14 EXPANSION OF THE UNIVERSE

The theory of relativity brought the insight that space and time are not merely the stage on which the piece is produced, but are themselves actors playing an essential part in the plot.
Willem de Sitter, "The expanding universe" (1931)

THE GREAT DISCOVERY

Doppler effect

From a historical viewpoint the Doppler effect paved the way to the discovery of the expanding universe. Nowadays we do not use the Doppler effect in cosmology, except in its classical Fizeau–Doppler form as a rough and ready guide. We examine the Doppler effect briefly and defer to Chapter 15 a more searching inquiry.

The spectrum of light from a luminous source contains bright and dark narrow regions, as shown in Figure 14.1, that are the emission (bright) and absorption (dark) lines produced by atoms. When a luminous source such as a candle or a star moves away from an observer, all wavelengths of its emitted radiation, as seen by the observer, are increased. Its spectral lines are moved toward the longer wavelength (redder) end of the spectrum and it is said to have a redshift. This redshift is detected by comparing the spectrum of the luminous source with the spectrum of a similar source that is stationary relative to the observer. The source may move away from the observer, or the observer may move away from the source, and in either case the separating distance increases and there is an observed redshift.

When the luminous source moves toward the observer, all wavelengths of its emitted radiation, as seen by the observer, are decreased. Its spectral lines are moved toward the shorter wavelength (bluer) end of the spectrum and it has a blueshift. The source may move toward the observer, or the observer may move toward the source, and in either case the separating distance decreases and there is an observed blueshift.

The shift in the observed wavelengths can be expressed in terms of the relative velocity by means of the classic Fizeau–Doppler relation (Figure 14.2). Let V be the relative velocity of a luminous source moving away; also let λ be the wavelength of an emitted ray of light and λ_0 the wavelength of the same ray received by the observer. According to the Fizeau–Doppler formula,

$$\frac{\text{observed wavelength}}{\text{emitted wavelength}} = \frac{\lambda_0}{\lambda} = 1 + \frac{V}{c}, \quad [14.1]$$

where c is the speed of light. The wavelength of a line in the spectrum of the light from the source, as seen by the observer, is measured and compared with the wavelength of the same line emitted by atoms in the observer's laboratory. The relative velocity V of the source is determined from the fractional difference in the emitted and received wavelengths:

$$\frac{V}{c} = \frac{\lambda_0 - \lambda}{\lambda}, \quad [14.2]$$

given by Equation [14.1]. But this classic formula can be used only when V is very much less than c (V less than $0.01c$). For higher velocities we must use a formula (referred to as the special relativity Doppler formula or just the Doppler formula) derived from special relativity theory and given in the next chapter (Chapter 15). Notice that when the source moves toward

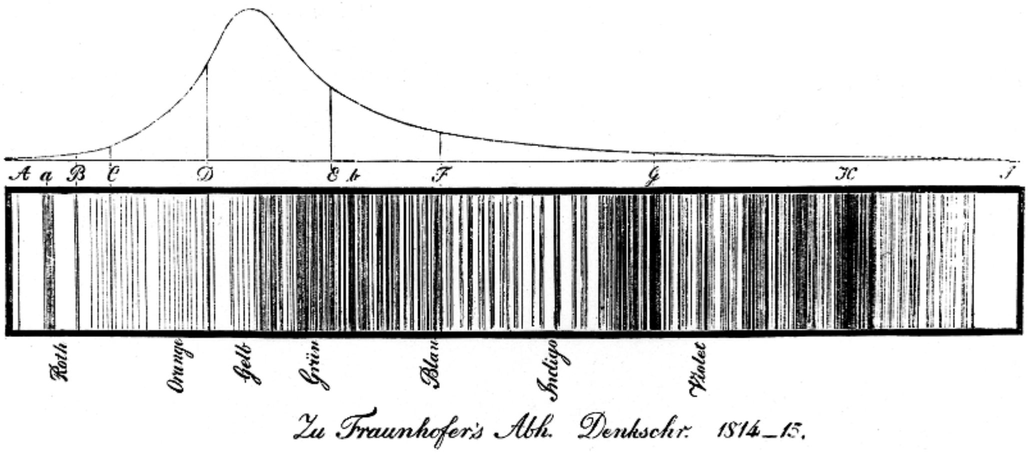

Figure 14.1. A spectrum of light from the Sun in which wavelengths increase from violet at the right to red at the left. This spectrum shows the dark absorption lines, as observed in 1814 by the German physicist Joseph von Fraunhofer. These lines are now known as Fraunhofer lines. The curve on the top shows that the intensity of sunlight peaks between yellow and green.

the observer the velocity V changes sign and becomes negative, and we can still use Equations [14.1] and [14.2]. The redshift of a source is defined as the fractional increase in wavelength:

$$z = \frac{\lambda_0 - \lambda}{\lambda} \qquad [14.3]$$

and is always denoted by z. From Equations [14.2] and [14.3] we find

$$V = cz. \qquad [14.4]$$

Thus a redshift of 0.01 means that the source is receding at 1 percent of the velocity of light, or 3000 kilometers per second.

Slipher's celestial speed champions

Vesto Slipher, an astronomer at the Lowell Observatory at Flagstaff, Arizona, in 1912 began to measure the shift in the spectral lines of light received from spiral nebulae. By 1923, as a result of his painstaking measurements, it was known that of the 41 galaxies studied, 36 had redshifts, and the other five, which included the Andromeda Nebula, had blueshifts. His measured redshifts multiplied by the velocity of light indicated recession velocities of many thousands of kilometers a second. The discovery of galaxies moving with these amazingly high velocities received wide publicity in the press. Slipher's results were also surprising because, if galaxies had random motions, one would expect that those moving away with redshifts would be approximately equal in number to those approaching with blueshifts. His observations revealed that most galaxies were moving away from us.

Discoverers of the expanding universe

The discovery of the expanding universe did not occur abruptly. To unfold the historical record we must anticipate certain developments that will be clearer in later chapters.

The first intimation of an expanding universe came in 1917 in the work of the Dutch astronomer-cosmologist Willem de Sitter. On the basis of theoretical studies, he predicted the existence of a "systematic displacement of spectral lines toward the red" in the light received from distant nebulae. There is little doubt that Slipher's redshift measurements and de Sitter's studies gave birth to the idea of an expanding universe.

In 1917, Einstein and de Sitter proposed two different kinds of universe, both based on Einstein's theory of general relativity that had been published in its final form

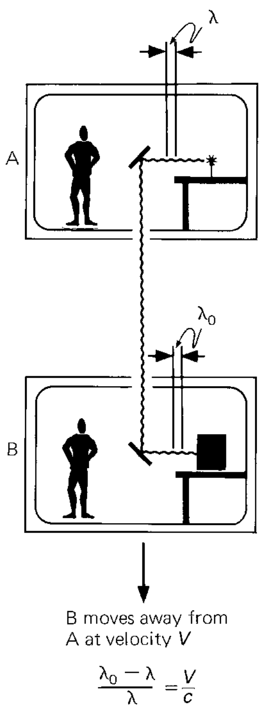

Figure 14.2. Observers A and B are in separate laboratories that move apart at velocity V. Both study the light emitted by atoms in their laboratories and find their emission and absorption lines have identical wavelengths in the two laboratories. Let λ be the wavelength of one of these spectral lines. A sends light to B, who finds that the wavelength λ of the transmitted line is received at λ_0. The classical Fizeau–Doppler formula states that the fractional increase $(\lambda_0 - \lambda)/\lambda$ is equal to V/c, where c is the velocity of light.

the previous year. Einstein's universe was uniform: it contained uniformly distributed matter and had uniformly curved spherical space. The main feature of Einstein's model of the universe was its static nature; it was unchanging, neither expanding nor collapsing. We must remember that at that time the astronomical universe was believed to be static on the cosmic scale. To conform with this belief, and to enable his universe to maintain a static state, Einstein introduced a cosmological constant into the theory of general relativity. The cosmological constant is equivalent to a repulsive force that opposes the force of gravity. By adjusting the value of the cosmological constant it is possible to make the repulsion counterbalance the gravity due to the uniform distribution of matter. The cosmological constant, denoted by Λ and sometimes called the lambda constant, was introduced in a rather *ad hoc* way, and Einstein sought to justify its use by appeal to Mach's principle. In his static universe, the local mass density was related directly to the cosmological constant, and Einstein believed that this method of tying together the local and global realms was in accord with Mach's philosophy.

But in the same year came the completely different de Sitter universe. It incorporated the cosmological constant, it was assumed to be static, and unlike the Einstein universe it contained no matter. This alternative universe, derived from general relativity theory, showed clearly that the cosmological constant did not guarantee a unique universe, as Einstein had hoped. The empty de Sitter universe might have been ignored as a curious freak were it not for one arresting property: when particles are sprinkled in it, they accelerate away from one another. It was thought that this "de Sitter effect," as it became known, has perhaps some bearing on the recessions and redshifts observed by Slipher.

The astronomer Carl Wirtz, inspired by Slipher's redshift measurements and the de Sitter effect, proposed in 1922 a velocity–distance relation. He used the apparent diameters of galaxies as distance indicators (the larger the distance the smaller the average apparent diameter) and the Fizeau–Doppler formula of Equation [14.4] and found that the recession velocity V increased with distance.

The apparent static nature of the de Sitter universe was a mathematical fiction. This universe appeared to be static because it contained nothing that could exhibit its actual dynamic state. Howard Robertson, a mathematician, later showed that a simple readjustment in the distinction between space and time made the de Sitter universe spatially homogeneous and flat, and in this form it was found to be expanding. An apt distinction then emerged: the Einstein universe was "matter without motion," and the de Sitter universe was "motion without matter."

Fundamental theoretical developments by the Russian scientist Alexander Friedmann in 1922 and 1924, and the Belgian cosmologist Georges Lemaître in 1927 and 1931, opened wide the door to a large class of candidate universes, all homogeneous, isotropic, expanding, and all containing matter.

Hubble: measurer of the universe

Edwin Hubble was an exacting observer who made many discoveries at the 100-inch Mount Wilson telescope. We have recounted how Hubble classified the galaxies by their appearance and how he showed that the Andromeda Nebula is a galaxy beyond the Milky Way. He developed into a fine art the distance-measuring techniques pioneered by Harlow Shapley and in 1924 began to determine the distances of galaxies. In 1929 he announced the momentous discovery that the redshifts of galaxies tend to increase with their distances. From this redshift–distance relation, and from the Fizeau–Doppler formula of Equation [14.4], he arrived at a velocity–distance law stating that the recession velocity of galaxies increases with distance. The result seemed incredible: the observable universe is expanding! Figures 14.3 and 14.4 show Hubble's results.

In the following years, Hubble extended his distance measurements. With redshifts determined by Milton Humason, Hubble made secure the concept of an expanding universe. By 1955, using the 200-inch telescope of Mount Palomar, Humason, Nicholas Mayall, and Allan Sandage had observed and analyzed the spectra of more than 800 galaxies and detected redshifts up to 0.2.

But Hubble was not the first to discover the expansion of the universe, nor did he ever claim that he had made such a discovery, and in *The Realm of the Nebulae*

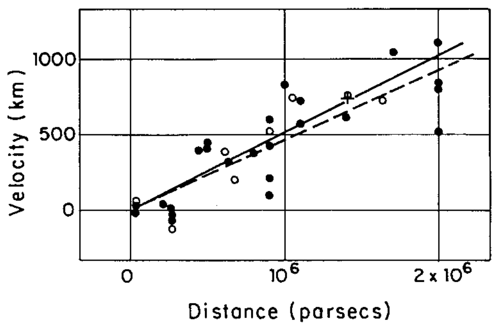

Figure 14.3. The recession velocity in kilometers per second of extragalactic nebulae plotted against their distances in parsecs, taken from Hubble's 1929 paper. The velocity used by Hubble is *cz* from the Fizeau–Doppler formula.

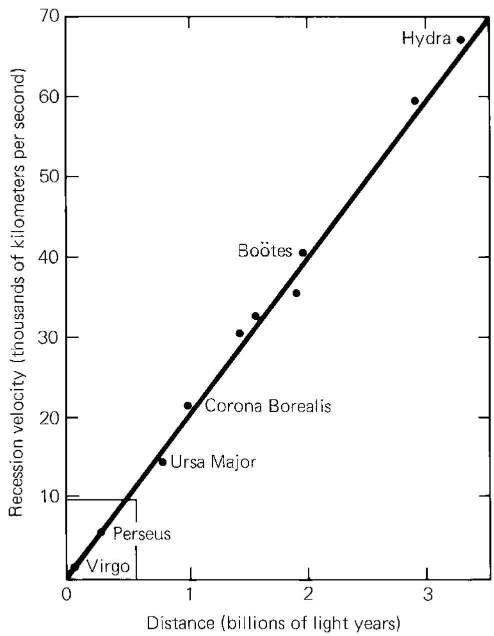

Figure 14.4. A more recent velocity–distance diagram showing the expansion of the universe. The lower left corner is the region surveyed by Hubble shown in Figure 14.3.

published in 1936 he expressed reservations concerning the velocity interpretation of extragalactic redshifts. Many astronomers since the time of Hubble have viewed with reservation the velocity–redshift relation of Equation [14.4]. This relation is valid for only very small values of redshift and, as we shall see, their reservations have been justified. Howard Robertson in 1928 showed that Slipher's redshifts and Hubble's previously published distances supported an approximate redshift–distance relation,

$$zc = \text{constant} \times \text{distance},$$

and this relation in the form $zc = H \times$ distance, or

$$zc = HL, \quad [14.5]$$

where L is the distance of the galaxy, is now known as the Hubble law, where the constant H is called the Hubble term. The Hubble term changes in time and its value at the present cosmic epoch is denoted by H_0. Using a Doppler-like formula, such as Equation [14.4], Robertson derived the velocity–distance law,

$$V = HL. \quad [14.6]$$

According to this law, the more distant a galaxy, the faster it moves away. When distance is doubled, the recession velocity V is also doubled. Robertson's "rough verification," as he called it, was tucked away in a theoretical paper in a physics journal not widely read by astronomers.

