ultrahot epoch during which $T \approx 10^{32}$ K), it is presumed that the strong, electroweak, and gravitational forces were joined to form a completely unified force. In the first 10^{-35} s following the Big Bang (the hot epoch, $T \approx 10^{29}$ K), gravity broke free of this unification while the strong and electroweak forces remained joined. This was a period when particle energies were so great ($>10^{16}$ GeV) that very massive particles such as those predicted by Supersymmetry (see Section 15.12) as well as quarks, leptons, and their antiparticles existed. Then, after 10^{-35} s, the Universe rapidly expanded and cooled (the warm epoch, $T = 10^{29}$ to 10^{15} K) and the strong and electroweak forces parted company. As the Universe continued to cool, the electroweak force split into the weak force and the electromagnetic force about 10^{-10} s after the Big Bang. After a few more minutes, protons and neutrons condensed out of the cooling quark-gluon plasma.

During the first half-hour of creation, the temperature was probably about 10^{10} K and the Universe could be viewed as a cosmic thermonuclear bomb fusing protons and neutrons into deuterium and then helium, producing most of
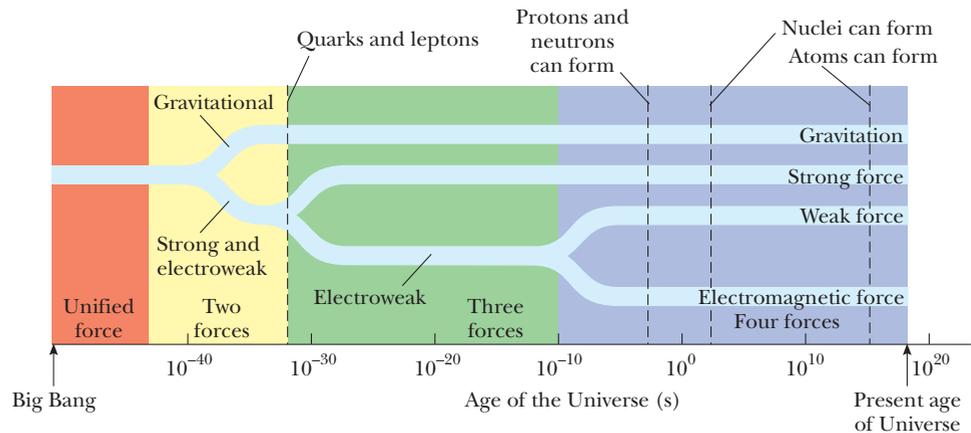

Figure 16.1   A brief history of the Universe from the Big Bang to the present. The four forces became distinguishable during the first nanosecond. Following this, all the quarks combined to form particles that interact via the nuclear force. However, the leptons remained separate and to this day exist as individual, observable particles.

[1]Data from the Wilkinson Microwave Anisotropy Probe in 2003 pinpointed the age of the Universe at 13.7 ± 0.2 billion years ("WMAP Spacecraft Maps the Entire Cosmic Microwave Sky with Unprecedented Precision," *Physics Today,* April 2003, 56(4): pp. 21–24.)
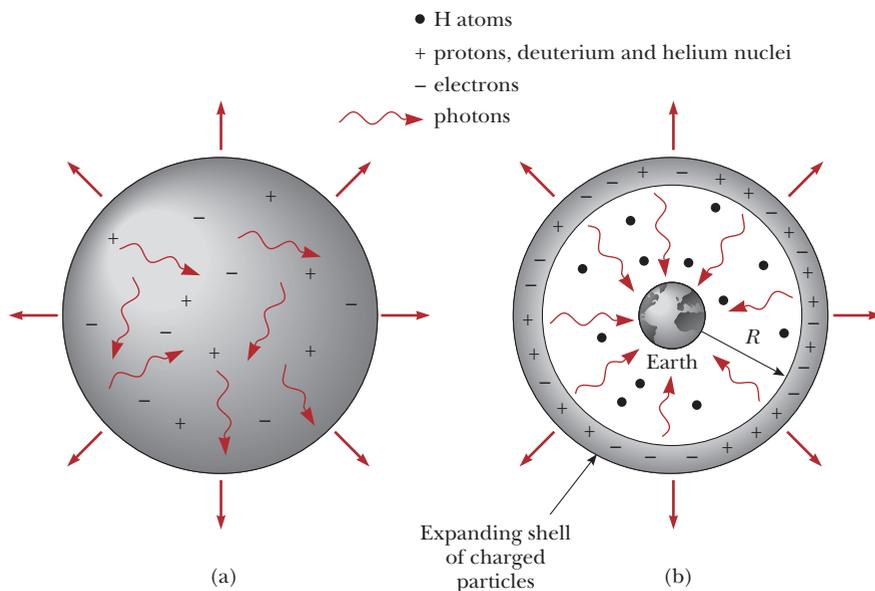
Figure 16.3 Radiation- and matter-dominated stages in the evolution of the Universe. (a) The radiation-dominated stage ($T > 3000$ K, age $< 400,000$ years) is an expanding Universe filled with protons, electrons, neutrons, photons, and neutrinos. Charged particles continually scatter photons, ensuring thermal equilibrium of radiation and matter. (b) The matter-dominated state ($T < 3000$ K, age $> 400,000$ years). Hydrogen atoms form, making the Universe transparent to photons. Photons emitted from an expanding ion shell at $R \approx 14$ billion lightyears are currently seen on Earth enormously Doppler-shifted to the red.

Figure 16.2 George Gamow, 1904–1968, Russian-American physicist. Gamow and two of his students, Ralph Alpher and Robert Herman, were the first to take the first half-hour of the Universe seriously. In a mostly overlooked paper published in 1948, they made truly remarkable cosmological predictions. They correctly calculated the abundances of hydrogen and helium after the first half-hour (75% H and 25% He) and predicted that radiation from the Big Bang should still be present, with an apparent temperature of about 5 K. In Gamow's own words, the Universe's supply of hydrogen and helium were created very quickly, "in less time than it takes to cook a dish of duck and roast potatoes." The comment is characteristic of this interesting physicist, who is known as much for his explanation of alpha decay and his theories of cosmology as for his delightful popular books, his cartoons, and his wonderful sense of humor. A classic Gamow story holds that, having coauthored a paper with Alpher, he made Hans Bethe an honorary author so that the credits would read "Alpher, Bethe, and Gamow" (to resemble the Greek letters α, β, and γ). (*AIP Emilio Segre Visual Archives*)

the helium nuclei that now exist. (See Fig. 16.2.) Until about 400,000 years after the Big Bang, the Universe was dominated by radiation. Energetic photons prevented matter from forming clumps, or even single hydrogen and helium atoms, because photon-atom collisions would instantly ionize any atoms that happened to form. In addition, photons experienced continuous Compton scattering from vast numbers of free electrons, resulting in a universe that was opaque to electromagnetic radiation. By the time the Universe was about 400,000 years old—it had expanded to one-thousandth of its current size and cooled to about 3000 K—electrons could bind to protons and helium nuclei to form atoms. Because of the drastic reduction in free charged particles, far fewer photons were absorbed or scattered, and the Universe suddenly became transparent to photons. Radiation no longer dominated the Universe, and clumps of neutral matter steadily grew—first atoms, then molecules, gas clouds, stars, and finally galaxies. The state of affairs for radiation- and matter-dominated periods is shown in Figure 16.3.

16.2 OBSERVATION OF RADIATION FROM THE PRIMORDIAL FIREBALL

In 1965, Arno A. Penzias and Robert W. Wilson of Bell Laboratories were testing a sensitive microwave receiver and made an amazing discovery. A pesky
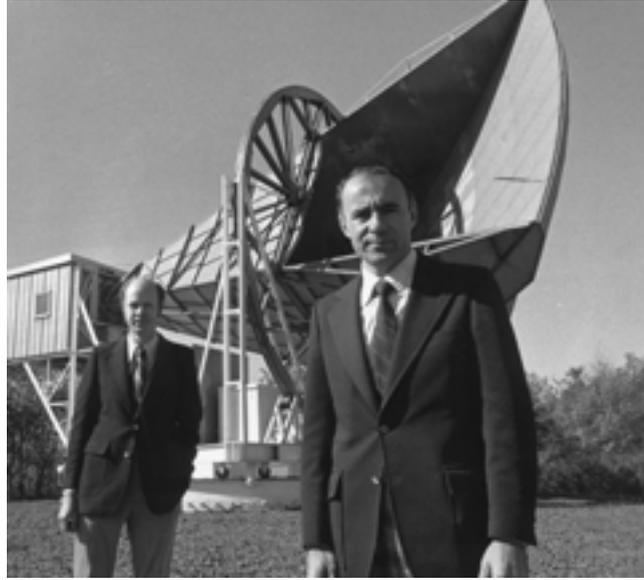
Figure 16.4   Robert W. Wilson (*left*) and Arno A. Penzias with the Bell Telephone Laboratories horn-reflector antenna. (*AIP Emilio Segre Visual Archives/Physics Today Collection*)

signal that produced a faint background hiss was interfering with their satellite communications experiments. Despite their valiant efforts, the signal remained. Ultimately it became clear that they were detecting microwave background radiation (at a wavelength of 7.35 cm) that represented some of the leftover electromagnetic "glow" from the Big Bang.

The microwave horn that served as their receiving antenna is shown in Figure 16.4. The intensity of the detected signal remained unchanged as the antenna was pointed in different directions. The fact that the radiation had equal strengths in all directions suggested that the entire Universe was the source of this radiation. Booting a flock of pigeons from the 20-foot horn and cooling the microwave detector both failed to remove the "spurious" signal. Through a casual conversation, Penzias and Wilson discovered that a group at Princeton had predicted the residual radiation from the Big Bang and were planning an experiment to confirm their theory. The excitement in the scientific community was high when Penzias and Wilson announced that they had already observed an excess microwave background of a 3-K blackbody source, which was the predicted temperature of the cooled, residual Big Bang radiation at the present time.

Because Penzias and Wilson took their measurements at a single wavelength, they did not completely confirm the radiation as 3-K blackbody radiation. One can imagine the excitement in the late 1960s and 1970s as experimentalists fought obscuring atmospheric absorption and, point by point, added intensity data at different wavelengths. The results shown in Figure 16.5, which include precise measurements taken around the maximum intensity wavelength of 0.1 cm by the Cosmic Background Explorer Satellite (COBE), unambiguously confirm that the radiation is that of a black body at 2.7 K. Figure 16.5 is probably the most clear-cut evidence that the Big Bang
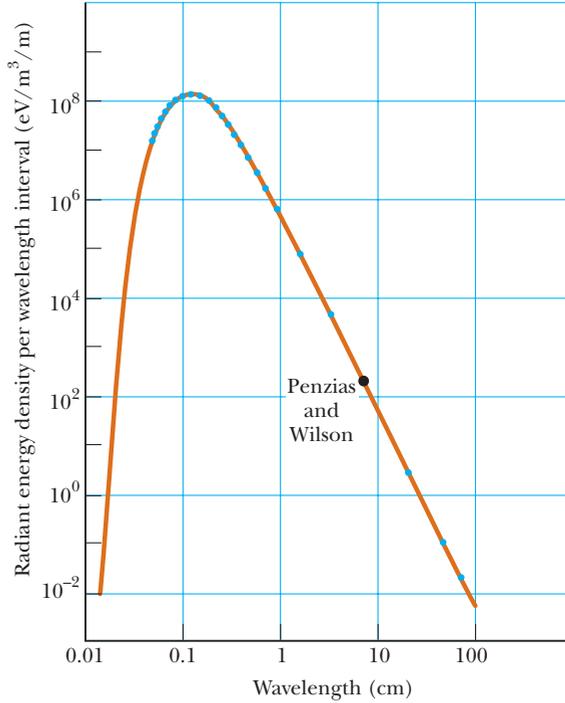
Figure 16.5 Theoretical blackbody (brown curve) and measured radiation spectra (blue points) of the Big Bang. Most of the data were collected from the Cosmic Background Explorer (COBE) satellite. The datum of Penzias and Wilson is indicated.

occurred. It also presents the earliest view of the Universe, a much earlier view than that available with the largest optical and radio telescopes. The 1978 Nobel Prize was awarded to Penzias and Wilson for their important discovery.