At this stage in the development of modern cosmology it had yet to be realized that the velocity–distance law of Equation [14.6] is rigorously true at all distances in all expanding uniform universes. This realization emerged in the context of the expanding space paradigm that was established in the mid-1930s.

Decline of the Hubble term

Both Robertson in 1928 and Hubble in 1929 found for H a value of 150 kilometers a second per million light years (or nearly 500 kilometers a second per megaparsec). In other words, the recession velocity increased with distance by 150 kilometers per second for every million light years (or 500 kilometers per second for every megaparsec).

In 1952, Walter Baade discovered the two stellar populations. We have seen how his discovery showed that the distances of galaxies had previously been greatly underestimated. A second revision in distances came in 1958 when Allan Sandage discovered that what had previously been supposed to be bright stars in more distant galaxies were in fact very luminous regions of hot gas, and these more remote galaxies were therefore at even greater distances than previously supposed. These revisions in distance estimates have enlarged the scale of the universe and reduced the original value of H by a factor between 5 and 10. Most estimates nowadays place the value of the Hubble term between 15 and 30 kilometers a second per million light years (or 50 and 100 kilometers per megaparsec). A main theme in the history of twentieth-century cosmology has been the progressive decline in the value of the Hubble term determined by astronomers. The uncertainty in H stems from the extraordinary difficulty of measuring the distances of remote extragalactic systems. The Hubble term is customarily expressed in the form:

$$H_0 = 100h \text{ kilometers a second per megaparsec}, \quad [14.7]$$

or roughly

$$H_0 = 30h \text{ kilometers a second per million light years}, \quad [14.8]$$

and the zero subscript denotes the present cosmic epoch. The parameter h lies probably somewhere between 0.5 and 1. The Hubble term is everywhere the same in uniform space but varies in time. For this reason, in this book, we avoid using the confusing term "Hubble constant." We are not sure of the exact value of H_0 and shall sometimes assume that $h = 0.5$ and hence $H_0 = 50$ kilometers a second per megaparsec, or roughly 15 kilometers a second per million light years. Thus, at a distance of 1 billion light years, the recession velocity is 15 000 kilometers a second, or one-twentieth the velocity of light.

Two laws with the same name

The *redshift–distance law* $zc = HL$, Equation [14.5], is the observers' linear law first established by Slipher's redshift measurements and Hubble's distance determinations. Its proper name is the *Hubble law*. From the time of its discovery most cosmologists have realized that in its linear form it is only approximately true.

On the other hand, the *velocity–distance law* $V = HL$, Equation [14.6], is the theorists' linear law that follows automatically from the assumption that expanding space is uniform (isotropic and homogeneous). This law, often improperly referred to as the Hubble law, is of central importance in modern cosmology and is rigorously true in all uniform universes. At any moment in cosmic time the velocity of recession of the galaxies increases linearly with their distance. Considerable confusion exists because the approximate redshift–distance law and the exact velocity–distance law are indiscriminately referred to as the Hubble law.

The connecting link between the two laws is the linear velocity–redshift relation $V = cz$ from the Fizeau–Doppler formula of Equation [14.4]. This formula is an approximation only, thus explaining why Hubble's redshift–distance law is also only an approximation, valid for small redshifts of z much less than unity. This will become clearer in the next chapter. No linear velocity–redshift relation exists that is true for all redshifts in all universes, and the correct velocity–redshift relation must be derived from basic principles for each universe.

The technical definitions of recession velocity and distance were ambiguous, and the relation between recession velocity and redshift was obscure, until the mid-1930s. To this day they remain ambiguous and obscure to those not actively engaged in cosmological research. In the following, we introduce the *expanding space paradigm*, and define recession velocity, distance, and various other terms by means of imaginary experiments with an expanding rubber sheet.

THE EXPANDING SPACE PARADIGM

The expanding space paradigm emerged in the 1930s amid much controversy concerning the meaning of extragalactic redshifts. Arthur Eddington in "The expansion of the universe" in 1931 enunciated the paradigm when he wrote of the galaxies, "it is as though they were embedded in the surface of a rubber balloon which is being steadily inflated." Slowly emerged the idea that the universe consists of expanding space! The lesson we must learn from general relativity is that space can be dynamic as well as curved.

According to the expanding space paradigm, the universe does not expand in space, instead it consists of expanding space. From this statement flows all the simplicity and complexity of modern cosmology. The galaxies do not move through space, but instead float stationary in space. Their separating distances increase because the space between the galaxies expands. The paradigm helps us to understand the velocity–distance law, and also, in the next chapter, the nature of cosmological redshifts.

THE EXPANDING RUBBER SHEET UNIVERSE

The time has come to introduce ERSU – short for "Expanding Rubber Sheet Universe" – a make-believe two-dimensional flatland with which we shall perform imaginary experiments to illustrate the properties of an expanding universe.

A two-dimensional model has one drawback. It consists of a surface that expands in three-dimensional space. Our universe of three-dimensional space does not expand in a universe of higher-dimensional space, instead it consists of expanding space. This can be deduced from the containment principle: the universe is not in space but contains space.

In our imaginary experiments we stand back and with a godlike view survey the surface as it is everywhere at an instant in time. We see the universe in much the same way as seen by the cosmic explorer (Chapter 9) who moves from place to place instantaneously. But occasionally we must quit our godlike view and take the wormlike view of two-dimensional observers living in the surface.

Because there is no cosmic edge we must imagine an expanding surface that is indefinitely large. Instead of an expanding sheet we may imagine, if we wish, a spherical balloon that is steadily inflating, as suggested by Eddington. A flat surface serves just as well and in some ways is simpler.

Experiment 1: Dilation, shear, and rotation

We start by supposing that the flat surface exhibits every kind of two-dimensional motion consistent with flatness. On the surface we draw a large number of triangles and notice how in the course of time the triangles change in size, shape, and orientation. Three basic kinds of motion exist: dilation, shear, and rotation.

Those triangles that only dilate (or contract) exhibit shape-preserving motions, and the regions they occupy expand (or contract) isotropically (see Figure 14.5).

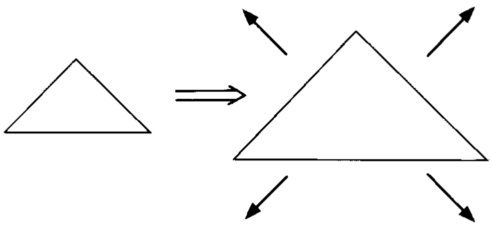

Figure 14.5. Dilation only: triangles change their area but not their shape and orientation.

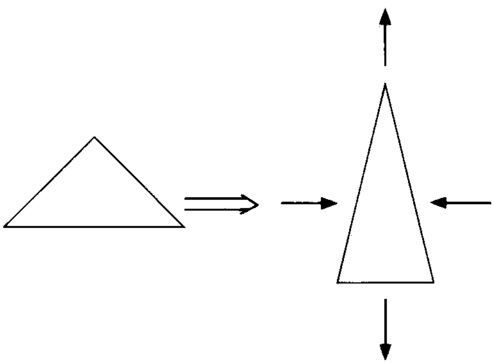

Figure 14.6. Shear only: triangles change their shape but not their area and orientation.

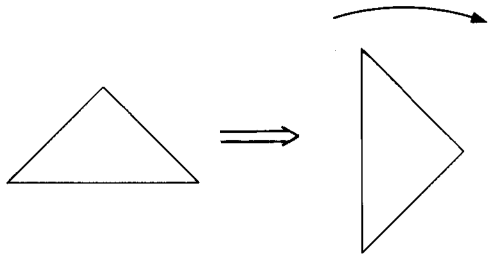

Figure 14.7. Rotation only: triangles change their orientation but not their area and shape.

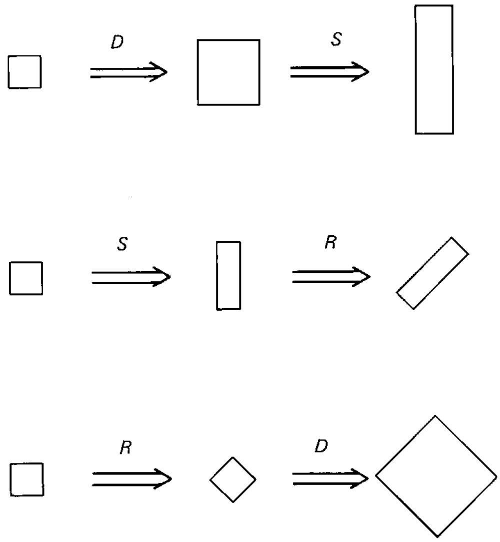

Figure 14.8. Combinations of dilation (D), shear (S), and rotation (R).

Those triangles that only shear exhibit area-preserving motions, and in the regions they occupy there is no dilation or rotation (see Figure 14.6).

Those triangles that only rotate exhibit both shape-preserving and area-preserving motions, and in the regions they occupy there is no dilation or shear (see Figure 14.7).

In general, complex motions combine dilation, shear, and rotation, and when all three vary from place to place, we have inhomogeneous motion (see Figure 14.8).

Experiment 2: Homogeneity

We draw identical triangles everywhere on the surface at a particular moment in time.

At this particular moment, ERSU represents a universe of homogeneous constitution. The preservation of this homogeneity requires that all triangles change in the same way and by the same amount in each interval of time. This means the motion of the surface is also homogeneous and has at every place in space at a common instant in time the same amount of dilation, shear, and rotation. We have now extended the meaning of homogeneity to include all forms of motion. A homogeneous universe in which all places are alike remains homogeneous when the motion is also homogeneous. We are of course ignoring the small-scale irregularities of astronomical systems.

A homogeneous universe having dilation only is shape-preserving; it expands (contracts) equally in all directions, and its motion is isotropic as well as homogeneous. This means, in ERSU, that all triangles increase (decrease) in size and retain their shapes and orientations. A homogeneous (all places are alike) and isotropic (all directions are alike) universe displays the simplest kinematics: shearfree and irrotational. We have previously referred to such a universe as uniform.

A homogeneous universe having only shear is area-preserving; it expands unequally in different directions, and its motion is anisotropic. This means, in ERSU, that all triangles change their shapes but preserve their areas: circles become ellipses and squares become oblongs. The observed highly isotropic cosmic background radiation shows that our dilating universe has little or no shear.

A homogeneous universe having only rotation is shape-preserving and area-preserving. Dilation and shear are zero, and in ERSU triangles preserve their shapes and areas but change their orientations. Because the motion is homogeneous, the rotation is not about a single central point but about all points in the surface. Remember, as observers we are in the surface and must not think of three-dimensional space. Rotation of our three-dimensional universe of space is not easy to imagine. It is the same everywhere and consists of anisotropic motion about one of three perpendicular axes. Such rotation can be detected by shooting a particle at a distant target and noticing that its trajectory curves away from the target. The compass of inertia does not rotate with the universe (as Mach claimed); if it did, we would not know if the universe were rotating or not. The observed highly isotropic cosmic background radiation shows that the universe has little or no rotation.

Experiment 3: Uniformity and cosmic time

When the expansion is homogeneous and isotropic, triangles in ERSU dilate everywhere in the same way, and do not change their shape or rotate. In such uniform expansion, figures change everywhere in space in exactly the same way and preserve their form (shape) in time. In subsequent experiments we shall assume that ERSU expands uniformly.

Homogeneity of the universe also means that all clocks in the universe – apart from timekeeping variations owing to local irregularities – agree in their intervals of time. With our godlike view we see clocks everywhere ticking away in constant agreement; but as denizens of flatland, with only a wormlike view, we must summon the explorer to our aid, who goes around the universe at infinite speed adjusting all clocks to show a common time. On subsequent tours the explorer finds the clocks running in synchronism, showing the same time. This universal time is known as *cosmic time*. All local departures from cosmic time are due to the motions and gravitational fields of individual astronomical systems. Because we are disregarding local irregularities, we shall disregard also local irregularities in time.

Experiment 4: Drawn circles are not "galaxies"

With chalk we draw a circle on the surface of the uniformly expanding ERSU and declare that it represents a galaxy (see Figure 14.9). As the sheet expands we observe that the "galaxy" gets bigger. This result is

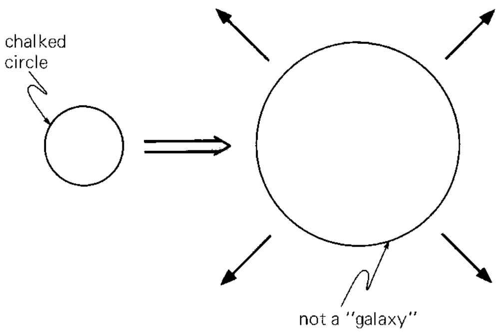

Figure 14.9. A circle drawn on the surface of the expanding rubber sheet also expands and cannot therefore represent a galaxy.

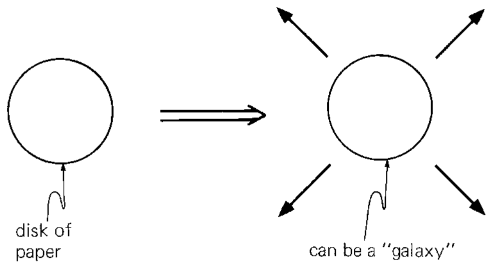

Figure 14.10. A disk of paper retains its size when placed on the surface of the expanding sheet and therefore can represent a galaxy of fixed size.

misleading. A real galaxy is held together by its own gravity and is not free to expand with the universe. Similarly, if the chalked circle is labeled "Solar System," "Earth," "atom," or almost anything, the result would be misleading because most systems are held together by various forces in some sort of equilibrium and cannot partake in cosmic expansion. If we label the chalked circle "cluster of galaxies" the result could also be wrong because most clusters are bound together and cannot expand. Superclusters are vast sprawling systems of numerous clusters that are weakly bound and can expand almost freely with the universe.

This experiment teaches us a useful lesson. We detect expansion because our measuring instruments do not expand but have fixed sizes. If everything were like the chalked circle, free to expand, then clearly there would be no way of detecting expansion. It is an amusing thought that perhaps the universe is not expanding but is static, and we fail to notice this because all atoms – and this means ourselves, our laboratories, and observatories – are all shrinking. With tongue in cheek, Eddington in 1933, in *The Expanding Universe*, said the theory of the "expanding universe" might also be called the "theory of the shrinking atom." He said: "We walk the stage of life, performers of a drama for the benefit of the cosmic spectator. As the scenes proceed he notices that the actors are growing smaller and the action quicker. When the last act opens the curtain rises on midget actors rushing through their parts at frantic speed. Smaller and smaller. Faster and faster. One last microscopic blur of intense agitation. And then nothing."

Experiment 5: Disks of paper are "galaxies"

We place on the surface of the rubber sheet a disk of paper to represent a galaxy or any other bound system (see Figure 14.10). As the sheet expands the disk stays constant in size. This result is not misleading, and we have found a way in ERSU of correctly representing a galaxy or a cluster of galaxies of fixed size. We sprinkle uniformly over the surface disks of various sizes to represent the galaxies (see Figure 14.11). Strictly speaking, the disks should represent the largest bound systems, the clusters of galaxies, but for convenience we shall continue to refer to them as "galaxies."

Experiment 6: World map and world picture

A selected galaxy is surrounded by receding galaxies and its wormlike inhabitants might therefore think that they occupy the cosmic center from which everything is flying away. But because ERSU, like our universe, is uniform, with no cosmic center and no edge, this impression of being at the center is shared with the inhabitants of all galaxies.

We, the ERSU experimenters, look down on the surface and observe that it is uniform; we see that the sprinkled disks, on average, are everywhere the same, and the surface

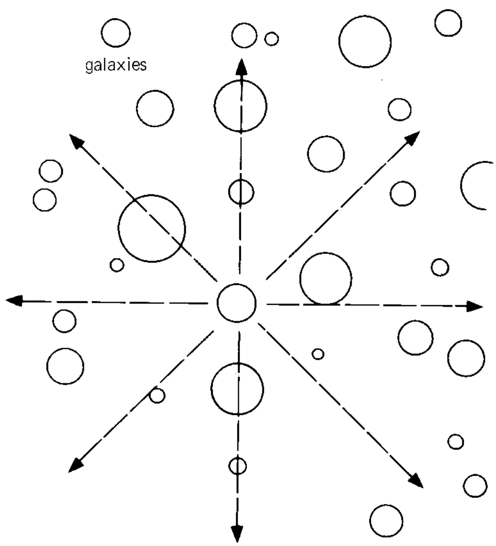

Figure 14.11. A large number of disks are placed on the expanding surface. About any one disk the other disks recede isotropically. This gives a godlike view of space at an instant in time and is what Milne called the world map.

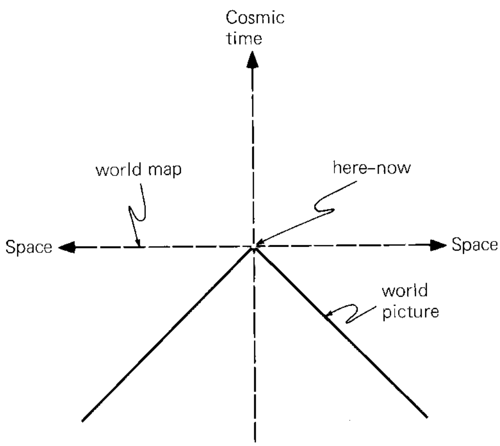

Figure 14.12. A spacetime diagram that shows the backward lightcone on which the observer sees the universe. This wormlike view of the universe is what Milne called the world picture.

everywhere expands in the same way. We are godlike spectators seeing everything everywhere in space at an instant of cosmic time. This is what Edward Milne called the *world map* (see Figure 14.12). It is also what is seen by the cosmic explorer who rushes around and instantaneously sees that all places are alike.

Wormlike denizens in a particular galaxy look out in space and back in time and cannot see the way things are everywhere in space at the moment of observation. They cannot observe the world map. Instead, they observe things distributed on their backward lightcone. They observe what Milne called the *world picture*. The world map is godlike, the world picture is wormlike. Both would be similar if the speed of light were infinitely great.

Regrettably, we in our universe, like the inhabitants in ERSU, are limited to the wormlike view. From the Galaxy we see other galaxies scattered about us isotropically and moving away isotropically. All directions are alike. Only by invoking the location principle can we conclude that probably all places are alike. The location principle bridges the world picture and the world map.

Experiment 7: Velocity–distance law

Our next experiment shows that ERSU obeys the velocity–distance law (see Figure 14.13). We choose any disk and label it A. A second disk, labeled B, at a certain distance, moves away from A at a certain velocity. A third disk, labeled C, in the same direction as B and at twice the distance, moves away from A at twice the velocity. It must. The expansion is homogeneous, and therefore C moves away from B at the same velocity that B moves away from A. This argument, extended to disks E, F, G, . . . , all equally spaced in the chosen direction at a moment in time, shows that the recession velocity relative to A is always strictly proportional to distance. The equal spaces between disks all increase in the same way and the disks remain equally spaced. This result is independent of the location of disk A. ERSU thus shows us that homogeneity is preserved when the expansion is homogeneous:

recession velocity = constant × distance.

The linear expansion law of Equation [14.6]

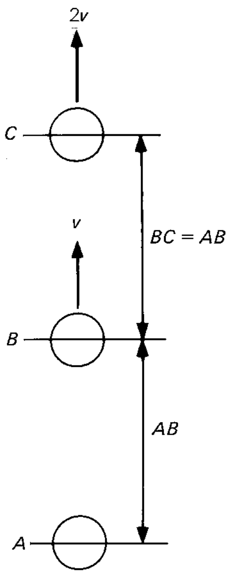

Figure 14.13. Three disks A, B, and C are arranged in a straight line with the distance between A and B equal to the distance between B and C. In uniform expansion the velocity that B recedes from A is the same as the velocity that C recedes from B. Hence C recedes from A at twice the velocity that B recedes from A, showing that the recession velocity is proportional to distance.

is a direct consequence of time-invariant homogeneity. The recession velocity relative to any disk at rest on the surface obeys the same law. If the expansion is anisotropic (faster in one direction than in another), the "constant" has different values in different directions. But in isotropic expansion, which is our primary interest, the constant has the same value for recession in all directions. The law we have derived by this experiment is the velocity–distance law:

$$V = HL, \tag{14.9}$$

and at the time of observation H is denoted H_0.

We must pause and take note that recession velocity V, Hubble term H_0, and distance L all require thought and careful interpretation, for each is open to misunderstanding. The velocity–distance law is obviously true in the world map for us looking down on ERSU, but not so for the poor observers inhabiting their disks who see only the world picture.

The expression recession velocity needs careful handling. On the surface of ERSU the disks (our imaginary galaxies) are stationary; they move apart because the surface is expanding. The disks do not move on the surface, but are at rest and are carried apart by the expansion of the surface. Similarly, the galaxies in the universe are stationary, yet recede from one another because intergalactic space expands. The galaxies are not hurling through space; they are at rest in space and are carried apart by the expansion of space. Recession velocity is therefore not an ordinary velocity in the usual sense and is unlike the velocities encountered on Earth, in the Solar System, or in the Galaxy. They are not Newtonian velocities or the velocities used in special relativity. For this reason we must be careful in cosmology when using the word "velocity," and to avoid confusion we shall most of the time use recession velocity or just recession to indicate relative motion owing to the expansion of space.

The Hubble term H (present value denoted by H_0) is the same everywhere in space at a common instant in cosmic time, but is usually not constant in time. The expansion may have been faster in the past, in which case the value of H was greater than H_0; and if the expansion was slower, the value of H was smaller than H_0. Observers look out from their galaxy to great distances and look back great periods of time, and see H having different values at different distances. To them the velocity–distance law is true only for short distances; at larger distances the law breaks down because H appears to change with distance. The velocity–distance law is true in the world map visualized by the theorist but not in the world picture seen by the observer.

The measurement of distances is no great problem to the theorist who can always use a tape measure (see Chapter 10). But the observer who looks out in space and back in time uses rough-hewn scales of distance and has a dreadful problem trying to

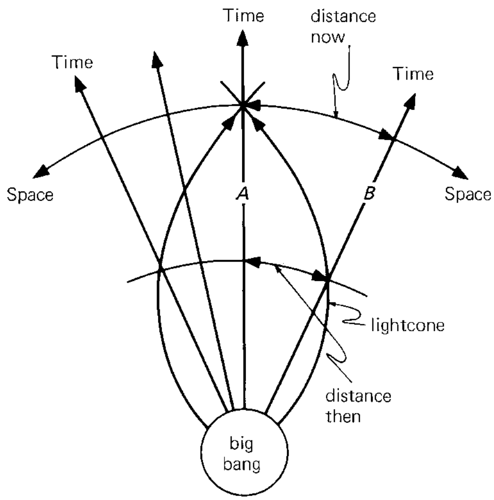

Figure 14.14. Galaxies shown as world lines diverging from the big bang. An observer in galaxy A looks out into space and back in time and sees galaxy B as it was in the past. There are thus two distinctly different distances between galaxies A and B: the "distance now" of B from A, and the "distance then" at the time B emitted the light that A now sees. The distance now is the tape-measure distance in the world map that is used in the velocity–distance law.

determine how far away are the faint galaxies. The difficulty is twofold. First, everything seen is distributed on the sky and distances are not apparent but must be inferred. Second, distances in an expanding universe change with time (see Figure 14.14). We may speak of the distance of a galaxy now at the time of observation or at the time when the galaxy emitted the light now seen. In an expanding universe the distance a galaxy now has is greater than the distance it had when it emitted the light now seen. The distance now is the theorists' tape-measure distance in the world map, the distance at the time of emission is the observers' distance in the world picture. In the velocity–distance law we must use the theorists' distances that galaxies have now at a common instant in time in order to determine H_0. All world-picture distances must be adjusted before they can be used in the velocity–distance law of the world map. The determination of the Hubble term H_0 requires the mapping of the world picture into the world map.