16.3   SPECTRUM EMITTED BY A RECEDING BLACKBODY

Since the observed 3-K blackbody spectrum is the single strongest piece of evidence for the Big Bang theory of the Universe, it is important to examine this topic in more detail. In particular, we want to show what happens to a blackbody spectrum in an expanding Universe.

Consider blackbody radiation from the Big Bang at the time when the Universe first became transparent to photons, 400,000 years *after the Big Bang* (ABB). Recall from Chapter 3 that the wavelength distribution for blackbody radiation is

$$u(\lambda,T)\,d\lambda = \frac{8\pi hc}{\lambda^5}\left(\frac{1}{e^{hc/\lambda k_{\mathrm{B}}T} - 1}\right)d\lambda \qquad (3.20)$$

where $u(\lambda,T)\,d\lambda$ is the energy per unit volume of the radiation with wavelength between λ and $\lambda + d\lambda$ emitted by a blackbody at temperature T. Suppose that between then (400,000 years ABB) and now (14,000,000,000 years ABB), the Universe has expanded and a photon of original wavelength λ is

now Doppler-shifted to λ', where λ' is longer than λ. Our question is: What happens to the form of the original wavelength distribution $u(\lambda, T)\ d\lambda$?

Let $f(\lambda', T')\ d\lambda'$ be the current energy per unit volume of the residual Big Bang radiation as measured from the Earth. Since the shell of charged particles that emitted the radiation is moving away from the Earth at extremely relativistic speed (Fig. 16.3), we must use the relativistic Doppler shift for light from a receding source to relate λ' to λ. As shown in Chapter 1, Example 1.6, this connection is

$$\lambda' = \frac{\sqrt{1 + (v/c)}}{\sqrt{1 - (v/c)}}\ \lambda \tag{16.1}$$

where v is the speed of recession of the charged shell. For our purposes, Equation 16.1 may be written more simply as

$$\lambda' = F\lambda \tag{16.2}$$

which shows that the present wavelength λ' is greater than the original wavelength λ by the factor $F = \sqrt{1 + (v/c)}/\sqrt{1 - (v/c)}$. (Equation 16.2 may be generalized to mean that all distances in the Universe have grown by a factor of F since 400,000 years ABB.) To find the relation between the currently observed spectrum $f(\lambda', T')\,d\lambda'$ and the original blackbody distribution $u(\lambda, T)\,d\lambda$, we substitute $\lambda = \lambda'/F$ into the expression for $u(\lambda, T)\,d\lambda$:

$$u(\lambda, T)\ d\lambda = \frac{8\pi hc}{\lambda^5}\left(\frac{1}{e^{hc/\lambda k_B T} - 1}\right)d\lambda = \frac{8\pi hc}{(\lambda'/F)^5}\left(\frac{1}{e^{hc/(\lambda'/F)k_B T} - 1}\right)\frac{d\lambda'}{F}$$

or

$$\frac{u(\lambda, T)\ d\lambda}{F^4} = \frac{8\pi hc}{(\lambda')^5}\left(\frac{1}{e^{hc/\lambda' k_B T'} - 1}\right)d\lambda' \tag{16.3}$$

where $T' = T/F$. We can identify the right side of Equation 16.3 as a (current) blackbody spectrum $f(\lambda', T')\,d\lambda'$ because it has the standard functional form of a blackbody spectrum with wavelength λ' and temperature T'. Thus we can write Equation 16.3 as

$$\frac{u(\lambda, T)\ d\lambda}{F^4} = f(\lambda', T')\ d\lambda' \tag{16.4}$$

where $f(\lambda', T')\,d\lambda'$ has the explicit form

$$f(\lambda', T')\ d\lambda' = \frac{8\pi hc}{(\lambda')^5}\left(\frac{1}{e^{hc/\lambda' k_B T'} - 1}\right)d\lambda' \tag{16.5}$$

with $T' = T/F$. *Equation 16.4 shows that the radiation from a receding blackbody still has the same spectral distribution, but its temperature T' drops by a factor of F and its energy per unit volume decreases by a factor of F^4.* In a frame of reference moving with the charged particles that last scattered the radiation, the characteristic temperature of the radiation is about 3000 K. Using the current observation of a 3-K cosmic background from Earth, and substituting into $F = T/T'$, we find $F = 3000\ K/3\ K \approx 1000$. This means distances in the Universe have grown by a factor of 1000 since 400,000 years ABB and that the speed of recession from the Earth of the charged shell which emitted the radiation is very high.

Exercise 1   Find the speed of recession from the Earth of the charged shell that emitted the cosmic background radiation.

Answer   0.999998 c

16.4   RIPPLES IN THE MICROWAVE BATH

The discovery of the cosmic microwave background (CMB) radiation bathing the Earth brought with it a problem, however—the radiation was too uniform. Scientists believed that slight fluctuations in the temperature of the CMB from spot to spot in the sky had to occur in order to act as nucleation sites for the formation of the galaxies, stars, and planets that now exist. In 1992, George Smoot (b. 1945), an American astronomer at the Lawrence Berkeley Laboratory, found, by carefully analyzing data collected by COBE, that the background was not perfectly uniform as was first thought, but instead contained irregularities that corresponded to temperature variations of 1 part in 100,000. While variations of this size may seem insignificant, temperature variations of just this size have been shown by computer simulation to lead to the galaxy clusters we see today.[2] The Wilkinson Microwave Anisotropy Probe (WMAP), launched in June 2001, has provided the most definitive data to date showing that distinct hot and cold spots in the CMB have angular sizes precisely predicted by cosmological theory. (See Fig. 16.6.) Additionally, WMAP has confirmed the composition of the Universe at different epochs as well as the geometry of the Universe.

What Caused the CMB Ripples?

The hot and cold spots in the CMB are believed to correspond, respectively, to density compressions and rarefactions (sound waves) in the gaslike system of
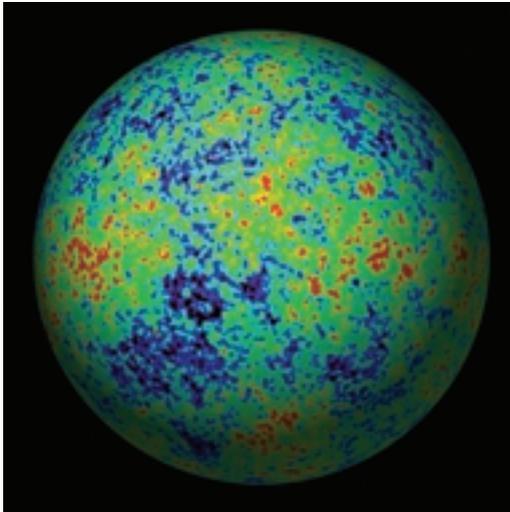

Figure 16.6   A "baby picture" of the Universe taken when it was only 380,000 years old. The Earth is at the center of this celestial sphere and red corresponds to warmer regions and blue to colder. (*NASA/WMAP SCIENCE TEAM*)

[2]See "Reading the Blueprints of Creation," by Michael A. Strauss, *Sci. Am.,* February 2004.

photons, electrons, protons, and neutrons of the radiation-dominated era when radiation was trapped by matter. About 380,000 years ABB, the Universe became transparent to radiation, and photons could travel in straight lines for billions of lightyears without scattering. Photons emitted from compressed hot areas were more energetic than those from rarified cooler areas, and when these photons reached the Earth, after a straight-line trip of 45 billion lightyears, they preserved the information about the hot and cold spots produced by sound waves in the early Universe. This explanation, however, just pushes the cause of the CMB ripples back one step, since it is natural to ask what produced the sound waves in the first place. Here, we must invoke the concept of inflation, already introduced in 1981 by Alan H. Guth (b. 1947, American physicist) to explain an important difficulty with the Big Bang theory called **the horizon problem.**

The Horizon Problem.   **The horizon problem is the issue of the causal connectedness shown by the uniform temperature of the entire CMB when, in fact, any two points in the sky separated by more than 1 degree (about twice the Moon's diameter) contain CMB radiation from regions of the universe that were causally disconnected at 380,000 years ABB.** That is, extrapolating the paths of photons from points 2 degrees apart in the sky backwards, using the known expansion rate of the Universe, results in a separation of the photons' points of origin such that a signal would have to travel at about 10 times the speed of light to correlate the behavior of the two regions. This dilemma is called the horizon problem since the limit of the region we can observe—our horizon—is defined by signals traveling at the speed of light. Thus it is difficult to understand the uniform temperature of the CMB to 1 part in 100,000 on the basis of the standard Big Bang theory. What we are left with is that by sheer coincidence all the causally disconnected regions of the early universe *happen* to have a homogeneous density and temperature.

Inflation

In 1981, Guth proposed a more natural solution to this problem, the theory of inflation.[3] **The theory of inflation states that at about 10^{-35} s ABB the Universe went through a period of rapidly accelerating expansion for about 10^{-32} s and exponentially increased in size with time.** It is thought that after inflation the volume of the Universe had increased about 10^{100} times more than it would have if it had expanded at the standard expansion rate. This solves the horizon problem because before inflation, the matter and radiation were so closely packed together that collision processes would easily keep matter and radiation in equilibrium, producing uniform density and temperature in the CMB-emitting postinflation early Universe. It is most interesting that inflation occurred so rapidly it actually outpaced the speed of light, and sections of the Universe grew out of contact with each other during inflation. When inflation ended, the expansion decelerated, and regions of the Universe came back into each other's view. This situation is shown schematically in Figure 16.7 in terms of a space-time diagram (introduced in Section 1.7). Figure 16.7 shows the plasma of photons and charged particles of the radiation-dominated era

---

[3]See "The Inflationary Universe," by Alan H. Guth and Paul J. Steinhardt, *Sci. Am.,* May 1984.
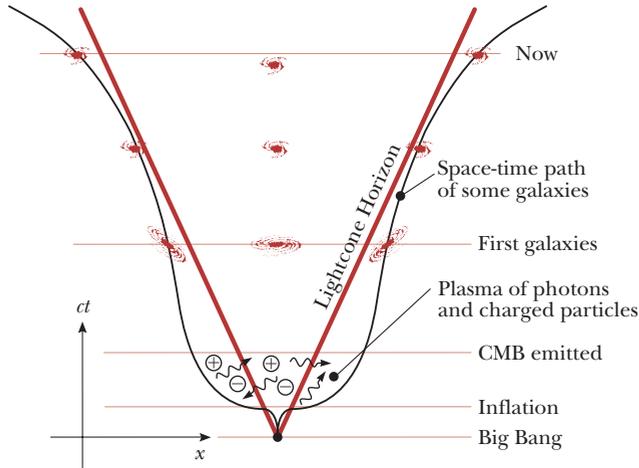
Figure 16.7 A schematic diagram of the inflationary expanding universe. The space-time diagram shows periods of deceleration and acceleration in the cosmic expansion during which some galaxies move in and out of the observable region of the universe.

expanding far beyond the horizon (lightcone) during inflation, with different regions dropping out of each others' view. Note that after inflation ended and the expansion decelerated, some of these regions came back into view. Finally, the current accelerating expansion is also suggested by showing some present galaxies moving past the horizon and out of each others' view again.