Experiment 8: The Hubble sphere

The recession velocity increases with distance and equals at a certain distance the velocity of light (see Figure 14.15). This distance is c/H_0 and is the Hubble length L_H:

$$L_H = \frac{c}{H_0}. \qquad [14.10]$$

This expression is obtained by writing $V = c$ and $L = L_H$ in Equation [14.6]. With a Hubble term of H_0 equal to $30h$ kilometers a second per million light years, we find

$$L = 10h^{-1} \text{ billion light years.} \qquad [14.11]$$

If $h = 0.5$, the recession equals the velocity of light at distance 20 billion light years. Notice that the velocity–distance law can be written in the form:

$$\frac{V}{c} = \frac{L}{L_H}, \qquad [14.12]$$

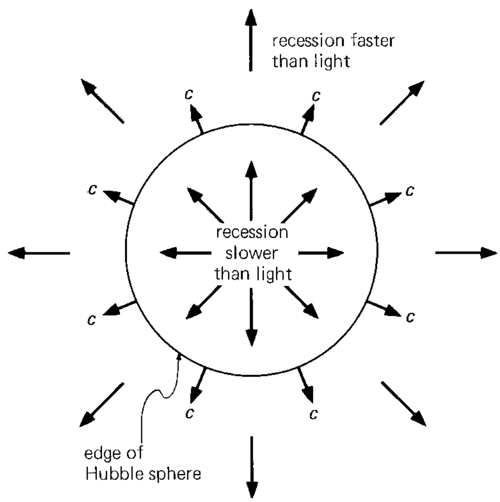

Figure 14.15. At the edge of the Hubble sphere the recession velocity of the galaxies is transluminal (equal to the velocity of light); inside the Hubble sphere all galaxies recede subluminally (slower than the velocity of light); and outside all galaxies recede superluminally (faster than the velocity of light). The observable universe is approximately the size of the Hubble sphere. A more exact definition is given in Chapter 21.

where V is the recession velocity of a galaxy at distance L in the world map. A Hubble length $L_H = c/H_0$ serves as a cosmic yardstick, and when we speak of distances of cosmic magnitude we have in mind distances comparable with the Hubble length.

In ERSU we draw about any disk a large circle whose radius we call the Hubble length. Inside this Hubble circle, which we shall call a sphere, the recession is subluminal (less than the velocity of light), outside the recession is superluminal (greater than the velocity of light), and at the edge of the Hubble sphere the recession is transluminal (equal to the velocity of light). Each disk (galaxy) is the center of its Hubble sphere.

The universe, like ERSU, has no edge and cannot terminate abruptly at the boundary of the Hubble sphere. A cosmic edge at which the recession from our Galaxy equals the velocity of light, even if it existed (as often implied in popular literature), could not be at the same distance from all other galaxies. There is no cosmic edge and galaxies farther away than the Hubble distance recede faster than the velocity of light. How, the beginning cosmologist asks, can galaxies move faster that light? The answer is that galaxies are not moving through space but are moving apart by the expansion of intergalactic space. No galaxy can move through space faster than light and in its local space it obeys always the rules of special relativity. But recession is a result of the expansion of space that obeys the rules of general relativity, and is not like motion through space that obeys the rules of special relativity. Recession velocity is without limit, and in an infinite universe a galaxy at infinite distance has infinite recession. Those persons who find it difficult to understand that recession is without limit usually make the mistake of thinking that the receding galaxies are like projectiles shooting away through space. This is an incorrect view. The correct view is of galaxies more or less at rest in expanding space.

This important experiment demonstrates that the expansion of space does not obey the rules of special relativity, and the recession velocity is not limited by the speed of light.

Experiment 9: The steady-state expanding universe

The steady-state expanding universe is easily simulated with ERSU. As the surface expands and the disks move apart, we continually sprinkle new disks on the surface so that the average separating distance between disks remains always the same. The surface presents an unchanging appearance because of the "continuous creation" of disks. Expansion also never changes and the Hubble term H_0 therefore stays constant and the Hubble sphere has constant radius.

Experiment 10: Comoving galaxies and peculiar motion

Galaxies stationary in expanding space are said to be comoving. They comove with the expansion. Clocks on comoving galaxies all measure cosmic time. Our gadabout cosmic explorer has set all clocks on comoving bodies to read the same time. In subsequent tours of the homogeneous universe the explorer finds that these clocks are in constant agreement. All comoving bodies have their world lines perpendicular to a cosmic space of uniform curvature and uniform expansion (Figure 14.14).

In all previous experiments with ERSU we have supposed that the disks comove with the expanding sheet. But few galaxies are exactly comoving. They dither around relative to their neighbors and have independent motion in their local space. This independent motion superposed on the expansion is known as peculiar motion. Because of peculiar motion the world lines of galaxies are not straight but slightly crinkled. We can easily imagine that all disks in ERSU have independent motion and jitter around slowly on the expanding surface.

Observers see the peculiar velocities of other galaxies superposed on the flow of recession. Normally the observers in one galaxy cannot distinguish between the recession and peculiar motion of other nearby galaxies. Peculiar motion tends to be random and the velocities of many galaxies in a particular region can be averaged to find the recession velocity of that region. Observers must take into account the peculiar motion of their own galaxy, and this can in principle be done by averaging the peculiar motion of neighboring galaxies, or better still, by determining the anisotropy of the cosmic background radiation.

It sounds easy, particularly for us looking down on ERSU, but in the real universe the determination of peculiar velocities is very difficult. Fortunately, recession dominates at large distances and peculiar motions can then often be neglected. Some nearby galaxies are approaching us and others are moving away: farther away, a few are approaching and most are receding; and even farther away, none are approaching and all are receding. Typical peculiar velocities of galaxies in small clusters are 300 kilometers a second. Beyond $10h^{-1}$ million light years, at redshifts greater than 0.001, recession dominates. In rich clusters many galaxies have peculiar velocities as great as 3000 kilometers a second, and in their case the recession dominates beyond $100h^{-1}$ million light years, at redshifts greater than 0.01.

Our own peculiar motion in the universe consists of the Earth moving about the Sun (30 kilometers a second with annual variation in direction), the Sun moving around in the Galaxy (200 kilometers a second), the motion of the Galaxy in the Local Group (approximately 100 kilometers a second), the peculiar motion of the Local Group (350 kilometers a second) toward the Virgo cluster, and the motion of the Local Supercluster (300 kilometers a second) toward the Hydra Centaurus supercluster, giving a net peculiar velocity of 600 kilometers a second in the direction of the constellation Leo. Determining the peculiar velocity of the Local Group in the universe by optical means is not easy, and the most reliable information comes from the anisotropy of the cosmic background radiation.

The implications of the experiments performed so far are startling. First, in Newtonian theory we are taught there is no such thing as absolute rest and all motions are purely relative. Yet in cosmology a comoving body is in a state of absolute rest that can in principle be verified. All peculiar velocities have absolute values that can be determined relative to the state of rest in the local comoving frame. Second, in special relativity we are taught that there is no preferred way of decomposing spacetime into space and time and that all decompositions are relative. Yet in cosmology spacetime separates naturally into uniformly curved expanding space and orthogonal cosmic time. Third, we are taught that a body cannot move faster than light, but in cosmology we find that comoving bodies obey a velocity–distance law in which recession velocities can exceed the velocity of light and are without limit.

What we have been taught in Newtonian mechanics and special relativity theory applies to local peculiar motions in the laboratory, the Solar System, and the Galaxy, and not at all to cosmic motion. All local velocities are peculiar within the cosmic frame of reference and cannot exceed the velocity of light. Recession velocity, however, is not a local phenomenon; it is the result of the expansion of space and does not conform to the rules of special relativity. In summary, we may say that motion in an expanding universe is compounded from recession and peculiar velocities; recession velocities are due to the expansion of space and are without limit, and peculiar velocities are due to motion through space and conform to special relativity.

Experiment 11: Sub-Hubble sphere

The sub-Hubble sphere contains the nearby universe in which astronomical peculiar motions dominate over cosmic recession. Thus

$$L_{\text{sub-}H} = \frac{V_{\text{pec}}}{H_0} = \frac{V_{\text{pec}}}{c} L_H, \qquad [14.13]$$

where $L_{\text{sub-}H}$ is the sub-Hubble radius and V_{pec} is a typical large peculiar velocity. A typically large peculiar velocity, let us say, is 1000 kilometers a second, and the present radius $L_{\text{sub-}H}$ of the sub-Hubble sphere is hence $10h^{-1}$ megaparsecs, or approximately $30h^{-1}$ million light years. Meaningful distant cosmological observations are made beyond the sub-Hubble sphere where the Hubble flow becomes fully developed and dominates over peculiar motions. Galactic redshifts less than 0.003 are sub-Hubble and not cosmologically significant. If we take into account peculiar velocities as large as 3000 kilometers a second, the sub-Hubble sphere approaches $100h^{-1}$ million light years and sub-Hubble redshifts are as large as 0.01. This raises troubling questions: Does the linear Hubble law of Equation [14.5] break down before leaving the sub-Hubble sphere? Is it possible that the linear Hubble law applies nowhere, neither in the sub-Hubble sphere (where measurements are of no cosmological significance) nor in the Hubble sphere itself? Must we interpret observations with a more exact redshift–distance formula appropriate to a specific model of the universe? We glimpse here one of the reasons why the determination of H_0 is so difficult.

The sub-Hubble sphere is dominated by the irregularities of astronomical systems. Beyond the sub-Hubble sphere, astronomical irregularities are much less pronounced and tend to be less important. The length $L_{\text{sub-}H}$ is a measure of the scale of the largest irregularities. In response to the question, over how large a region should we average matter to discover the cosmic density, the answer is the size of the sub-Hubble sphere.

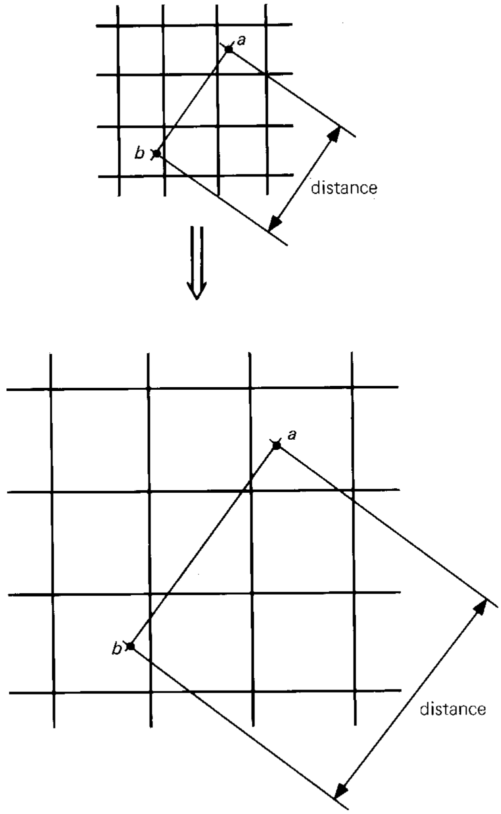

Figure 14.16. This shows a comoving coordinate system consisting of a network of intersecting lines drawn on the expanding surface. All coordinate distances, such as that between points *a* and *b*, remain constant in value. The real distance, of course, increases.

Experiment 12: Comoving coordinates

With chalk we draw on the expanding surface of ERSU a network of lines. These intersecting lines form what is called a comoving coordinate system (see Figure 14.16). The coordinate lines are fixed on the surface. Whether the lines are straight or curved is not very important; what matters is that comoving disks are fixed in their relation to these coordinate lines, and their positions are specified by appropriate coordinate values. All distances between comoving disks remain constant when measured in comoving coordinates, and these constant distances are called comoving coordinate distances. A comoving coordinate system enables us to distinguish between recession and peculiar motion. Recession velocities apply to bodies stationary in the comoving coordinate system, and peculiar velocities apply to bodies moving relative to the comoving coordinates.

Experiment 13: An empty universe

We remove all disks from the surface of ERSU except the one occupied by our observers. These observers are left with no way of determining their peculiar motion; moreover, they cannot tell if the surface is expanding, static, or contracting. Confronted with this situation of undressed space of indeterminate kinematic behavior, they might ruefully recall Mach's words: "When, accordingly, we say that a body preserves unchanged its direction and velocity in space, our assertion is nothing more or less than an abbreviated reference to the entire universe." By the entire universe, Mach had in mind the "remote heavenly bodies" that we have, figuratively, removed from ERSU. Even the ubiquitous explorer is puzzled and tempted to believe that empty space is meaningless and the physical properties of space are dependent on the presence of matter. Then we learn the trick of scattering around a few tiny comoving particles, or (what amounts to the same thing) of chalking on the sheet a network of comoving coordinates. This illustrates what happened originally with the empty universe proposed by Willem de Sitter: nobody knew that it was expanding until it was sprinkled with test particles and equipped with a network of comoving coordinates. It was then found to have kinematic properties even though it contained no matter.

Experiment 14: Idealized universe

The distribution of disks in ERSU is clumpy on small scales and not obviously uniform except on large scales. These random irregularities are distracting when we wish only to study the large-scale behavior of the universe. In the last experiment we removed all the disks. Let us now take these disks, grind them into powder, and then smoothly and uniformly distribute the powder over the expanding surface of ERSU. This represents an idealized universe in which all galaxies are smoothed out into a continuous fluid of uniform density.

Idealized universes have their uses. It is of interest, for instance, to know what happens to light when it propagates in a universe free of all irregularity. The large-scale effects of the universe are first determined and corrections for irregularities can be added later. Cosmologists take the view that an idealized universe is a convenient fiction, useful for easy calculations, and in Howard Robertson's words a sort of cosmic undergarment onto which the ostentatious detail of the real world is tacked.

An idealized universe is also useful for studying the origin of galaxies. These vast celestial systems have not always existed, and prior to their formation the universe was much less irregular than at present. The idealized universe may therefore resemble the way things were once upon a time. How the original unstructured universe evolved into its present highly structured state is a major research area in cosmology.

Idealized universes – perfectly homogeneous and isotropic – are known as Robertson–Walker models after Howard Robertson and Arthur Walker showed rigorously that universes obeying the cosmological principle have a spacetime that uniquely separates into a curved expanding space and a cosmic time that is common to all comoving observers.

By studying the expansion of a small region of an idealized universe we automatically learn how all other small regions expand, and by piecing these regions together we learn how the whole universe expands. The behavior of large regions, even the universe itself, is mirrored in the behavior of small regions. We shall exploit this intriguing aspect of cosmology in Chapter 17.

MEASURING THE EXPANSION OF THE UNIVERSE

The universal scaling factor

Distances between galaxies (or clusters of galaxies) increase in an expanding universe, whereas distances inside galaxies and even clusters of galaxies do not increase. This difference, of vital importance in astronomy and other sciences, is a distraction in cosmology. The solution is simple. Because we are interested only in cosmic phenomena, we abolish all astronomical systems and use instead an idealized universe. In this way we become free to consider how all distances vary in time without the bother of distinguishing between large expanding and small nonexpanding regions. In a smoothed universe, distances can be as small or as large as we please.

We consider only uniform (homogeneous and isotropic) expansion. Over an interval of time all distances between comoving points increase by the same factor. If one distance increases by 1 percent, then all distances increase by 1 percent. A comoving triangle, for example, has its three sides scaled by the same factor and the dilated triangle retains its original shape. There exists a universal scaling factor, often (for historical reasons) denoted by R, that increases in time in a uniformly expanding universe, and at any instant in cosmic time has the same value everywhere in space (see Figure 14.17). All distances of the tape-measure kind between comoving points increase in proportion to R, all areas of two-dimensional figures increase in proportion to R^2, and all volumes of three-dimensional figures increase in proportion to R^3.

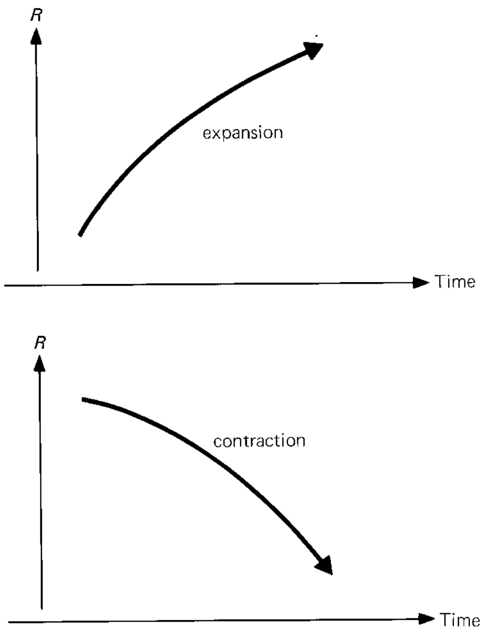

Figure 14.17. The scaling factor R – everywhere the same in space – changes in time; its value increases in an expanding universe and decreases in a contracting universe.

Comoving coordinates and the scaling factor

The scaling factor has many important uses. We can best show this by beginning with comoving coordinates. A network of intersecting lines drawn on the surface of ERSU is an example of a comoving coordinate system. All comoving points are separated by coordinate distances that stay constant during expansion. Such a coordinate system is fixed in expanding space. With this coordinate system we can say that the actual distance L is the coordinate distance multiplied by the scaling factor:

$$L = R \times \text{coordinate distance.} \qquad [14.14]$$

The "actual distance" is the tape-measure distance: the distance that would be measured by stretching a tape measure in a uniformly curved surface; it is the straight-line (shortest or geodesic) distance between two points. The coordinate distance stays constant, whereas the distance itself increases in proportion to R. Observers use other distances, such as luminosity distance and distance by apparent size (see Chapter 19), but these distances unfortunately offer no help in understanding the fundamentals of cosmology.

Instead of a flat surface with a network of lines drawn on it, as in ERSU, we could use a rubber balloon with latitude and longitude coordinates drawn on its surface. Latitude and longitude are comoving coordinates, and a point on the surface of an expanding balloon retains its position relative to these coordinates. In this application, the scaling factor becomes the radius of the balloon's surface. The distance between two points on the surface is the constant coordinate distance (expressed in terms of latitude and longitude) multiplied by the radius R. Originally the scaling factor was referred to as the radius of the universe, and this is why it is denoted by the symbol R. The expression "radius of the universe" may be misleading, however, because some universes are flat, and a more neutral expression, such as "scaling factor," is less ambiguous.

Distances, areas, volumes, and densities

Usually we are not interested in the actual value of the scaling factor R, only in how it changes, and how its value compares at different stages in cosmic time.

Let R_0 be the value of the scaling factor at the present time, and let L_0 be the present distance between two comoving points:

$$L_0 = R_0 \times \text{coordinate distance.}$$

At any other time the scaling factor is R, and the distance between the same two comoving points is

$$L = R \times \text{coordinate distance,}$$

and only the scaling factor has changed. Therefore

$$\frac{L}{L_0} = \frac{R}{R_0}, \qquad [14.15]$$

which is obvious, for if the universe doubles its size, then the scaling factor is increased twofold and all distances between comoving points are also increased twofold. Pursuing similar arguments, we can say that if an area comoves, then its value A in terms of its present value A_0 is

$$\frac{A}{A_0} = \left(\frac{R}{R_0}\right)^2, \qquad [14.16]$$

Similarly with volumes:

$$\frac{V}{V_0} = \left(\frac{R}{R_0}\right)^3, \qquad [14.17]$$

where V (not to be confused with recession velocity) is a comoving volume whose present value is V_0. When distances double in size, comoving areas increase fourfold, and comoving volumes increase eightfold.

Suppose an expanding volume V contains N particles and that no particles are created or destroyed. The density of particles, call it n, is the number in a unit of volume, such as a cubic centimeter. Hence $n = N/V$. The present density n_0 is the fixed number N divided by the present volume: $n_0 = N/V_0$. We know how volumes vary, and hence densities vary as:

$$n = n_0 \left(\frac{R_0}{R}\right)^3. \qquad [14.18]$$

With this important result it is possible to find the density in the past or the future from the present density. The average density of matter in the universe is about 1 hydrogen atom per cubic meter. Back in the past when the scaling factor was 1 percent of its present value, the density was a million times greater and equal to 1 hydrogen atom per cubic centimeter. This is a typical value for the density of galaxies, and we infer that galaxies as we know them had not formed when the universe was smaller than 1 percent of its present size. Incidentally, when we say that the universe "changes in size" we imply not that it is finite but only that the scaling factor R changes.

THE VELOCITY–DISTANCE LAW

The scaling factor increases with cosmic time in an expanding universe. But how fast does it increase? Thought on this matter soon makes clear that the rate of increase of R must have something to do with the Hubble term.

Consider a comoving body at a fixed coordinate distance and at an actual distance $L = R \times$ coordinate distance. As R increases, the distance L increases and the body recedes. The faster R increases, the faster the body recedes. The recession velocity V of a comoving body is the rate at which its distance L increases. This equals the rate of increase of R multiplied by the constant coordinate distance; that is

$$V = \text{rate of increase of } R \times \text{coordinate distance.} \qquad [14.19]$$

For convenience we use Newton's notation and let $\dot{R}$ stand for the rate of increase of

R; therefore

$$V = \dot{R} \times \text{coordinate distance.} \qquad [14.20]$$

On inserting the expression $L = R \times$ coordinate distance into Equation [14.20], we obtain

$$V = L\frac{\dot{R}}{R}. \qquad [14.21]$$

This is the velocity–distance law in which the Hubble term is given by the expression

$$H = \frac{\dot{R}}{R}. \qquad [14.22]$$

The Hubble term is everywhere the same in space and in most universes varies in time. This important derivation of the velocity–distance law:

$$V = HL, \qquad [14.23]$$

reveals that this fundamental law is nothing more than the consequence of uniform expansion. A scaling factor that is everywhere the same in space, and varies in time, automatically yields the linear velocity–distance law.

Because H changes in time, the velocity–distance law at the present cosmic epoch is

$$V = H_0 L,$$

as in Equation [14.6]. As stressed previously, this is the theorists' law and is not the observers' Hubble law:

$$zc = H_0 L,$$

of Equation [14.5].

Hubble time and the age of the universe

The Hubble time (or period) would be the age of the universe if expansion were constant (see Figure 14.18). The Hubble time is

$$t_H = \frac{1}{H_0} = \frac{L_H}{c}$$
$$= 10h^{-1} \text{ billion years,} \qquad [14.24]$$

and is the Hubble distance divided by the speed of light. Often the Hubble time is referred to as the expansion time. If h has a value somewhere between 0.5 and 1, t_H lies between 10 and 20 billion years.

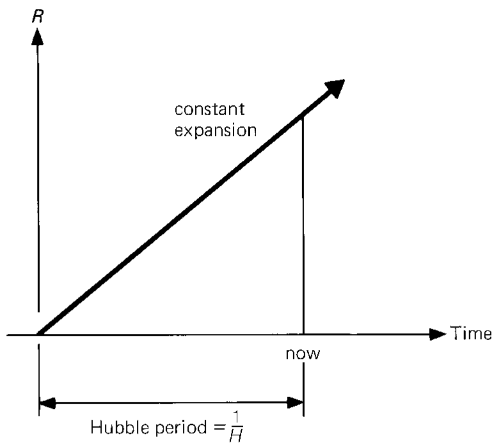

Figure 14.18. A Hubble period would be the age of the universe if the universe expanded at a constant rate.

In almost all universes studied by cosmologists the scaling factor does not increase at a constant rate in time but either accelerates or decelerates. In an accelerating universe R increases more rapidly in time and the actual age is greater than a Hubble time. In a decelerating universe R increases more slowly in time and the actual age is shorter than a Hubble time. At present we cannot tell precisely the age of the universe, and a Hubble time serves as a rough measure of age. We must be on our guard, however, because in some universes a Hubble time is a grossly misleading indicator of age. The de Sitter and steady-state universes of infinite age are examples. In the steady-state universe nothing ever changes in a cosmic sense, and therefore the Hubble term stays constant (see Figure 14.19). Because $\dot{R} = HR$, and H is constant, $\dot{R}$ increases as R, and this is an accelerating universe.

The Hubble sphere and the observable universe

Broadly speaking, the observable universe spans the Hubble sphere. If the age of the universe is roughly a Hubble period, the

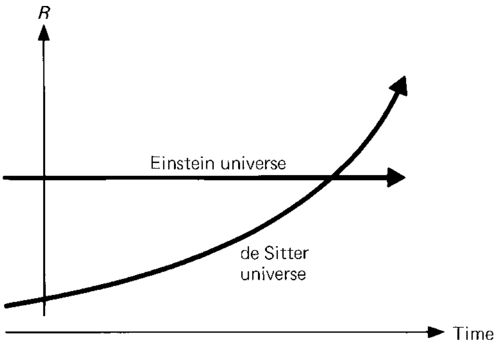

Figure 14.19. The static Einstein universe and the exponentially expanding de Sitter universe. In the Einstein universe the Hubble term is zero and the scaling factor is constant in time. In the de Sitter universe, as in the steady-state universe, the Hubble term is constant; and the scaling factor R increases exponentially with time.

distance light travels in this time is approximately a Hubble length. We cannot see things outside the observable universe because their light is still traveling and has yet to reach us. The older the universe, the more we can see of it (Chapter 22).

ACCELERATING AND DECELERATING UNIVERSES

We have seen that a tape-measure distance increases according to the rule

$$L = R \times \text{coordinate distance},$$

and the recession velocity of a comoving body of constant coordinate distance is the rate of increase of distance ($V = \mathrm{d}L/\mathrm{d}t = \dot{L}$) according to

$$V = \dot{R} \times \text{coordinate distance}.$$

Acceleration is just the rate of increase of velocity ($\mathrm{d}V/\mathrm{d}t$), and if we use the symbol $\ddot{R}$ to denote the rate of increase of $\dot{R}$, we have

$$\text{acceleration} = \frac{\mathrm{d}V}{\mathrm{d}t} = \ddot{R} \times \text{coordinate distance} \qquad [14.25]$$

because, as before, the coordinate distance of the comoving body is constant. We now use our first relation, $L = R \times$ coordinate distance, and find

$$\text{acceleration} = L\frac{\ddot{R}}{R}. \qquad [14.26]$$

The term $\ddot{R}/R$ was once referred to as the acceleration. More popular is the deceleration term, indicated by the symbol q and defined by

$$q = -\frac{\ddot{R}}{RH^2}. \qquad [14.27]$$

The deceleration term, like the Hubble term, is constant everywhere in space at a common instant in time, and generally changes in time.

When the rate of expansion never changes, and $\dot{R}$ is constant, the scaling factor is proportional to time t ($R = \text{constant} \times t$), and the deceleration term is zero, as in Figure 14.18. When the Hubble term is constant, the deceleration term q is also constant and equal to -1, as in the de Sitter and steady-state universes, shown in Figure 14.19. In most universes the deceleration term changes in time, as illustrated in Figure 14.20.