The full predictive power of the theory of inflation is really only seen, however, when it is realized that it also provides answers to several other deep problems that had been troubling cosmologists. One unsolved mystery was the observation that the present Universe seemed to obey Euclidean geometry (sum of the angles of a triangle is 180°— one says the Universe is "flat") over immense distances, which meant that at about 1 s ABB the Universe had to be exquisitely flat, precariously balanced in mass to a precision of about 1 part in 10^{15}. While standard Big Bang theory gives no reason why the Universe started off in such an unusual way, inflationary theory does—inflation drives an arbitrarily curved Universe to such a flat state that the presently observed Universe would still appear flat. This occurs in much the same way that blowing up a balloon (enough) causes a fixed area on the balloon's surface to approach flatness to any desired degree. Another puzzle facing physicists was the failure to detect magnetic monopoles in spite of extensive experimental searches. (There is strong theoretical evidence that these very massive stable particles should have been abundantly produced in the Big Bang.) Inflation solves this dilemma by showing that magnetic monopoles were so thinly spread out by the enormous inflationary expansion that their calculated density in the part of the Universe we can presently observe is below detectable limits.

The cause of inflation is believed to be a new kind of quantum field, the inflaton, which carried the potential energy needed to cause the rapid expansion and trigger sound waves in the early Universe. Furthermore, a consistent theory has been developed that directly relates microscopic quantum fluctuations in the inflaton field to observed macroscopic CMB temperature variations. This is a theory that essentially shows that inflation magnified quantum

fluctuations into the temperature variations we see in the CMB today. Although describing this theory in detail is beyond the scope of this chapter, we can describe in a crude way how the pattern of sound waves inferred from the temperature variations in the CMB fit into inflationary theory.

Sound Waves in the Early Universe

Because inflation triggered sound waves with different frequencies all at the same time of 10^{-35} s ABB, the waves are believed to have all started out in phase and oscillated throughout the radiation-dominated era for 380,000 years. If this was the case, astronomers should be able to measure the strengths of the fundamental and overtones—the power spectrum—in the acoustic oscillations of the early Universe in a process similar to that of measuring the sound spectrum of a musical instrument. In the case of the CMB, cosmologists measure the magnitude of temperature variations versus the angular size of hot and cold spots to get the power spectrum. When these measurements are made, a series of peaks with different strengths are found, agreeing with the predictions of inflationary theory in strength and frequency and confirming that all sound waves were indeed generated at the same moment by inflation.

Another success of inflationary theory applied to the radiation-dominated era is proof that the Universe obeys Euclidean geometry. Since cosmologists can precisely calculate the length of the fundamental sound wave in the early Universe (f_0 is measured, the velocity of sound v in the plasma is known, and $\lambda = v/f_0$), and the distance CMB photons have traveled to the Earth from opposite ends of the fundamental is known, the sides of a triangle, literally as big as all outdoors, are known, and the sum of its angles can essentially be checked. When this is done, it is found that space is indeed quite flat or Euclidean, in agreement with other types of measurement.

16.5   OTHER EVIDENCE FOR THE EXPANDING UNIVERSE

Evidence from Observational Astronomy

Most of the key discoveries supporting the theory of an expanding Universe, and indirectly the Big Bang theory of cosmology, were made in the 20th century. In 1913, Vesto Melvin Slipher (1875–1969), an American astronomer working at the Lowell Observatory, reported that most spiral galaxies[4] were receding from the Earth at speeds up to several million miles per hour. Slipher was one of the first to use the methods of Doppler shifts in spectral lines to measure velocities. The key to this method is to locate in galactic spectra the characteristic lines of some element whose wavelengths have been measured on Earth. The speed of the receding galaxy (for recession speeds small compared to c) is then directly proportional to the wavelength shift to the red, as shown in Figure 16.8. The symbol Z is used by astronomers to denote the

---

[4]This was before Hubble had figured out what galaxies really were, very distant "island universes" consisting of billions of stars. In Slipher's time, galaxies with spiral arms were called spiral nebulae (nebulae are glowing clouds of gas and dust), since individual stars in the glowing clouds could not be distinguished.
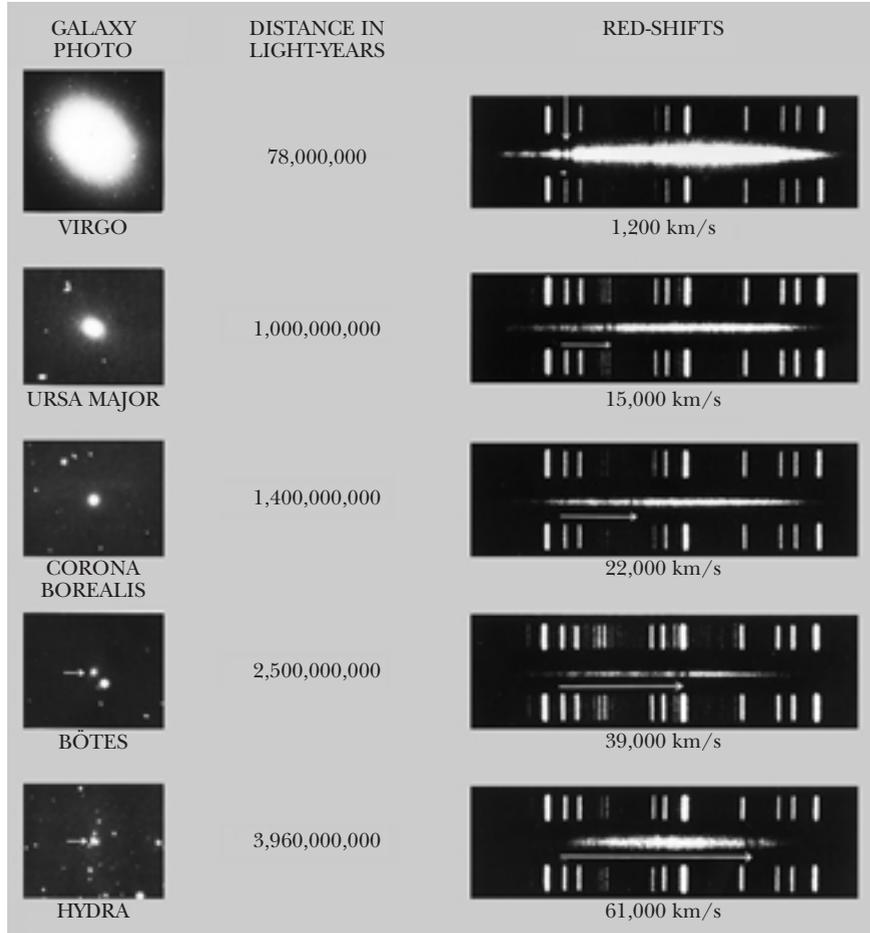
GALAXY PHOTO	DISTANCE IN LIGHT-YEARS	RED-SHIFTS
VIRGO	78,000,000	1,200 km/s
URSA MAJOR	1,000,000,000	15,000 km/s
CORONA BOREALIS	1,400,000,000	22,000 km/s
BÖTES	2,500,000,000	39,000 km/s
HYDRA	3,960,000,000	61,000 km/s

Figure 16.8   Redshift of the H and K lines of calcium for five galaxies. Redshifts are expressed as speeds $c(\Delta\lambda/\lambda)$. Calibration spectra taken in the laboratory appear above and below each galaxy's spectrum.

measured **redshift** of receding astronomical objects, where

$$Z \equiv \frac{\Delta\lambda}{\lambda} = \frac{\lambda - \lambda_0}{\lambda_0} \tag{16.6}$$

λ_0 is the emission wavelength of a particular line measured in the lab, and λ is the measured wavelength of the same line in a receding galaxy. Slipher found that the number Z was positive in most cases and had the same value for all lines in the spectrum of a particular galaxy. In general, redshift Z is related to galaxy recession velocity v by the relativistic Doppler shift expression (Eq. 16.1), which is frequently written in the form

$$Z = \sqrt{\frac{c + v}{c - v}} - 1 \tag{16.7}$$

v can be explicitly given in terms of Z as

$$v = c \left(\frac{Z^2 + 2Z}{Z^2 + 2Z + 2} \right) \tag{16.8}$$

Figure 16.9 Milton L. Humason (1891–1972). He and Edwin Hubble did the most to demonstrate the expanding Universe experimentally. The two astronomers present an interesting contrast. Humason was originally a mule-team driver and janitor at Mount Wilson. He taught himself astronomy and became a consummate experimentalist. (*Photo by Margaret/AIP Emilio Segre Visual Archives*)

Slipher's measurements of galactic speeds were quite accurate, but the distances to the galaxies were virtually unknown, although there was some indication that more-distant objects had higher recessional speeds.

Shortly after Slipher's discovery, a flowering of ideas occurred in theoretical cosmology. In 1916, Einstein published his general theory of relativity (a new theory of gravity replacing Newton's), in which he formulated the differential equations connecting the distribution of matter in the Universe with the curvature of space-time. Einstein solved these equations and found that in order to have a Universe static in time,[5] he had to add a "cosmic repulsion" force between galaxies characterized by a **cosmological constant, Λ.** This was a rather unusual force, since it was independent of mass and grew stronger with increasing galaxy separation. In the same year, the Dutch mathematician and astronomer Willem de Sitter found another static solution to Einstein's equations. This solution predicted an increasing redshift for lines emitted by more-distant objects. De Sitter's solution stirred much interest because it seemed capable of being confirmed experimentally.

Major breakthroughs in the cosmological problem came in 1922 and 1924 when Alexander Friedmann,[6] (1888–1925, Russian mathematician), showed that two *nonstatic* models were predicted by Einstein's field equations of general relativity: an expanding Universe, called an **open Universe** and a contracting Universe, called a **closed Universe.** At first Einstein found it difficult to exchange the eternal, steady, static Universe for an expanding, dynamic Universe that presumably had a beginning in time. He wrote to de Sitter, "This circumstance of an expanding Universe irritates me" and "to admit such possibilities seems senseless."[7] However, by 1933 Einstein was finally convinced of the impeccable logic of Friedmann's arguments. Any remaining doubt about the nonstatic nature of the Universe was erased by the powerful experimental results of American astronomers Edwin P. Hubble (1889–1953) and Milton Humason (1891–1972) on the redshifts of distant nebulae (see Figs. 16.9 and 16.10). We return to a consideration of Freidmann's important work, the foundation of current cosmology theory based on general relativity, in Section 16.6.

Hubble and Company Observe Galaxies

Hubble first established that spiral nebulae (distant clouds of glowing gas and dust) were actually composed of billions of stars, and then went on to measure

---

[5]The average distance between clumps of matter on a large scale is constant with time.

[6]Freidmann was an amateur meteorologist and balloonist who taught Gamow relativity theory. Gamow reported that meteorology killed Freidmann, since he died of a severe chill and pneumonia following one of his meteorological balloon flights.

[7]There has been much scientific irritation over expanding Universe theories. Presumably the irritation derives from the fact that the Big Bang theory places the cause of creation beyond the reach of physical investigation. Fred Hoyle (b. 1915, an English astronomer) has never accepted the Big Bang theory and prefers instead a steady-state model of the Universe in which matter is created out of nothing at a rate necessary to compensate for the increasing distances between visible galaxies. Others have expressed their irritation as follows:

Sir Arthur Eddington: "I have no axe to grind in this discussion but the notion of a beginning is repugnant to me . . . I simply do not believe that the present order of things started off with a bang . . . The expanding Universe is preposterous and incredible; it leaves me cold."

Walter Nernst: "To deny the infinite duration of time would be to betray the very foundation of science."

Phillip Morrison: "I find it hard to accept the big bang theory. I would like to reject it."

Allan Sandage: "It is such a strange conclusion. . . . It cannot really be true."