When the deceleration term is positive, there is deceleration (slowing down of expansion), and when it is negative, there is acceleration (speeding up of expansion). From the curves in Figure 14.21 we see that in a decelerating universe, where q is positive, the age of the universe is shorter than a Hubble period; and in an accelerating universe, where q is negative, the age of the universe is longer than a Hubble period.

The present values of the Hubble term H_0 and deceleration term q_0 show how the scaling factor is now changing. The precise value of the Hubble term is unknown and it is believed that the h parameter lies somewhere between 0.5 and 1. The rate of increase in the scaling factor is possibly slowing down and the expansion is decelerating. From observations it is very difficult to determine precisely the value of the deceleration term. The available evidence at present suggests that q_0 is perhaps as small as 0.05, but many cosmologists believe that it may be as large as 0.5 because of "missing

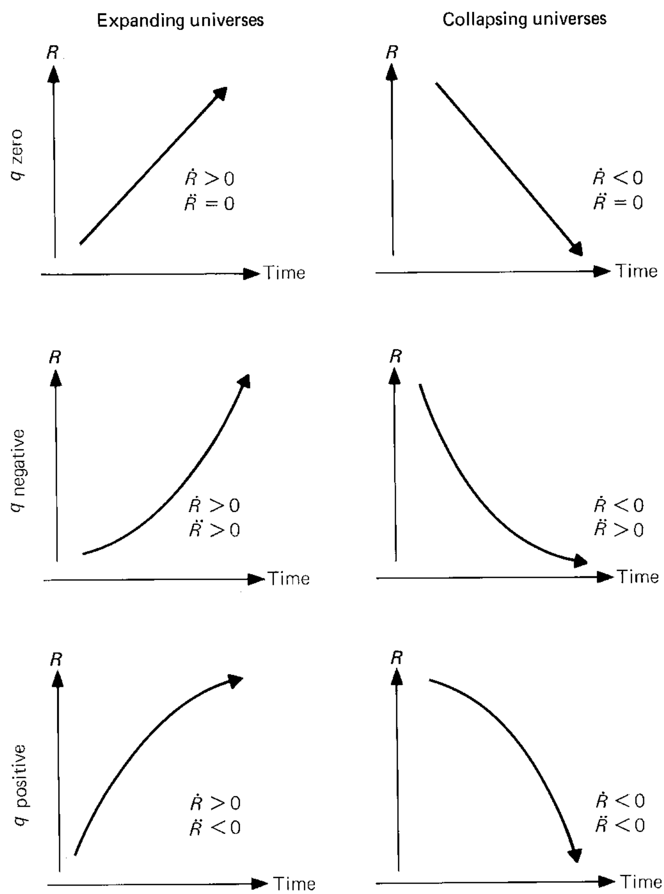

Figure 14.20. This array of diagrams shows universes classified by their values of the Hubble term H and the deceleration term q.

mass." The estimated values of H_0 and q_0 have changed frequently over the decades, and even current estimates must be viewed with reservations. Fortunately, much in cosmology can be discussed in general kinematic and geometric terms without having an exact knowledge of the present values of the Hubble and deceleration terms.

CLASSIFYING UNIVERSES

Geometrical classification

Universes are classified in various ways. Here we mention three quite simple methods of classifying uniform universes. The first is the geometrical method based on curvature. In this method there are three classes:

(a) $k = 0$: flat space (open)
(b) $k = 1$: spherical space (closed)
(c) $k = -1$: hyperbolic space (open)

defined by the curvature constant k and discussed in Chapter 10.

Kinematic classification

The way in which the scaling factor varies, based on the values of H and q, gives us a second method of classification, as shown in Figure 14.20. All models can be characterized by whether they expand or contract, and accelerate or decelerate. To the four classes

(a) $(H > 0, q > 0)$: expanding and decelerating
(b) $(H > 0, q < 0)$: expanding and accelerating

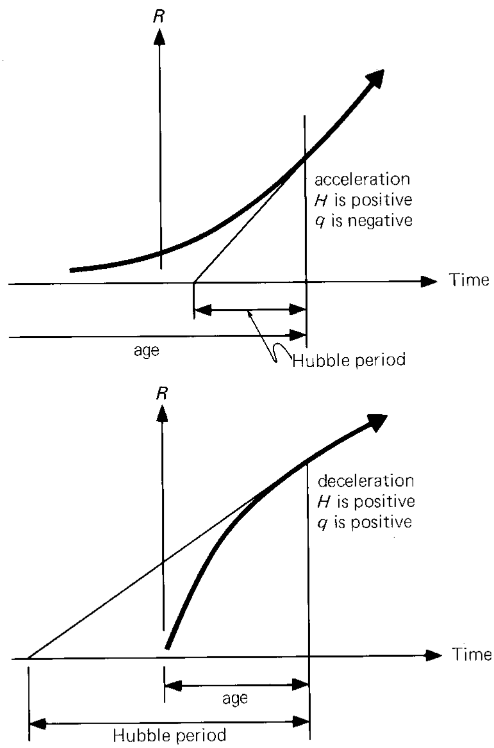

Figure 14.21. Expansion in accelerating and decelerating universes. Notice that in accelerating universes the age is greater than the Hubble period, and in decelerating universes the age is less than the Hubble period.

(c) $(H < 0, q > 0)$: contracting and decelerating
(d) $(H < 0, q < 0)$: contracting and accelerating

we can add three classes

(e) $(H > 0, q = 0)$: expanding, zero deceleration
(f) $(H < 0, q = 0)$: contracting, zero deceleration
(g) $(H = 0, q = 0)$: static.

There is little doubt that we live in an expanding universe, and hence only (a), (b), and (e) are possible candidates.

Bang–whimper classification

Another simple method is the bang–whimper classification. Our universe was more dense in the past than now, and also much hotter. The cosmic background radiation is generally accepted as evidence of a dense and hot early universe. Universes that start or end at high density, or pass through a high-density phase, are of the big bang type, and the descriptive name "big bang" was coined by Fred Hoyle. The universes that begin or die "not with a bang but a whimper" (in T. S. Eliot's words) we shall call whimper universes.

A big bang occurs whenever the scaling factor R is either zero or extremely small, and a whimper is a long-drawn-out state that occurs when R is large and without limit. Thus a big bang means: at some time R is close to 0 and density is close to ∞; a whimper means R approaches ∞ and the density approaches 0; and the symbol ∞ denotes infinity. Because a universe can begin either as a bang or a whimper, and can end either as a bang or a whimper, there are altogether four classes (see Figure 14.22):

(a) bang–bang: has finite lifetime
(b) bang–whimper: has infinite lifetime
(c) whimper–bang: has infinite lifetime
(d) whimper–whimper: has infinite lifetime.

Of these four classes, we note that only (a) has a finite lifetime.

Because we live in an expanding universe we can rule out the whimper–bang class (c)

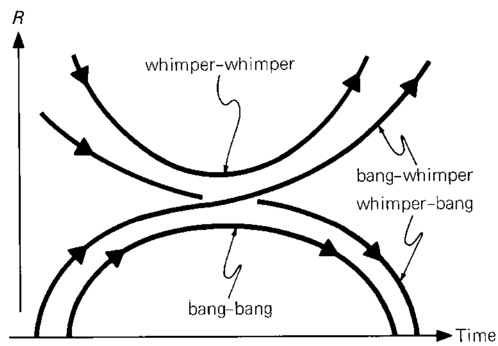

Figure 14.22. The four classes of universes in the bang–whimper classification scheme.

that continually contracts. The whimper–whimper class (d), although possible, is unlikely because the turnaround, or bounce, must presumably occur at a density great enough to create the cosmic background radiation, and a dense turnaround qualifies as a big bang. We are left with only two possible classes of universe: the bang–bang class (a) that expands from a big bang and then collapses back to a big bang, with us at present in the expanding phase; and the bang–whimper class (b) that expands continually for an infinite period of time.

We shall see later in the case of the Friedmann universes that the possible classes are: geometrically, all three curvatures are possible; kinematically, only $(H > 0, q > 0)$ and $(H < 0, q > 0)$ are possible; and physically, only the bang–bang and bang–whimper classes are possible.

REFLECTIONS

1 *In 1886, in his observatory at Tulse Hill just outside London, William Huggins was the first to observe the displacement of stellar spectral lines predicted in 1848 by the French scientist Armand Fizeau. The "classical Doppler" formula (Equation 14.1) was in fact formulated by Fizeau. Previously, in 1843, Christian Doppler had argued that the pitch of sound waves should be affected by the velocity of the source. Doppler argued that this effect should occur not only with sound waves but also with light waves. Thus he correctly predicted the "Doppler effect" for both sound and light. But Doppler erred when he proposed that the color difference between stars in binary systems was due to this Doppler effect. He argued that the approaching star would be blue and the receding star would be red. Although on the right track, he was a long way off in estimating the amount of spectral displacement, as shown by Fizeau. The French scientist Armand Fizeau, who was the first to measure the speed of light in a terrestrial environment, formulated the expression (Equation 14.1) often attributed to Doppler. Through the 19th century the velocity displacement of spectral lines was the Doppler effect and the actual amount of displacement was given by the Fizeau–Doppler formula. I shall continue this forgotten custom and in the case of light use the term "Fizeau–Doppler formula" for nonrelativistic velocities and "relativistic Doppler formula" for relativistic velocities.*

2 *The following article appeared in the* New York Times, *page 61, on January 19, 1921:*

> *"DREYER NEBULA NO. 584*
> *INCONCEIVABLY DISTANT*
> *Dr. Slipher Says the Celestial Speed Champion Is 'Many Millions of Light Years' Away.*
>
> *By Dr. Vesto Melvin Slipher, Assistant Director of the Lowell Observatory, Flagstaff, Ariz.*

FLAGSTAFF, Ariz., Jan. 17. – The Lowell Observatory some years ago undertook to determine the velocity of the spiral nebulae – a thing that had not been previously attempted or thought possible. The undertaking soon revealed the quite unexpected fact that spiral nebulae are far the most swiftly moving objects known in the heavens. A recent observation has shown that the nebula in the constellation Cetus, number 584 in Dreyer's catalogues, is one of very exceptional interest.

"Like most spiral nebulae, this one is extremely faint, and to observe its velocity requires an exceedingly long photographic exposure with the most powerful instrumental equipment. This photograph was exposed from the end of December to the middle of January in order to give the weak light of the nebula's spectrum time to impress itself upon the plate. It is necessary to disperse the nebular light into a spectrum in order to observe the spectral lines, and to measure the amount that they are shifted out of their normal positions, for it is this displacement of the nebula's lines that discloses and determines the velocity with which the nebula is itself moving. The lines in its spectrum are greatly shifted showing that the nebula is flying away from our region of space with a marvelous velocity of 1100 miles per second. This nebula belongs to the spiral family,

which includes the great majority of the nebulae. They are the most distant of all celestial bodies, and must be enormously large.

"If the above swiftly moving nebula be assumed to have left the region of the sun at the beginning of the earth, it is easily computed, assuming the geologists' recent estimate of the earth's age, that the nebula now must be many millions of light years distant.

"The velocity of this nebula thus suggests a further increase to the estimated size of the spiral nebulae themselves as well as to their distances, and also further swells the dimensions of the known universe."

3 *Gerald Whitrow wrote in "Hubble, Edwin Powell" that Hubble's work "made as great a change in man's conception of the universe as the Copernican revolution 400 years before. For, instead of an overall static picture of the cosmos, it seemed that the universe must be regarded as expanding, the rate of the mutual recession of its parts increasing with their relative distance."*

- *In* The Expanding Universe *(1933), Eddington wrote: "The unanimity with which the galaxies are running away looks almost as though they had a pointed aversion to us. We wonder why we should be shunned as though our system were a plague spot in the universe.... But the theory of the expanding universe is in some respects so preposterous that we naturally hesitate to commit ourselves to it. It contains elements apparently so incredible that I feel almost an indignation that anyone should believe in it – except myself."*

4 *We have previously seen that matter affects the geometry of spacetime. We now see that matter also affects the dynamics of spacetime. Indeed, in some expanding universes containing uniformly distributed matter, space is flat and the matter affects only the dynamics and not the geometry of space. In general, in the presence of a uniform distribution of matter: a curved and static space is possible; a flat and expanding space is possible; but a flat and static space is impossible.*

- *Consider two comoving particles separated by distance L. From general relativity, their relative acceleration in uniform space is given by the equation*

$$\ddot{L} = \frac{\Lambda}{3}L - \frac{4\pi G}{3}\left(\rho + \frac{3P}{c^2}\right)L, \qquad [14.28]$$

where Λ is the cosmological constant, ρ the density, and P the pressure. In the empty de Sitter universe, ρ and P are both zero, and therefore $\ddot{L} = (\Lambda/3)L$, and this acceleration is the "de Sitter effect." In the Einstein static universe, $\ddot{L} = 0$, and if $P = 0$ (as Einstein assumed), then $\Lambda = 4\pi G\rho$.