Figure 16.10 Edwin P. Hubble (1889–1953). Hubble was born to wealth, graduated from the University of Chicago, and was a Rhodes Scholar, lawyer, athlete, soldier, and astronomer with an unerring eye for important problems in astronomy. (© *Bettmann/CORBIS*)

their distances. By 1923, he had already determined that the Andromeda nebula was more than 1 million lightyears distant, well beyond the 100,000-lightyear extent of our own Milky Way galaxy. Hubble measured the distances to nearby galaxies, or "island universes," by observing the brightness of stars with fluctuating intensity, called *cepheid variables*. Since the intrinsic luminosity, L, the total electromagnetic energy radiated per second, was known from the period of light fluctuations, Hubble used the observed intensity flux, f, the energy per second per unit area in some frequency band received at the detector, and energy conservation to find the distance, R:

$$f = \frac{L}{4\pi R^2} \tag{16.9}$$

In the late 1920s, Hubble made the bold assertion that the entire Universe was expanding. From 1928 to 1936, he and Humason toiled at Mount Wilson to prove this assertion until they reached the limits of the 100-inch telescope. The results of their work and its continuation on a 200-inch telescope in the 1940s showed that the recession velocity, v, of a galaxy increases in direct proportion to its distance R from us (Fig. 16.11). This linear relationship, known as **Hubble's law,** may be written

$$v = H_0 R \tag{16.10}$$

Hubble's law

where H_0, **Hubble's constant,** has the currently accepted value

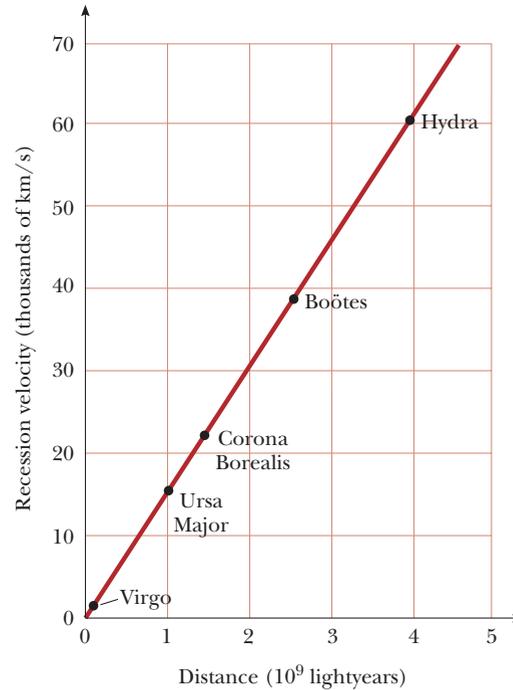$$H_0 = 23 \times 10^{-6} \text{ (km/s)/lightyear}$$

Figure 16.11   Hubble's law: The recession velocity of a galaxy is directly proportional to distance out to at least 5 billion lightyears. The five galaxies shown in Figure 16.8 are plotted here.

with an uncertainty of about 25%. This large uncertainty is produced by the great difficulty of measuring distances to far-off galaxies and galaxy clusters. The amazingly simple form of Hubble's law follows from **the cosmological principle, which states that the Universe is isotropic (has the same average features—mass density, expansion factor, and so on— in any direction) and homogeneous (has the same average features at any position) when large enough volumes of the Universe are sampled.** If we consider several equally spaced galaxies in a line moving at nonrelativistic velocities away from us in an expanding Universe, we easily find Hubble's law. (See Problem 2.) So, Hubble's law is a direct and expected consequence of an expanding homogeneous Universe.

The use of the term **Hubble's** *constant* for H_0 is a bit misleading since H_0 is constant with distance but varies slowly with time because gravity or other forces (so-called vacuum repulsion or dark energy), respectively, slow down or speed up the expansion of the Universe as time progresses. We can gain more insight into the meaning of Hubble's law and show that Hubble's "constant" is a function of time by introducing the **universal expansion scaling factor,** $a(t)$, which is the same anywhere in the Universe. $a(t)$ is defined by the equation

$$R(t) = R_0 a(t) \qquad (16.11)$$

where $R(t)$ is the distance from Earth to some receding galaxy, R_0 is the current separation in some appropriate length unit, and $a(t)$ is the dimensionless scaling factor with the present time denoted t_0, and $a(t_0) = 1$. Differentiating

Equation 16.11 with respect to time and substituting for R_0 from 16.11, we find

$$v = \dot{R} = R_0\dot{a} = \frac{R}{a}\,\dot{a} \qquad (16.12)$$

where v, R, a, and $\dot{a}$ are all functions of time. Finally, comparing 16.12 to the general form of Hubble's law written as

$$v(t) = H(t)R(t) \qquad (16.13)$$

we find

$$H(t) = \frac{\dot{a}(t)}{a(t)} \qquad (16.14)$$

Equation 16.14 shows that H is the time derivative of $\ln a$ and that a generally depends on time except in the case of exponential expansion (inflation), $a(t) = Ae^{bt}$. (See Problem 4.) Further, 16.14 shows that H is proportional to the first derivative with respect to time of a. Another parameter, $q(t)$, **the deceleration parameter,** determines the second derivative of a and tells whether the expansion is slowing down or speeding up with time.[8] H and q can be used to construct a first-order approximation to the function $a(t)$, but in general one needs to measure an infinite set of time derivatives of $a(t)$ or the function itself to fully characterize $a(t)$. We will further discuss this important scaling function, which describes the expansion of the Universe, in Section 16.6.

EXAMPLE 16.1

The basic form of Hubble's law, $v = H_0R$, follows from assuming that the Universe is undergoing a nonrelativistic uniform expansion and is homogeneous. (See Problem 2.) Thus this law in its basic form will begin to fail at some maximum distance, R_{max}, at which the recessional velocity is an appreciable fraction of the speed of light. Estimate the maximum distance at which Hubble's law applies without relativistic correction by assuming that it fails when $v = c/4$.

Solution

$$R_{max} = \frac{v}{H_0} = \frac{c}{4H_0} = \frac{3.0 \times 10^5 \text{ km/s}}{80 \times 10^{-6} \text{ km/s/lightyear}} \approx 4 \times 10^9 \text{ lightyears}$$

Hubble's law can be corrected for relativistic effects, so that interesting departures from this law can be observed such as accelerations and decelerations in the expansion of the Universe at different epochs.

New Evidence from Exploding Stars

Telescopes look farther back in time as they view more distant stars and galaxies, so by looking at objects over as great a range of distances as possible we can establish the entire history of the expansion of the Universe. In particular, measuring the velocity (from redshift) and distance (from apparent brightness or flux) of astronomical objects over a wide range of distances gives the history of the expansion of the Universe. One can tell if the Universe departed from the

---

[8]The exact form of q is given by $q = -a(\ddot{a}/\dot{a})^2$, where $q > 0$ indicates the expansion is slowing down and $q < 0$ indicates the expansion is speeding up.
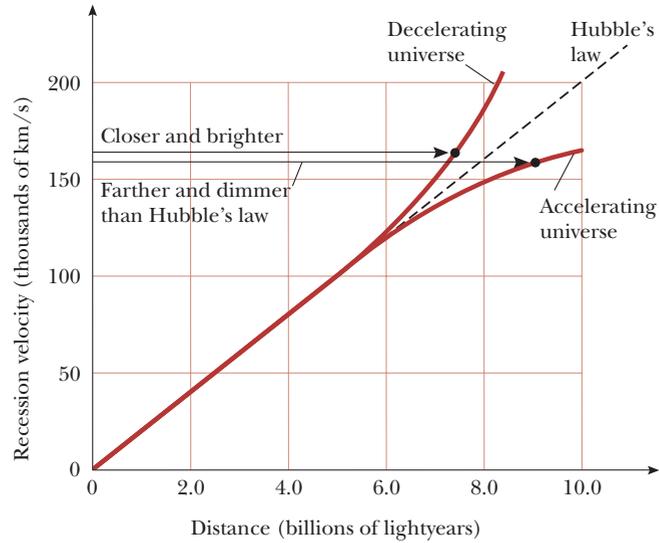
Figure 16.12   Deviations from Hubble's law (uniform expansion) for decelerating and accelerating expansions.

uniform expansion of space predicted by Hubble's law and if the expansion accelerated or decelerated at different times by comparing the measured values of velocity and distance of astronomical objects to those predicted by Hubble's law. For example, if the Universe is currently decelerating, the velocity of a very distant galaxy, determined from the redshift of light emitted by this galaxy in the early Universe, would be somewhat greater than the velocity predicted by Hubble's law; see Figure 16.12. As can be seen from the figure, an equally valid signature of deceleration is that a galaxy with a given velocity will be closer and brighter than expected from Hubble's law. Similarly, an accelerating expansion would cause galaxies in the early Universe to be more distant and dimmer than predicted by Hubble's law, as shown in Figure 16.12.

Attempts to measure the distance to very distant galaxies and groups of galaxies in the second half of the 20th century in order to detect departures from Hubble's law proved very uncertain, with great statistical scatter in data points. The failure of galaxies to act as standardized light sources at large distances is now known to be caused by evolutionary changes in galaxies—early galaxies formed at 2–5 billion years ABB have been shown to have different shapes and sizes compared to more current galaxies 9–14 billion years ABB. Thus, galaxies are too variable in structure and intrinsic brightness over very long periods of time to serve as uniform markers of the expansion. However, a particular class of exploding stars undergoing "standardized" thermonuclear detonation are so bright, so uniformly luminous, and so well understood physically that they serve as excellent tracers of the Universe's expansion. These objects, as bright as 10 billion suns for a time, are called type Ia supernovae (SNe Ia) and are formed after the collapse of white dwarf stars. Since spectra from high-Z $(Z \geq 1)$ supernova are shifted into the near IR where obscuring emission from the Earth is high, supernovae chasers have had to use the Hubble Space Telescope (HST) to accumulate the best results on the history of cosmic expansion.
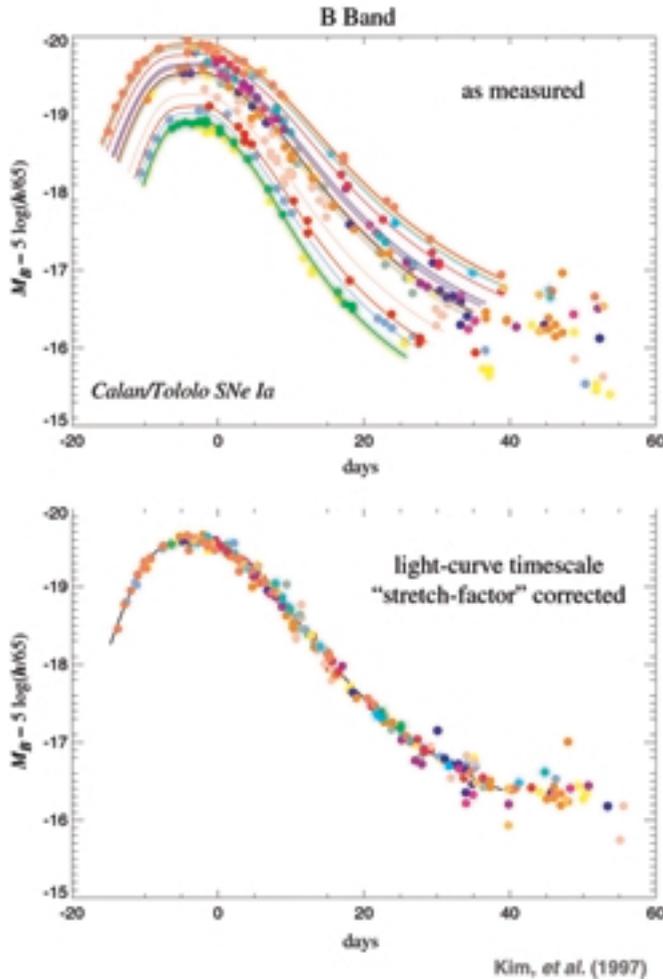
Figure 16.13   (a) Light curves for SNe Ia at different distances moving with different recession speeds. (b) Universal light curve resulting from correction of the data shown in (a) for various factors, including cosmic expansion and time dilation. *(Courtesy the Supernova Cosmology Project, supported by the Director, Office of Science, High Energy Physics, U.S. Department of Energy under Contract No. DE-AC03-76SF00098)*

Two teams, one led by Saul Perlmutter (Supernova Cosmology Project)[9] and one by Adam Riess (Space Telescope Institute Baltimore) have discovered most of the distant new Ia supernovae by surveying large sections of sky to locate and measure the rising and falling light curves characteristic of SNe Ia. Figure 16.13a shows as-measured light curves (flux versus time) of SNe Ia with different Z's, describing SNe located at different distances and having different recession velocities. Note that the width and peak values of the light curves are all greatly different. When the light curves are corrected for cosmic expansion, relativistic time dilation, intervening dust absorption, and instrumental effects, the light curves coalesce and show the same intrinsic peak luminosity and light curve width (Fig. 16.13b). *This universal light curve shows that*

[9]See the article by Saul Perlmutter, *Physics Today,* April 2003, p. 53.

individual type Ia supernovae explosions are incredibly uniform and that SNe Ia are excellent probes of cosmic history, resulting in distance and lookback time measurements with uncertainties of only about 5%. The correction to the light curve for cosmic expansion is made because the distance to the SN was $1/(1 + Z)$ of its present value at the time of emission of light now seen as redshifted by Z. (So for a $Z = 1$ SN, the SN was half its current distance when the light was emitted or the Universe was half its present linear size.) The relativistic time correction involves the same factor of $1 + Z$. Because the SN was receding from us at high velocity at the time its light was emitted, we see the detonation slowed down by a factor of $1 + Z$, which means that for a $Z = 1$ supernova, an explosion lasting 40 days in its rest frame appears to us to take 80.

What HST Found.   The first results in 1998 showed that the light from SNe which had exploded around 9 billion years ABB was fainter than predicted by Hubble's law. Consequently these objects were thought to be farther away than expected and that the expansion of the Universe was currently speeding up. Such a controversial finding immediately led to suggestions of other causes of the dimming, such as intergalactic dust, which would make the SNe seem farther away than they actually were. **In 2002, a new imaging instrument installed on the HST called the Advanced Camera for Surveys greatly boosted Hubble's ability to find supernovae.** Many more ancient SNe that exploded from 5 to 7 billion years ABB were located and measured, a real trick since there is only about one SN explosion per century in a large galaxy and such ancient SNe are very dim. The results conclusively show that the current acceleration is real, and that prior to about 9 billion years ABB, the expansion of the Universe was actually decelerating! Because dust filling intergalactic space would make more-distant (higher redshift) SNe dimmer yet, and the more-distant SNe are actually much brighter than the results predicted for dust and not dimmer, the latest results rule out dust. Figure 16.14 shows a plot of the departure of measured brightness of supernovae from what is expected at a specific redshift for a Universe governed by the freely expanding Hubble model. (Note that increasing redshift corresponds to decreasing time after the Big Bang, with $Z = 0.5$ corresponding to 9 billion years ABB and $Z = 2$ to about 3 billion years ABB.) The data points are averages of several measured SNe with the same redshift, and the curve through the points shows a transition from recent acceleration to past deceleration at about $Z = 0.5$. Model curves showing constant acceleration (positive slope) and deceleration (negative slope) over the age of the Universe are seen to be definitely ruled out by the new high-Z measurements.

Big Bang Nucleosynthesis

Another important confirmation of Big Bang theory comes from the predictions of cosmic abundances of light elements made in the first half-hour after the Big Bang. When the equations of the hot Big Bang are combined with the well-understood physics of low-energy nuclear reactions, detailed predictions of the abundances of certain light elements emerge that are in good agreement with observed values. In addition, considerations of nucleosynthesis determine the amount of neutrons and protons in the Universe. Let's investigate this in a little more detail.
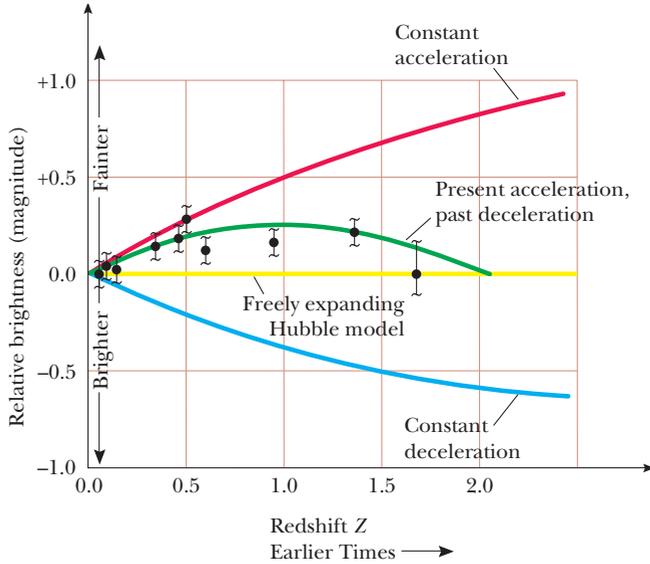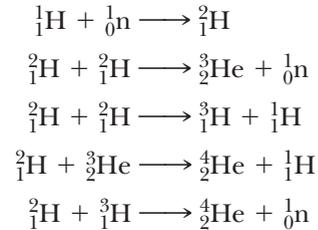
Figure 16.14 Observed brightness (flux) of supernovae relative to the brightness predicted by Hubble's law. The relative brightness is measured in terms of astronomical magnitude, which is smaller for brighter objects. Constant acceleration and deceleration of the cosmic expansion are indicated by the red and blue curves, respectively. The observed data points are best fit by the green curve, showing a transition from current acceleration to past deceleration at $Z = 0.5$ and with cosmological parameters $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$. The freely expanding Hubble line has $\Omega_m = 0$, $\Omega_\Lambda = 0$.

George Gamow and his collaborators were the first to make serious calculations of the production of elements following a hot Big Bang. In a series of papers starting in 1946, they predicted that 75% of the observable mass in the Universe was hydrogen and about 25% was helium, cooked up in a Big Bang that they predicted should show residual blackbody radiation at a temperature of 5 K. They also calculated the changing neutron-to-proton ratio as a function of time (this ratio is crucial in determining what reactions occur) but were stymied in their attempts to show that *all* the elements from H to U were made in the Big Bang by the occurrence of a mass 5 bottleneck. This means that at the temperature, pressure, and density calculated for the expanding Universe in the first few minutes, there were no mass number $A = 5$ nuclei formed for long enough times to allow $A = 6$ and heavier nuclei to be formed by combination with a neutron, proton, or other synthesized nuclei. It is now known that 98% of the observable mass in the cosmos in the form of hydrogen, helium, and trace amounts of lithium was produced in the first few minutes of the Universe. All heavier elements formed later in stellar interiors (earliest stars ~0.3 billion years ABB) and were spit into interstellar space by supernovae detonations.[10] Although helium can indeed be produced in stars, calculations first made by Fred Hoyle and Roger Taylor in 1964 showed that if

[10]The answer to the mass 5 puzzle was found in 1952 by E. E. Salpeter, who showed that in the hot, dense cores of red giant stars there is an unstable $A = 5$ state which lasts barely long enough to make $A = 6$ nuclei.

all the helium were made in stars, galaxies would be 10 to 100 times brighter than they are observed to be—thus confirming the primordial origin of cosmic helium.

The most important reactions considered in helium production at the temperatures and densities of the first few minutes ABB are

$$\ce{^1_1H + ^1_0n -> ^2_1H}$$

$$\ce{^2_1H + ^2_1H -> ^3_2He + ^1_0n}$$

$$\ce{^2_1H + ^2_1H -> ^3_1H + ^1_1H}$$

$$\ce{^2_1H + ^3_2He -> ^4_2He + ^1_1H}$$

$$\ce{^2_1H + ^3_1H -> ^4_2He + ^1_0n}$$

Current calculations of all the elements made in the first few minutes are quite sophisticated, involving shifting numbers of protons, neutrons, and other synthesized nuclei and hundreds of interdependent nuclear reactions of nuclei with As as high as 23. Figure 16.15 shows the results of such calculations, starting from thermal equilibrium values of neutrons and protons and showing the production of hydrogen, helium, deuterium, lithium, and beryllium as a function of time and temperature in the early Universe. In particular, most of the helium, deuterium, and lithium in the Universe today was produced shortly after the Big Bang, and the observed abundances agree with the calculations shown in Figure 16.15.
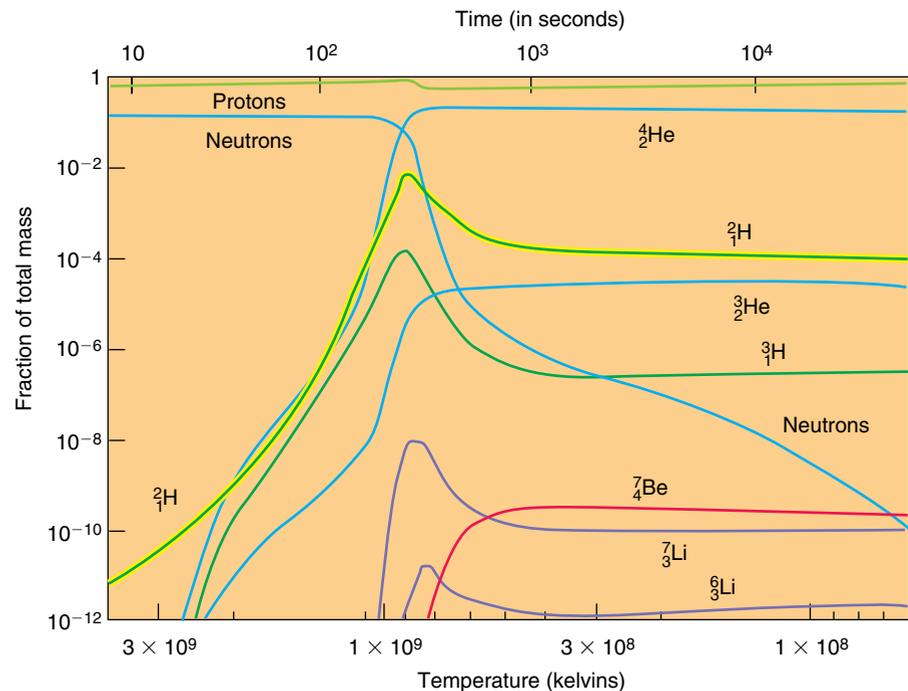

Figure 16.15 Light element abundances in the first minutes. (*Courtesy of Robert Wagoner, Stanford University*)