5 *In 1930, Arthur Eddington introduced the idea of an expanding rubber surface in an article entitled "On the instability of Einstein's spherical world." He wrote, "Observationally, galaxies 'at rest' will appear to be receding from one another since the scale of the whole distribution is increasing. It is as though they were embedded in the surface of a rubber balloon which is being steadily inflated." Let us inflate a spherical balloon, and with a soft pen draw on its surface coordinate circles of latitude and longitude, and a few scattered "galaxies" (see Figure 14.23). Notice, as the balloon inflates and deflates, that the lines of latitude and longitude behave as a comoving system of coordinates, and the mutually receding galaxies are stationary relative to the coordinate lines. The balloon analogy was also mentioned by Eddington in 1933 in* The

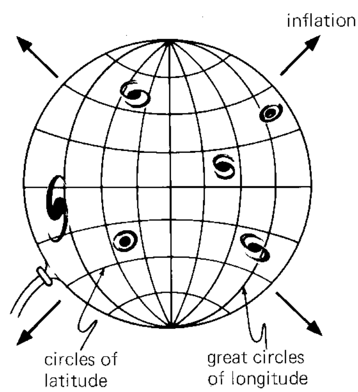

Figure 14.23. These circles of latitude and great circles of longitude on the surface of an expanding balloon illustrate the nature of comoving coordinates.

Expanding Universe*: "For a model of the universe let us represent spherical space by a rubber balloon. Our three dimensions of length, breadth, and thickness ought all to lie in the skin of the balloon; but there is room for only two, so the model will have to sacrifice one of them. That does not matter very seriously. Imagine the galaxies to be embedded in the rubber. Now let the balloon be steadily inflated. That's the expanding universe."*

6 *Hermann Weyl, a mathematician and philosopher, who was a pioneer in general relativity theory and quantum mechanics, wrote in 1922 (at the end of World War I) in the preface of his book* Space, Time, Matter*, "To gaze up from the ruins of the oppressive towards the stars is to recognize the indestructible world of laws, to strengthen faith in reason, to realize the 'harmonia mundi' [harmony of the worlds] that transcends all phenomena, and that never has been, nor will be, disturbed." Weyl in 1923 supposed that the galaxies have diverging world lines, as shown in Figure 14.24, and the galaxies in effect are stationary in a uniform space that is perpendicular to the world lines. In this space the galaxies share a common time (cosmic time). The velocity–distance law, implicit in Weyl's principle, emerged a decade later amidst considerable controversy. The debate began at a British Association science meeting in 1931 and was published as a collection of contributions in* Nature *under the title "The evolution of the universe." From this symposium Edward Milne emerged as a principal contributor, and in a series of publications he formulated the cosmological principle and stressed the fact that the velocity–distance relation is the consequence of time-invariant homogeneity. It seems to have been forgotten by many cosmologists that Hubble's approximate redshift–distance law derives from observation, whereas the exact velocity–distance law derives from theory and the cosmological principle. The two laws are not the same and should not be referred to indiscriminately as Hubble's law. Why? Because one might then think either that the redshift–distance law is rigorously true for all distances (which it is not), or the velocity–distance law is only approximately true (which it is not).*

The cosmological principle formed the foundation of the work by Howard Robertson, Arthur Walker, Richard Tolman and others on homogeneous and isotropic spacetimes (now enshrined in the Robertson–Walker metric) that consist of uniformly curved or flat expanding three-dimensional space with cosmic time as a fourth dimension.

But Milne rejected general relativity and strenuously opposed the expanding space paradigm. He refused to attribute to space (which "by itself has no existence") the properties of curvature and expansion. In protest he developed his own theory, which he called kinematic relativity. Of the expanding space paradigm, he said in 1934, "This concept, though mathematically significant, has by itself no physical content; it is merely the choice of a particular mathematical apparatus for describing and analyzing phenomena. An alternative procedure is to choose a static space, as in ordinary physics, and analyze the expansion phenomena as actual motions in this space" ("A Newtonian expanding universe"). But a bounded finite cloud of galaxies expanding at the boundary at the speed of light in an infinite static space restores the cosmic center and the cosmic

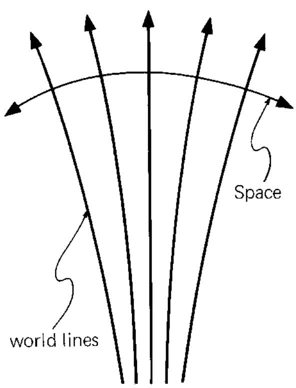

Figure 14.24. The Weyl principle: nebulae have diverging world lines, and are stationary in a space that is perpendicular to the world lines. Hermann Weyl argued in 1923 that the nebulae would recede from one another with apparent velocities that increase with their separation.

edge, and is contrary to modern cosmological beliefs.

7 *The determination of the distances of very distant galaxies is extremely difficult. If we rely only on less distant galaxies, which have redshifts small enough to justify the use of the Fizeau–Doppler formula, the determination of H_0 becomes uncertain, as history has shown. But the greater the redshifts, the greater the problem of determining the distances; furthermore, the nonlinear corrections to the Fizeau–Doppler formula, necessary at larger redshifts, depend on the characteristics of the cosmological model that observations have yet to determine. It seems like a no-win situation; corrections to the observations depend on knowing the model, and the choice of a model depends on knowing the correct observations.*

- *Observers use various kinds of distance indicators, such as "luminosity distance," "apparent-size distance," "number-count distance," and "redshift distance." But the Hubble term H_0 of the linear velocity–distance law is defined in terms of geometric tape-measure distances, and therefore the observers' distance indicators must be translated into tape-measure distances to determine H_0. Translating distance indicators into geometric distances is tricky. In this book, unless otherwise stated, we use only the geometric distances (as in Chapter 19) of the kind one would obtain with a tape measure stretched in the world map. These are the unambiguous, clearly defined, and easily understood distances used in the velocity–distance law, and in terms of these distances the velocity–distance law is linear and recession velocities are without limit.*

8 *Can we prove that all places are alike in space? Can we, in other words, prove that the universe is homogeneous in the world map? Remember, we cannot observe the world map; we see only a world picture that slices back through space and time. From this world picture, which is isotropic but not homogeneous, we try to construct a world map showing what the universe is like everywhere at the present moment in cosmic history. We must project the world picture onto the world map of the present cosmic epoch by transforming the observers' distance indicators into tape-measure distances and by allowing for evolution and expansion. Hence we must assume that the laws of physics are everywhere the same and that things evolve everywhere as they do in our neighborhood. This presupposes the existence of an underlying homogeneity. We prove geometric homogeneity with arguments that presuppose physical homogeneity. At best, we can show that the observed world picture is consistent with a homogeneous world map; we can never prove by direct observation that all places are alike. It is possible for an infinite number of isotropic universes to mimic at some instant our world picture and yet have world maps that are inhomogeneous. Yet each of these deceptive universes requires that we have special location at a cosmic center. On philosophical grounds and by appeal to the location principle we dismiss these possibilities as unlikely. We favor homogeneity because special location is improbable.*

- *It is a curious consequence of homogeneity that on the average everybody in the universe thinks alike. Our gross averaging is done over a cosmologically significant element of volume of hundreds of millions of light years in size and hundreds of millions of years in duration. All variety in thought and outlook is to be found within a cluster of galaxies and not by exploring the uttermost depths of the universe.*
- *An object – an organism or a planet – consists of a finite number of particles. A finite number of particles can be arranged into only a finite number of distinctly different configurations of finite size. If the particles are rearranged an infinite number of ways, then each configuration of finite probability occurs an infinite number of times. Consider now a uniform universe containing space (flat or hyperbolic) of infinite extent. The observable part extends out roughly a Hubble distance of between 10 and 20 billion light years. The unobservable universe that lies beyond extends endlessly. Trillions of trillions of Hubble distances are nothing compared with*

infinity. And if all places are alike, there exists out there an infinite number of identical copies of all things that exist here: an infinite number of Solar Systems having identical Earths, having identical human populations living identical lives. All things of finite probability are repeated an infinite number of times in an infinite universe. The principle of plenitude returns with a vengeance! See "Life in the infinite universe" by G. F. R. Ellis and G. B. Brundrit (1974).

*This form of infinite plenitude was discussed by the German philosopher and poet Friedrich Nietzsche in "the great game of chance that constitutes the universe." He wrote: "In infinity, every possible combination must have been realized, and must also have been realized an infinite number of times" (*The Will to Strength, *1886).*

While on this theme we should consider also the steady-state universe in which everything goes on forever in the same way. It is a universe of infinite and uniform space in which everything is eternally the same. Out there in space at this instant are an infinite number of identical Harrisons writing this identical book. Moreover, in every cosmic element of proper volume every configuration of finite probability has been repeated in the past and will be repeated in the future an infinite number of times. Uniqueness is a forbidden word. At this moment in time an infinite number of Harrisons exists in space and at this place in space an infinite number of Harrisons have existed in the past and will exist in the future. An infinite space of homogeneous content has never appealed to me, and I have felt repelled by the steady-state theory of the universe from its inception because of its eternal sameness in time. What's the point of infinite plenitude when once is usually more than enough? It's easy to understand why some cosmologists, if only for philosophical considerations, favor homogeneous universes that are finite in space and time.

9 *The term big bang, used to denote a dense beginning, was first used by Fred Hoyle in 1949 in his series of BBC radio talks on astronomy, first published in* The Listener *and later in* The Nature of the Universe, *1950. The word whimper, used to denote a universe that does not begin or end with a big bang, was used by T. S. Eliot in "The hollow men":*

This is the way the world ends
Not with a bang but a whimper.

George Ellis has used whimper differently to indicate universes that collapse nonuniformly, with different regions inside their own trapped surfaces (Chapter 20), and do not terminate in a single big bang. Our alternative use of the word is not likely to confuse the reader, and is perhaps more in accord with Eliot's poem. Instead of bang and whimper we could use fire and ice from Robert Frost's "Fire and ice":

Some say the world will end in fire,
Some say in ice.
From what I've tasted of desire
I hold with those who favor fire.
But if it had to perish twice,
I think I know enough of hate
To say that for destruction ice
Is also great
And would suffice.

- *In classical theory, a cosmic singularity occurs when density is infinitely great. A bang type of universe has a singularity when the scaling factor R is zero. What happens at infinite density is not known, and for physical reasons (see Chapters 13 and 20) it is likely that a singular state of this nature is unattainable. The extreme density attained in gravitational collapse, however, is still referred to as a singularity.*
- *Many persons have disliked the notion of a big bang. In* The Expanding Universe *(1933), Arthur Eddington wrote, "Since I cannot avoid introducing this question of a beginning, it has seemed to me that the most satisfactory theory would be one which made the beginning not too unaesthetically abrupt." In 1931, in "The expansion of the universe," he wrote, "Philosophically, the notion of a beginning of the present order of Nature is repugnant to me." His aversion to a big bang was shared by others, including advocates of the steady-state universe.*

Cosmic birth and death (Chapter 25) were common notions in mythology, and fears of such ideas may (I am guessing) stem from the way that people now live. Most of us are no longer members of extended family communities, surrounded by relatives, young and old, among whom birth and death are common events. We live instead singly or in small families, isolated from one another, and birth and death are unfamiliar events occurring out of sight in hospitals. Eddington, who was outspoken in his dislike of cosmic birth and rejected catastrophic cosmic death, was a bachelor who lived with his sister. When a cosmologist presents an argument for a particular type of universe, perhaps we should not read too much into the science but wonder about that person's religion, philosophy, and even psychology.

10 *The Hubble sphere contains all galaxies receding subluminally (less than the velocity of light); galaxies at the edge of the Hubble sphere recede transluminally (at the velocity of light); and galaxies outside the Hubble sphere recede superluminally (greater than the velocity of light). The edge of the Hubble sphere at radial distance* $L_H = c/H_0$ *recedes at velocity* $U_H = \mathrm{d}L_H/\mathrm{d}t$, *or*

$$U_H = c(1+q). \qquad [14.29]$$

Galaxies at the edge recede at the velocity of light c, and the edge overtakes the galaxies at relative velocity

$$U_H - c = cq.$$

In all decelerating universes ($q > 0$), *the Hubble sphere expands faster than the universe and contains an increasing number of galaxies. In all accelerating universes* ($q < 0$) *the Hubble sphere expands slower than the universe and contains a decreasing number of galaxies. If* N_H *is the number of galaxies in the Hubble sphere, it can be shown that*

$$\frac{\mathrm{d}N_H}{\mathrm{d}t} = 3qHN_H. \qquad [14.30]$$

This expression cannot be used in the steady-state universes of $q = -1$ *because of the continual creation of new galaxies; in this universe* U_H *and* N_H *are constant.*

11 *The metric equation for the distance dL between two adjacent points in a spherical space of curvature* $1/R^2$ *and comoving coordinates a,* θ*, and* ϕ *is*

$$\mathrm{d}L^2 = R^2[\mathrm{d}a^2 + \sin^2 a(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)],$$

from Equation [10.15]. The coordinates a, θ*, and* ϕ *are physically dimensionless (as for colatitude and longitude on the surface of a sphere), and with distances measured in light-travel time we see that R has the dimensions of time. By changing the symbol a into r, we have*

$$\mathrm{d}L^2 = R^2[\mathrm{d}r^2 + \sin^2 r(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)]. \qquad [14.31]$$

To any point the distance from an arbitrary origin $r = 0$, *is* $L = R\int \mathrm{d}r = Rr$ *and the comoving coordinate distance is simply r. In Equation [14.31], R becomes the scaling factor that varies in time, and r,* θ*, and* ϕ *are the fixed coordinates of comoving bodies. More generally,*

$$\mathrm{d}L^2 = R^2[\mathrm{d}r^2 + S^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)], \qquad [14.32]$$

in which

$S = r$:	flat space ($k = 0$)
$S = \sin r$:	spherical space ($k = 1$)
$S = \sinh r$:	hyperbolic space ($k = -1$)

and Equation [14.32] applies to the three homogeneous and isotropic spaces of curvature constant k and curvature k/R^2.