Critical Density, Ω_m, and Dark Matter

A final point on nucleosynthesis involves the strong dependence of calculated abundances of deuterium, helium-3, and lithium-7 on the density of neutrons and protons at the time of formation of these three elements in the early Universe. Since the density of n + p in the early Universe may be extrapolated forward by using the expansion factor of the Universe, we can obtain an estimate of the **present cosmic mass density of neutrons and protons,** which is difficult to measure directly. When this is done, one finds a value of

$$\rho_{p+n} = (3 \pm 1) \times 10^{-31} \text{ g/cm}^3$$

Cosmologists designate the ratio of the observed effective mass density (mass density + energy density/c^2) of the Universe to the calculated critical mass density, ρ_c, by $\Omega_m = \rho_{obs}/\rho_c$. The critical density, ρ_c, is the calculated mass density of the Universe needed to gravitationally halt the recession of one galaxy from another at infinite separation. ρ_c depends on the somewhat uncertain value of H_0 (see Example 16.2 below) and has the value

$$\rho_c = (1.1 \pm 0.7) \times 10^{-29} \text{ g/cm}^3$$

Thus the Big Bang model of nucleosynthesis predicts that the density of neutrons and protons (making up stars, black holes, intergalactic dust and gas—ordinary matter) is only 1%–10% of the critical density. Even the effective gravitational mass of all the photons and neutrinos in the Universe adds only another 1% at most to the mass of the Universe. The largest contribution to the mass of the Universe is believed to come from **cold dark matter,** which is inferred to exist from studies of the motion of outlying stars orbiting galaxies, the motion of galaxy clusters, and from images formed by gravitational lenses (composed of clusters of galaxies) of background galaxies. All these cases show a massive halo of invisible matter extending well beyond the visible material in galaxies and contributing a factor of around 5 to 10 times the visible mass density to the mass of the Universe. Adding in dark matter makes the current value of the ratio of the observed to critical mass density, Ω_m, about 0.3. **The actual nature of cold dark matter is uncertain, but it is not composed of ordinary atoms or elementary particles, which interact with photons as described by the standard model (Chapter 15). Most likely, dark matter consists of stable, massive, slow-moving exotic particles created at high energy in the first moments of the Big Bang and predicted by grand unified theories like Supersymmetry. The neutralino is the most likely dark matter candidate predicted by Supersymmetry.** The neutralino is the lightest supersymmetric particle, hence the most stable, and has no electric charge, so it is not affected by photons and only weakly interacts with ordinary matter. Unfortunately, this makes neutralinos hard to detect, and no conclusive detection of dark matter in earthborne detectors has been reported to date.[11]

Dark Matter

---

[11]For more information on the detection of dark matter, see "The Search for Dark Matter," by David B. Cline, *Sci. Am.,* March 2003, p. 50.

EXAMPLE 16.2   Critical Density of the Universe

We can estimate the critical mass density of the Universe, ρ_c, using classical energy considerations. The result turns out to be in agreement with the rigorous predictions of general relativity because of the simplifying assumption that the mass of the Universe is uniformly distributed.

Solution   Figure 16.16 shows a large section of the Universe with radius R with the critical density, containing a total mass M, where M consists of the total mass of matter plus the effective mass of radiation with energy E, E/c^2. A galaxy of mass m and speed v at R will just escape to infinity with zero speed if the sum of its kinetic energy and gravitational potential energy is zero. Thus,

$$E_{\text{total}} = 0 = K + U = \tfrac{1}{2}\,mv^2 - \frac{GmM}{R}$$

$$\tfrac{1}{2}\,mv^2 = \frac{Gm\frac{4}{3}\pi R^3 \rho_c}{R}$$

$$v^2 = \frac{8\pi G}{3}\,R^2 \rho_c$$

Because the galaxy of mass m obeys the Hubble law, $v = HR$, the preceding equation becomes

$$H^2 = \frac{8\pi G}{3}\,\rho_c \qquad \text{or} \qquad \rho_c = \frac{3H^2}{8\pi G}$$

Using $H = 23 \times 10^{-3}$ m/(s·lightyear), where 1 light-year $= 9.46 \times 10^{15}$ m and $G = 6.67 \times 10^{-11}$ N·m^2/kg^2, yields a present value of the critical density $\rho_c = 1.1 \times 10^{-26}$ kg/m^3. As the mass of a hydrogen atom is 1.67×10^{-27} kg, ρ_c *corresponds* to about 7 hydrogen atoms per cubic meter, an incredibly low density.
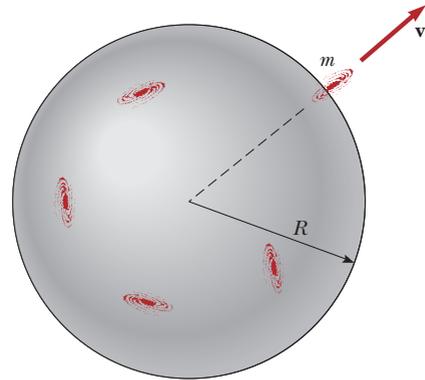

Figure 16.16   (Example 16.2) A galaxy escaping from a large cluster contained within radius R. Only the mass within R slows the mass m.

16.6   FREIDMANN MODELS AND THE AGE OF THE UNIVERSE

Freidmann's work established the foundation for describing the time evolution of the Universe based on general relativity. General relativity must be used in cosmological calculations because it correctly describes gravity, the most important force determining the Universe's structure, over immense cosmological distances. Newtonian theory can lead to errors when applied to the Universe as a whole because it assumes that the force of gravity is always attractive and is instantaneously transmitted. Although Freidmann did consider models both with and without Einstein's repulsive form of gravity (cosmological constant), it is easiest to see the general form of Big Bang behavior without introducing repulsive gravitational forces at this point.

Freidmann found three types of time-dependent universes, which may be described in terms of the universal expansion scaling factor $a(t)$. Figure 16.17 shows $a(t)$ (the separation between galaxies) as a function of time for the three cases labeled open universe, flat universe, and closed universe. Note that $a(t)$ alone has a value of zero at the lower-left corner of the graph, not t, and that the three curves start at different times in the past in order to give the same scaling factor at the present time, denoted t_0. **Open universes have less mass and energy than that needed to halt the expansion.** They start with a scale factor of zero and grow without limit, any given galaxy approaching a limiting
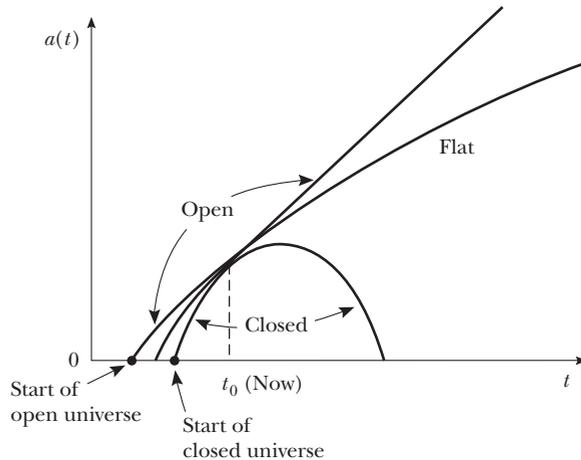
Figure 16.17 Time-dependent Freidmann universes. The figure shows the time de-
pendence of the universal expansion scaling factor $a(t)$ (proportional to the distance
between galaxies) for the three cases of open, closed, and flat universes. It also shows
that the Big Bang occurred at different times in the past for the different cases.

velocity that is higher the lower the mass-energy density within the universe.
Thus, the curve labeled open universe shows a typical case and really stands
for a set of curves describing universes with different densities, all less than the
critical density. **Closed universes have so much gravitationally attractive
mass-energy that the expansion is actually stopped and the recessional motion
of all the galaxies reversed so that the universe collapses and returns to the
state $a(t) = 0$.** Again the figure shows a single closed-universe curve, which
stands for a set of curves describing universes with densities greater than the
critical density. Both open and closed universes satisfy non-Euclidean geome-
tries. In between these two cases is the case of **the flat universe. A flat universe
has precisely the critical mass-energy density, and the recessional velocity of
any two galaxies approaches zero with increasing time.** A flat universe is
Euclidean.

The connection between Freidmann models and the age of the Universe is
easiest to see in the simple hypothetical case shown by the straight line, which is
the tangent to the $a(t)$ curves at the present moment, t_0 (Fig. 16.18). This line
shows the behavior of $a(t)$ for the case where there is no gravity in the universe
to slow the expansion. Since the slope of this line is $(da/dt)_{t_0} = a(t_0)/\tau$ and Hub-
ble's constant $H(t_0) = \dot{a}(t_0)/a(t_0)$, $\tau = 1/H$ is the age of the Universe in this no-
gravity case. τ **is called the Hubble time,** and we will show in the next section that
a flat universe with gravity has an age of $(2/3)\tau$ or $2/3H$. Since the effect of grav-
ity is to slow the expansion, as shown in Figure 16.18, an open universe has an
age between τ and $(2/3)\tau$, and a closed universe has an age less than $(2/3)\tau$.

In order to treat accelerations (current epoch) and decelerations (previous
epoch) of the expansion most generally, and to get a flavor of the equations of
general relativity describing an expanding Universe, we turn to the equations
themselves. One can show that general relativity requires the universal scaling
factor $a(t)$ to obey the differential equations:[12]

---

[12]See P. J. E. Peebles, *Principles of Physical Cosmology,* Princeton, NJ, Princeton University Press,
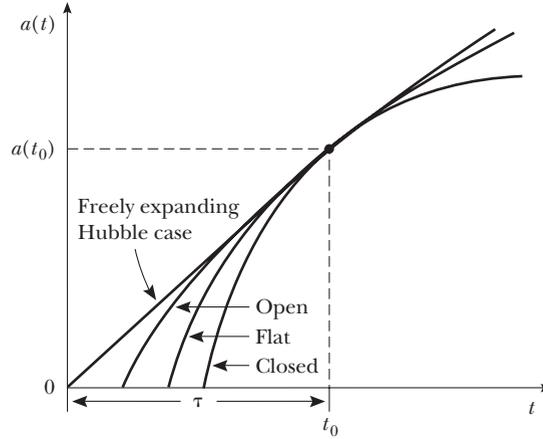1993, pp. 75 and 76.

Figure 16.18 The straight line is the tangent to the Freidmann $a(t)$ curves at the present moment and models the case where there is no gravitational attraction in the Universe. For this case, the age of the Universe is $1/H$. All the curves have steeper slopes than the tangent, indicating that these cases including gravity expanded at faster rates in the past and so reached the current size of the Universe more quickly. Consequently, open, flat, and closed universes have younger ages than the freely expanding Hubble case.

$$\left(\frac{da}{dt}\right)^2 = \frac{8\pi G\rho a^2}{3} - kc^2 \tag{16.15}$$

$$\frac{d^2a}{dt^2} = -\frac{4\pi aG}{3}\left(\rho + \frac{3P}{c^2}\right) \tag{16.16}$$

In these equations, a and ρ are functions of time, where ρ is the effective mass density of the Universe; that is, ρ includes the density of ordinary matter like hydrogen as well as the effective mass density of fields with energy density u, u/c^2. k is a parameter that determines whether the Universe is closed $(k = +1)$, open $(k = -1)$, or flat **($k = 0$, also called the Einstein–de Sitter case)**, P is the internal pressure, and the other symbols have their usual meanings. The first equation comes from the conservation of energy, and the second is a sort of Newton's second law, giving the rate of change of the expansion in terms of the effective mass density within the Universe. The unusual part of Equation 16.16 is the appearance of the internal pressure P of a substance that apparently adds to the attractive mass density of matter in the Universe. P appears because general relativity treats matter in the Universe on the average as an extremely dilute gas characterized by a mass density and an internal pressure, a measure of the gas's kinetic energy. The amount of effective mass density that P adds is much smaller than the contribution of the actual mass density of a substance under ordinary conditions since P is divided by c^2. For example, the contribution of the internal pressure of 1 cubic meter of air at standard temperature and pressure is about 10^{-12} of that of the mass density of the air. But it gets even more interesting. Freidmann considered the Universe to have $P = 0$ on average, which would be the case for a universe filled with cold, pressure-free dust particles. **In extreme cases, P can be positive and very large (inside stars and in the Big Bang) or even negative (for dark energy) and swamp out ρ to actually produce an acceleration of the expansion, as shown by Equation 16.16.** We will return to the dizzying subject of

the connection of mass density and internal pressure of different substances in the next section, but for now it suffices to realize that the term $(\rho + 3P/c^2)$ allows us to consider both attractive and repulsive forms of gravity in the Universe.

Equation 16.15 can be immediately solved for Hubble's constant in terms of fundamental parameters. Since $H^2(t_0) = \dot{a}^2(t_0)/a^2(t_0)$, 16.15 gives

$$H^2(t_0) = \frac{8\pi G\rho(t_0)}{3} - \frac{kc^2}{a^2(t_0)} \tag{16.17}$$

Equation 16.15 also may be used to show that the scale factor in the simple case of an Einstein–de Sitter Universe is proportional to $t^{2/3}$, where t is the time since the Big Bang. For this case, Equation 16.15 becomes

$$\frac{da}{dt} = \sqrt{\frac{8\pi G\rho}{3}}\, a \tag{16.18}$$

In order to separate variables and integrate this equation, we need to express $\rho(t)$ as a function of a. We take $t = 0$ at the singularity $a \rightarrow 0$ and $\rho \rightarrow \infty$ and assume a finite effective mass M in the first instants of the Universe. If the Universe simply expanded and no matter left or entered through its surface

$$M = \rho(t)\, \frac{4\pi a^3(t)}{3}$$

or

$$\rho(t) = 3M/4\pi a^3(t) \tag{16.19}$$

where $\rho(t)$ and $a(t)$ are the density and scaling factor at time t. Substituting 16.19 into 16.18 gives

$$\frac{da}{dt} = Ba^{-1/2} \tag{16.20}$$

where $B = \sqrt{2GM}$. Integrating 16.20 gives

$$\int_{a=0}^{a} a^{1/2}\, da = B\int_{t=0}^{t} dt$$

and finally,

$$a = Ct^{2/3} \tag{16.21}$$

where $C = \left(\frac{3B}{2}\right)^{2/3}$. Since $H = \dfrac{\dot{a}}{a}$, using 16.21 we find that the Einstein–de Sitter Universe has an age of $2/3H$, as claimed.

Although we have mentioned that light emitted from astronomical objects at increasingly distant time epochs means the light is increasingly redshifted as we see it now, Equation 16.21 gives the exact connection between redshift and past epoch for the Einstein–de Sitter model. The expansion of the universe between now (t_0) and the time of emission (t_e) of radiation now seen to be redshifted by Z is described by

$$1 + Z = \frac{a(t_0)}{a(t_e)} \tag{16.22}$$

For the Einstein–de Sitter model, this equation becomes

$$1 + Z = \frac{t_0^{2/3}}{t_e^{2/3}} \tag{16.23}$$

Thus, when we observe a supernova with a high Z of 2 in an Einstein−de Sitter universe, we are looking at historical conditions in a universe which was 1/3 of its present size and about 20% of its present age when the supernova detonated.

16.7   WILL THE UNIVERSE EXPAND FOREVER?

In a current mythology the ultimate fate of the Universe may be viewed as a cosmic struggle between two titans: the attractive gravitational force of matter and the repulsive force of dark energy spread uniformly through the vacuum. **The specific nature of dark energy or vacuum energy and its variation with time are unknown except that it has a negative pressure in our general relativity Equation 16.16 and so acts as a repulsive gravitational force.** In fact, the term $(\rho + 3P/c^2)$ in Equation 16.16 can be used to categorize the attractive or repulsive gravitational nature of substances both common and exotic, which are summed in this term:

Dark Energy

- For an ordinary gas at STP, the mass density term is 10^{12} times larger than the internal pressure term, both terms are positive, so the effect of gases is to produce gravitational attraction.
- For photons, which are mass 0 particles, it is convenient to write $(\rho + 3P/c^2) = 1/c^2(u + 3P)$, where u is the energy density or energy per unit volume of the photons. As the internal pressure of a photon gas is given by $P = u/3$, $1/c^2(u + 3P) = 2u/c^2$, a positive term, so that photons also exert a force tending to hold the Universe together.
- For cold dark matter, $P = 0$, and only the mass density of these exotic particles contributes to the attractive gravitational force.
- For all forms of *dark energy*, $P < 0$, and if P is sufficiently negative it will dominate all positive contributions and produce a net repulsive gravitational force in the Universe, which is believed to be the case now. For a "substance" called the quantum vacuum (to be discussed shortly) with an energy density u_{qv}, $P = -u_{qv}$, so $1/c^2(u + 3P) = -2u_{qv}/c^2$. The quantum vacuum produces a strongly repulsive force term that may dominate all attractive terms provided the quantum vacuum energy density is large enough. In fact, examination of Equation 16.16 shows that observation of accelerated expansion means dark energy exists. Denoting the energy density and pressure of dark energy by u_{de} and P_{de}, an observed acceleration of the expansion means $(\rho + 1/c^2(u_{de} + 3P_{de})) < 0$ or $P_{de} < -u_{de}/3$ for any form of dark energy. Different models of dark energy predict different time variations for u_{de} and different values for P_{de}, where P_{de} is limited as $-u_{de} \leq P_{de} \leq -u_{de}/3$.[13]

    It is very interesting that current observations show that dark energy actually exists and the effective mass density of dark energy, u_{de}/c^2, dominates the observed attractive mass density of the Universe Ω_m by a factor of about 2:1. The normalized dark energy density is expressed by writing $\Omega_\Lambda = u_{de}/u_c$, where u_{de} and u_c stand, respectively, for the observed dark energy and critical energy (mass) densities. When the constraint $\Omega_\Lambda + \Omega_m = 1$ from inflationary

---

[13]See the article by Bertram Schwarzchild, *Physics Today,* June 2004, p. 19.

theory is imposed on the most recent SN1e data, the best-fit values of $\Omega_m = 0.29 \pm 0.04$ and $\Omega_\Lambda = 0.71$ are found.[14]

The existence of dark energy had been suspected for several years, as well as a transition from an earlier epoch of deceleration to the current period of acceleration of the expansion. The transition was expected because the effective mass density term ρ falls as $1/a^3(t)$ as the Universe expands, and the $3P/c^2$ dark energy term is believed to be constant or more slowly decreasing than ρ. The actual observation of this expected switchover point provides conclusive evidence of the existence of dark energy and some information on how it changes with time. The latest results show that the transition from deceleration to acceleration occurred about 9 billion years ABB, or in terms of model-independent redshift, at a time corresponding to $Z_{transition} = 0.46 \pm 0.13$. The time dependence of the dark energy is described by the Z dependence of the dimensionless parameter w, the ratio of the dark energy's pressure to its energy density.

$$w = \frac{P_{de}}{u_{de}} \tag{16.24}$$

Writing

$$w(Z) = w_0 + \left.\frac{dw}{dZ}\right|_{t_0} Z \tag{16.25}$$

the best-fit values to the new supernovae data have

$$w(Z) = -1.31 + 1.48Z \tag{16.26}$$

To place these results in context, Einstein's cosmological constant has $w_0 = -1$ and $dw/dZ = 0$. If the current large value of $w = -1.31$ remains constant or becomes more negative with time, the Universe will not only continue to expand but the eventual unhappy fate of galaxies, solar systems, atoms, and nuclei will be to be ripped apart.

16.8   PROBLEMS AND PERSPECTIVES

While particle physicists have been exploring the realm of the very small, cosmologists have been exploring cosmic history back to the first moments of the Big Bang. Observation of events that occur when two particles collide in an accelerator is important for understanding the early moments in cosmic history, but perhaps the key to understanding the early Universe is to find a unified theory of quantum mechanics and gravity. Such a theory is urgently needed to adequately describe the first instants of the Big Bang. Cosmologists and particle physicists now find that they have many common goals and are joining hands to attempt to understand the physical world at its most fundamental level.

Lest the reader be fooled by our simplified overview of cosmology, it is important to realize that understanding is far from complete. There are many questions remaining concerning our incredible Universe so finely tuned that it has allowed the fragile and intricate complexities of life to arise 10 billion

---

[14]The latest SN1e results are from A. G. Riess et al., *Astrophys. J.* (in press), available at http://arXiv.org/abs/astro-ph/0402512

years ABB. The deepest questions concern the origin and observed size of the dark energy described by Ω_Λ. Current observations show that the dark energy density is surprisingly small, but exquisitely adjusted to just barely dominate the attractive mass density of the Universe—permitting stars, galaxies, planets, and people to form. Why surprisingly small? According to the most successful elementary particle model, the Standard Model, the quantum vacuum ought to be seething with fields and virtual particles constantly erupting into and disappearing from existence. All the associated energy and momentum of these fields and particles should produce a huge negative pressure or repulsive gravitation, yielding a dark energy density calculated to be as much as 10^{120} times the observed dark energy density. Leading physicists have identified this problem of dark energy as the most important one for 21st-century physics, likening it in importance to the problem of blackbody radiation in the 20th century. Just as joint efforts of experimentalists and theorists were needed to solve the blackbody problem, both new observations and theoretical insights will be needed to unravel the secret of dark energy. We end with the words of Nobel laureate Steven Weinberg concerning the importance of discovering the nature of dark energy:

> The task of discovering the nature of the "dark energy" is of obvious importance to cosmology. The apparent acceleration of the expansion of the universe is attributed to a dark energy residing in space itself, which also balances the kinetic energy of the expansion so as to give the universe zero spatial curvature, as observed in mapping fluctuations in the cosmic microwave radiation background. If the dark energy is a constant (the so-called cosmological constant) or growing, then the fate of the universe is sealed: it will continue expanding forever. If the dark energy is decreasing (as in some "quintessence" theories) then it was even more important in the past, and may have played a part in limiting the formation of the largest gravitationally bound structures. In any case, through its effect on the expansion of the universe, the dark energy affects all observations of astronomical objects at large redshift.
>
> The problem of the dark energy is also central to today's physics. Our best attempts at a fundamental theory suggest the presence of a cosmological constant that is many (perhaps as many as 120) orders of magnitude greater than the upper bound set by astronomical observations. For decades the problem seemed to be to find a symmetry or cancellation mechanism of some sort that would make the cosmological constant precisely zero. The single greatest failure of our most promising theories (such as string theories) is that they do not satisfy this requirement. Now that a dark energy has apparently been found, the problem is even harder: not just to explain why the dark energy is so tiny compared with what would have been expected theoretically, but also to explain why it happens to be of the same order of magnitude (roughly twice) as the energy in matter at the present moment in the history of the universe. It is difficult for physicists to attack this problem without knowing just what it is that needs to be explained—a cosmological constant or a dark energy that changes with time as the universe evolves—and for this they must rely on new observations by astronomers. Until it is solved, the problem of the dark energy will be a roadblock on our path to a comprehensive fundamental physical theory.

— Steven Weinberg
Department of Physics
University of Texas at Austin

SUGGESTIONS FOR FURTHER READING

1. In addition to the magazine articles cited in the footnotes the following articles are interesting, but note the dates for the most current results:

G. Gamow, "The Evolutionary Universe," *Sci. Amer.,* September 1956.

G. Gamow, "Gravity," *Sci. Amer.,* March 1961.

A. R. Sandage, "The Red-Shift," *Sci. Amer.,* September 1956.

C. Sagan and F. Drake, "The Search for Extraterrestrial Intelligence," *Sci. Amer.,* May 1975.

F. Wilczek, "The Cosmic Asymmetry Between Matter and Anti-Matter," *Sci. Amer.,* December 1980.

G. Veneziano, "The Myth of the Beginning of Time," *Sci. Amer.,* May 2004, pp. 54–65.

J. D. Bekenstein, "Information in the Holographic Universe," *Sci. Amer.,* August 2003, pp. 58–65.

A special report of four articles on cosmology in the February 2004 *Scientific American*: W. Hu and M. White, "The Cosmic Symphony," pp. 44–53; M. A. Strauss, "Reading the Blueprints of Creation," pp. 54–61; A. G. Riess and M. S. Turner, "From Slowdown to Speedup," pp. 62–67; and G. Dvali, "Out of Darkness," pp. 68–73.

2. Several good nonmathematical books are:

J. M. Pasachoff, *Astronomy: From the Earth to the Universe,* 6th ed., Brooks/Cole-Thomson Learning, Belmont, CA, 2002.

A. H. Guth, *The Inflationary Universe,* Helix Books, Addison-Wesley, Reading, MA, 1997.

L. Krauss, *Quintessence,* Basic Books, New York, 2000.

M. Rees, *Just Six Numbers,* Basic Books, New York, 2000.

3. Two books at the advanced undergraduate level are:

P. J. E. Peebles, *Principles of Physical Cosmology,* Princeton University Press, Princeton, NJ, 1993.

E. V. Linder, *First Principles of Cosmology,* Prentice Hall, 1997.

PROBLEMS

1. $Z \equiv \dfrac{\Delta\lambda}{\lambda} = \dfrac{\lambda - \lambda_0}{\lambda}$ is the symbol used by astronomers to denote the measured redshift of receding astronomical objects where λ_0 is the emission wavelength of a particular line measured in the lab and λ is the wavelength of the same line measured for a receding source. For example, type Ia supernovae used as distance markers in expansion rate studies have Z's as high as 2, galaxies have been observed with Z's in excess of 3, and quasars (star-like objects formed early in the Universe with enormous energy output) take the cake with Z's greater than 5. (a) Use the relativistic Doppler shift Equation 16.1 to show that the relation between red shift and recession velocity v, valid at all velocities, is

$$Z = \sqrt{\frac{1+\beta}{1-\beta}} - 1$$

where

$$\beta = \frac{v}{c}.$$

(b) Plot a graph of Z versus β for values of Z ranging from zero to five.

2. Show that the relative velocity of galaxies is directly proportional to their separation in a homogeneous, uniformly expanding Universe. Consider three collinear, equally spaced galaxies G_0, G_1, G_2, which are adjacent galaxies spaced L apart. An observer situated on the leftmost galaxy G_0 will observe the distances to G_1 and G_2 to be L and $2L$ and, for an expanding universe, G_1 and G_2 will have instantaneous velocities v_1 and v_2 away from G_0. At a time Δt later, G_1 will be a distance $L + v_1\Delta t$ away from G_0 and G_2 a distance $2L + v_2\Delta t$. **If the Universe is to remain homogeneous, all distances between galaxies must retain the same proportions as expansion proceeds.** This means we require the distance G_0G_2 to remain twice as large as G_0G_1 at time Δt. Thus, $2L + v_2\Delta t = 2(L + v_1\Delta t)$ or $v_2 = 2v_1$. Show that Hubble's law results from generalizing to p collinear galaxies with adjacent distance L, provided that all velocities are much less than c.

3. A quasar (believed to be a galaxy in formation with a super-massive black hole at its center) is a pointlike object that has tremendous energy output, is very distant from the Earth, and has a large redshift. Its speed can be measured from Doppler-shift measurements of the light it emits. A certain quasar recedes from the Earth at a speed of $0.55c$. (a) How far away is it? (b) Assuming that the quasar has moved with the speed $0.55c$ ever since the Big Bang, estimate the age of the Universe.

4. (a) Show that if Hubble's constant H is truly constant in time, the universe is undergoing inflationary expansion, that is $a(t) = Ae^{bt}$.

(b) Find H as a function of time for the case of an Einstein–de Sitter Universe.

5. Find the size and age of an Einstein–de Sitter Universe probed by $Z = 5$ quasars.

6. Using Hubble's law, estimate the wavelength of the 590-nm sodium line emitted from galaxies (a) 2×10^6 lightyears away from Earth, (b) 2×10^8 lightyears away, and (c) 2×10^9 lightyears away. (*Hint:* Use the relativistic Doppler formula for light emitted from a receding source):

$$\lambda' = \frac{\sqrt{1 + (v/c)}}{\sqrt{1 - (v/c)}} \lambda \qquad (16.1)$$

7. A distant quasar is moving away from Earth at such high speed that the blue 434-nm hydrogen line is observed at 650 nm, in the red portion of the spectrum. (a) How fast is the quasar receding? (b) Using Hubble's law, determine the distance from Earth to this quasar.

8. Hubble's law can be stated in vector form as $v = H\mathbf{R}$: Outside the local group of galaxies, all objects are moving away from us with velocities proportional to their displacements from us. In this form, it sounds as if our location in the Universe is specially privileged. Prove that Hubble's law would be equally true for an observer elsewhere in the Universe. Proceed as follows: Assume that we are at the origin of coordinates, that one galaxy cluster is at location $\mathbf{R}_1$ and has velocity $v_1 = H\mathbf{R}_1$ relative to us, and that another galaxy cluster has radius vector $\mathbf{R}_2$ and velocity $v_2 = H\mathbf{R}_2$. Suppose the speeds are nonrelativistic. Consider the frame of reference of an observer in the first of these galaxy clusters. Show that our velocity relative to her, together with the displacement vector of our galaxy cluster from hers, satisfies Hubble's law. Show that the displacement and velocity of cluster 2 relative to cluster 1 satisfy Hubble's law.

9. Use the Boltzmann distribution function $e^{-E/k_B T}$ to calculate the temperature at which 1% of a population of photons will have energy greater than 1.00 eV. The energy required to excite an atom is on the order of 1 eV. Thus, as the temperature of the Universe fell below the value you calculate, neutral atoms could form from plasma, and the Universe became transparent. The cosmic background radiation represents our vastly redshifted view of the opaque fireball of the Big Bang as it was at this time and temperature. The fireball surrounds us; we are embers.

10. The cosmic background radiation is blackbody radiation at a temperature of 2.73 K. (a) Use Wien's law to determine the wavelength at which this radiation has its maximum intensity. (b) In what part of the electromagnetic spectrum is the peak of the distribution?

11. It is mostly your roommate's fault. Nosy astronomers have discovered enough junk and clutter in your dorm room to constitute the missing mass required to close the Universe. After observing your floor, closet, bed, and computer files, they extrapolate to slobs in other galaxies and calculate the average density of the observable Universe as $1.20\rho_c$. How many times larger will the Universe become before it begins to collapse? That is, by what factor will the distance between remote galaxies increase in the future?

12. The early Universe was dense with gamma-ray photons of energy $\sim k_B T$ and at such a high temperature that protons and antiprotons were created by the process $\gamma \rightarrow p + \bar{p}$ as rapidly as they annihilated each other. As the Universe cooled in adiabatic expansion, its temperature fell below a certain value, and proton pair production became rare. At that time slightly more protons than antiprotons existed, and essentially all of the protons in the Universe today date from that time. (a) Estimate the order of magnitude of the temperature of the Universe when protons condensed out. (b) Estimate the order of magnitude of the temperature of the Universe when electrons condensed out.

13. If the average density of the Universe is small compared to the critical density, the expansion of the Universe described by Hubble's law proceeds with speeds that are nearly constant over time. (a) Prove that in this case the age of the Universe is given by the inverse of Hubble's constant. (b) Calculate $1/H$ and express it in years.

14. The gravitational self-energy of an object is the energy needed to assemble the object from individual elements initially spaced an infinite distance apart. The energy needed is the work done by an external force in an equilibrium process, that is an element of mass is moved so that the external force is equal and opposite to the gravitational force of attraction. Because the external force is opposite to the element's displacement, the work done or the gravitational self-energy is a negative quantity. The gravitational self-energy of N discrete mass elements is given by the sum of the potential energy of all pairs of elements and may be written

$$U = -G \sum_{\substack{\text{all pairs} \\ i \neq j}} \frac{m_i m_j}{r_{ij}}$$

An equivalent form is

$$U = -\frac{1}{2} G \sum_{\substack{i=1 \\ i \neq j}}^{N} \sum_{j=1}^{N} \frac{m_i m_j}{r_{ij}}$$

where we count each pair twice but correct with the factor of 1/2. (a) Use this expression to calculate the self-energy of the Milky Way Galaxy. Our Galaxy consists of $N = 2 \times 10^{11}$ stars with an average separation of $r_{ij} = 10^{21}$ m and mass $m_i = 2 \times 10^{30}$ kg.
(b) In the first part of this problem we did not consider the self-energy of individual stars. Remedy this problem by showing that the gravitational self-energy of a uniform sphere of mass m and radius r is

$$U_{\text{sphere}} = -\frac{3Gm^2}{5r}$$

(Hint: assemble the sphere from infinitesimal shells of thickness dr surrounding a solid spherical core of radius r.)
(c) Find the self-energy of the Sun if $m_s = 2 \times 10^{30}$ kg and $r_s = 7 \times 10^8$ m.

15. *Finding the temperature inside the Sun with the virial theorem.* The virial theorem is a powerful result which applies to the time average of the kinetic and potential energies of a group of particles confined within a finite volume by mutual inverse-square forces. The virial theorem may be written

$$\langle KE_{\text{total}} \rangle = -\frac{1}{2} \langle PE_{\text{total}} \rangle$$

where the brackets indicate time averages over sufficiently long times. This result holds even if all particle masses are not the same and if some of the particles simultaneously experience repulsive electric forces. The theorem basically says that a gravitationally bound system of particles, even a system starting with all particles at rest, ultimately develops a *KE* equal to one half of the absolute value of its *PE*.

In a typical star like the Sun, the average distance between atomic nuclei is usually greater than 10^{-12} m, so attractive nuclear forces (range $\approx 10^{-15}$ m) do not hold the Sun together. Rather, all stars are held together by gravity so the virial theorem applies. (a) Find an expression for the average temperature inside the Sun from the virial theorem using $\langle KE$ of a particle$\rangle = 3/2 \, k_{\text{B}} T$ and $\langle PE$ Sun$\rangle = -3 G m_{\text{S}}^2 / 5 r_{\text{S}}$ as shown in problem 14. Let m_{S} and r_{S} stand for the mass and radius of the Sun. Let N be the number of atoms in the Sun and $m = m_{\text{S}}/N$ be an average particle mass in the Sun.

(b) Find T_{av} for the Sun assuming $m_{\text{s}} = 2 \times 10^{30}$ kg, $r_{\text{s}} = 7 \times 10^8$ m, and that 60% of the Sun's mass is hydrogen, the rest helium.

16. *Collapse of a dust ball under Newtonian gravity.* Suppose that a spherical ball of dust with no internal pressure and with an initial radius R_0 and total mass M gravitationally collapses from rest. (a) Show that the outer shell of the collapsing dust ball obeys the equation of motion

$$\ddot{R} = -\frac{GM}{R^2}$$

(b) Solve this equation of motion and show that the time to collapse is given by

$$t_{\text{c}} = \frac{\pi}{2} \sqrt{\frac{R_0^3}{2GM}}$$

Hint: Write Newton's law as $\ddot{R} = f(R)$. Let $f(R) = dF(R)/dR$ and multiply both sides of $\ddot{R} = dF/dR$ by $\dot{R}$. The resulting equation may be immediately integrated with respect to time by recognizing that $\ddot{R}\dot{R} = \frac{d}{dt}\left(\frac{\dot{R}^2}{2}\right)$. Integrating again with respect to time, this general technique yields t as a function of R where $t = -\int \frac{dR}{\sqrt{2(F(R) + C)}} + C_1$ and C and C_1 are constants determined by the initial conditions.

(c) Find the time of collapse of a dust ball with a mass and radius equal to that of the Sun.
($R_{\text{S}} = 6.96 \times 10^{10}$ cm, $M_{\text{S}} = 1.99 \times 10^{33}$ g, $G = 6.67 \times 10^{-8}$ cgs units)
(d) What prevents the Sun from collapsing in this way?