The relativity line-element expressing ds in terms of intervals of time dt and space dL is

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \mathrm{d}L^2, \qquad [14.33]$$

as in Equation [11.4]. With dL from Equation [14.32] and dt understood as an interval in cosmic time, we obtain the famous Robertson–Walker line element

$$\mathrm{d}s^2 = \mathrm{d}t^2 - R^2[\mathrm{d}r^2 + S^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)], \qquad [14.34]$$

anticipated by many cosmologists, notably by Georges Lemaître, Howard Robertson, Richard Tolman, and Arthur Walker. Because of the study of its significance by Robertson and Walker it is referred to as the Robertson–Walker geometry, or metric, or line element. Various forms of Equation [14.34] are obtained by transforming the radial coordinate r. The form shown is most convenient because the radial coordinate distance r corresponds to a linear tape-measure distance.

- *Much of modern cosmology flows from the Robertson–Walker (or R–W) line element. At this stage we mention three issues:*

(i) The R–W equation defines cosmic time t as orthogonal to uniformly curved, expanding space.

(ii) The R–W equation automatically yields the velocity–distance law. Let us arrange the coordinate system such that we are at the origin $r = 0$. A comoving galaxy at distance $L = Rr$ recedes at velocity $V = \mathrm{d}L/\mathrm{d}t$. Because r is constant for the galaxy, the change in its distance in time dt is

$$\mathrm{d}L = r\,\mathrm{d}R = L\,\mathrm{d}R/R = LH\,\mathrm{d}t,$$

where $H = \dot{R}/R$, and hence the recession velocity of the galaxy is $V = LH$. At a fixed value of H (and therefore at a fixed instant in cosmic time) the recession velocities of all galaxies are proportional to their tape-measure distances L. Thus we see that the velocity–distance law is implicit in the R–W line-element because it applies to spaces of time-invariant homogeneity.

(iii) The R–W line element enables us to relate the world picture (on the backward lightcone) and the world map (in which the velocity–distance law applies). Light and anything moving at the velocity of light propagates on null-geodesics defined by $ds = 0$. Again we assume for convenience that we are at the origin $r = 0$ and consider radial rays of light for which $d\theta = 0$, $\mathrm{d}\phi = 0$. The R–W line element, with $\mathrm{d}s = 0$, reduces to $\mathrm{d}t = \pm R\,\mathrm{d}r$ for increments in radial distance, and the plus sign applies to outgoing rays on the forward lightcone and the negative sign to incoming rays on the backward lightcone. Hence, a coordinate distance on the backward lightcone (the world picture) is given by

$$r = \int_t^{t_0} \mathrm{d}t/R, \qquad [14.35]$$

where t is the time of emission of the ray observed at the present time t_0. Thus the actual distance to the source at the present time is

$$L = R_0 r = R_0 \int_t^{t_0} \mathrm{d}t/R, \qquad [14.36]$$

and its distance at the time of emission is

$$L_{\mathrm{emit}} = Rr = R\int_t^{t_0} \mathrm{d}t/R = (R/R_0)L. \qquad [14.37]$$

These equations help us to relate observed distances in the world picture to tape-measure distances in the world map and provide the bridge that must be crossed to determine H_0 and q_0 from observations.

12 *The way in which the scaling factor R varies in time must be found by means of a dynamic model of the universe. This will be discussed in later chapters. In the power-law models, R varies as t^n, where n is a constant. Of these bang–whimper universes, the most important is the Einstein–de Sitter model of $n = 2/3$. If we assume that*

$$R = R_0(t/t_0)^n, \qquad [14.38]$$

we find

$$H = n/t, \qquad [14.39]$$

$$q = (1 - n)/n, \qquad [14.40]$$

hence $n = (1 + q)^{-1}$. The subscript zero is added when we wish to denote present values. The deceleration q is constant in all power-law models. The universe expands when $H > 0$, or $n > 0$, and decelerates when $q > 0$, or $n < 1$. The Hubble length in the power-law models is

$$L_H = ct/n = ct(1 + q), \qquad [14.41]$$

and the Hubble time is

$$t_H = t/n = t(1 + q). \qquad [14.42]$$

Notice that $t/t_H = 1/(q+1)$, *illustrating the general rule that the age of the universe* t *is less than the Hubble time* t_H *in a decelerating universe of* $q > 0$. *Finally, the Hubble sphere expands at velocity*

$$U_H = c/n = c(1+q). \quad [14.43]$$

In the Einstein–de Sitter model of $n = 2/3$, *we find* $q = 1/2$, $L_H = 3ct/2$, $t_H = 3t/2$, $U_H = 3c/2$, *and the edge of the Hubble sphere overtakes the galaxies at* $c/2$.

Using Equations [14.35] and [14.38], we find that the present distance of an observed body that emitted light at time t *is*

$$L = \frac{n}{1-n} L_H (1 - x^{1-n}), \quad [14.44]$$

and the distance of this body at the time it emitted the light that we now see is

$$L_{\text{emit}} = \frac{n}{1-n} L_H x^n (1 - x^{1-n}), \quad [14.45]$$

where $x = t/t_0$. *These equations apply in a big bang universe of* $0 < n < 1$. *We see from Equation [14.44] that the maximum present distance (the maximum value of* L*) of a body is* $nL_H/(1-n) = t_0/(1-n)$, *and this value, because of expansion, is more than the maximum distance* t_0 *in a static universe of* $n = 0$. *And from Equation [14.45] we see that the emission distance is zero when* $t = t_0$ *and also when* $t = 0$.

PROJECTS

1 Take an elastic cord, fix markers to it, such as clothes pegs, and slowly stretch it, as in Figure 14.25. The comoving markers represent the galaxies in an expanding universe. Notice that distances are the tape-measure kind and that the markers move apart according to the velocity–distance law. We see how separating distances between the markers are related to comoving coordinate distances by means of a scaling factor. We see also that when stretched slowly or rapidly the elastic string demonstrates the change in time of the scaling factor.

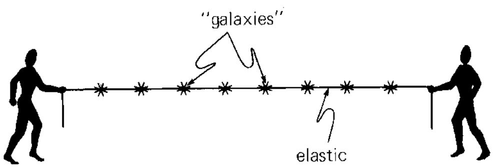

Figure 14.25. A one-dimensional expanding universe consisting of an elastic string having attached markers representing galaxies. The markers have fixed comoving coordinates and their relative motions illustrate the velocity–distance law.

2 The term big bang was first used by Fred Hoyle in his series of BBC radio talks on astronomy published in *The Nature of the Universe*, 1950. "This big bang idea seemed to me to be unsatisfactory even before examination showed that it leads to serious difficulties. For when we look at our own Galaxy there is not the smallest sign that such an explosion ever occurred. This might not be such a cogent argument against the explosion school of thought if our Galaxy had turned out to be much younger than the whole Universe. But this is not so. On the contrary, in some of these theories the Universe comes out to be younger than our astrophysical estimates of the age of our own Galaxy. . . . On philosophical grounds too I cannot see any good reason for preferring the big bang idea. Indeed it seems to me in the philosophical sense to be a distinctly unsatisfactory notion, since it puts the basic assumption out of sight where it can never be challenged by direct appeal to observation." Discuss Hoyle's remarks.

3 Let the smoothed-out density of the universe be equivalent to 1 hydrogen atom per cubic meter. Now gather this matter together into uniformly distributed marbles and find their separating distance. Gather the matter together into uniformly distributed stars similar to the Sun and find their separating distance. Now do the same for galaxies of 10^{11} solar masses.

4 Explain why, in a universe 1 year old, we cannot see farther than a distance of approximately 1 light year.

5 What is wrong with the idea of a universe beginning at a point in space?

6 Explain the difference between recession velocity and ordinary or peculiar velocity, and comment on the following remark taken from a textbook in astronomy: "There are mathematical models of the universe that have galaxies ... going even faster than the velocity of light. Of course, the laws of relativity forbid this, and such models are only of academic interest."

7 Draw diagrams showing how the scaling factor R may vary in time. Show in a single diagram the variation of R with time in an imaginary universe that passes through successive periods of expansion, contraction, deceleration, and acceleration, and label these periods with $H > 0$, $H < 0$, $q > 0$, and $q < 0$.

8 In what universe is (a) R constant? (b) H constant? (c) q constant?

9 Derive Hubble's redshift–distance law from the mathematicians' velocity–distance law. Explain the approximations used. What is the distance of galaxies of redshift $z = 0.01$ and 0.1? (Give the answer in terms of h and in light years.)

10 Show that when $R = t$, and therefore $H = 1/t$, $q = 0$, the Hubble sphere contains a constant number of galaxies.

FURTHER READING

Davies, P. C. W. *The Runaway Universe*. Dent, London, 1978.

Eddington, A. S. *The Expanding Universe*. Cambridge University Press, Cambridge, 1933.

Ellis, G. F. R. "Innovation, resistance and change: the transition to the expanding universe," in *Modern Cosmology in Retrospect*. Edited by B. Bertotti et al. Cambridge University Press, Cambridge, 1990.

Gamow, G. *The Creation of the Universe*. Viking Press, New York, 1952.

Hetherington, N. S. "Hubble's cosmology." *American Scientist* 78, 142 (1990).

Huchra, J. P. "The Hubble constant." *Science* 256, 321 (1992).

North, J. D. *The Measure of the Universe: A History of Modern Cosmology*. Oxford University Press, Clarendon Press, Oxford, 1965.

Sandage, A. R. "Cosmology: a search for two numbers." *Physics Today* (February 1970).

SOURCES

Andrade, E. N. da C. "Doppler and the Doppler effect." *Endeavour* p. 14 (January 1959).

Brush, S. G. *The Temperature of History: Phases of Science and Culture in the Nineteenth Century*. Franklin and Company, New York, 1978.

Burbidge, G. "Modern cosmology: the harmonious and discordant facts." *International Journal of Theoretical Physics* 28, 983 (1989).

Eddington, A. S. "On the instability of Einstein's spherical world." *Monthly Notices of the Royal Astronomical Society* 90, 668 (1930).

Eddington, A. S. "The evolution of the universe." Symposium. *Nature*, Supplement, p. 699 (October 24, 1931).

Eddington, A. S. "The expansion of the universe." *Monthly Notices of the Royal Astronomical Society* 91, 412 (1931).

Eddington, A. S. *The Expanding Universe*. Cambridge University Press, Cambridge, 1933.

Ellis, G. F. R. "The expanding universe: A history of cosmology from 1917 to 1960," in *Einstein and the History of General Relativity*. Editors, D. Howard and J. Stachel. (Einstein Study Series.) Birkhauser, Boston, 1988.

Ellis, G. F. R. and Brundrit, G. B. "Life in the infinite universe." *Quarterly Journal of the Royal Astronomical Society* 20, 37 (1974).

Friedmann, A. "On the curvature of space." *Zeitschrift für Physik* 10, 377 (1922). "On the possibility of a world with constant negative curvature." *Zeitschrift für Physik* 21, 326 (1924). Both translated in *Cosmological Constants*. Editors, J. Bernstein and G. Feinberg. Columbia University Press, New York, 1986.

Harrison, E. R. "A century of changing perspectives in cosmology." *Quarterly Journal of the Royal Astronomical Society* 33, 335 (1992).

Harrison, E. R. "Hubble spheres and particle horizons." *Astrophysical Journal* 383, 60 (1991).

Harrison, E. R. "The redshift–distance and velocity–distance laws." *Astrophysical Journal* 403, 28 (1993).

Hoyle, F. *The Nature of the Universe*. Blackwell, Oxford, 1950.

Hubble, E. *The Realm of the Nebulae*. Yale University Press, New Haven, 1936.

Hubble, E. *The Observational Approach to Cosmology*. Oxford University Press, Clarendon Press, Oxford, 1937.

Huggins, W. "Further observations on the spectra of some stars and nebulae, with an attempt to determine therefrom whether these bodies are moving towards or away from the Earth." *Philosophical Transactions* 158, 529 (1868).

Hujer, K. "Sesquicentennial of Christian Doppler." *American Journal of Physics* 23, 51 (1955).

Kragh, H. *Cosmology and Controversy: The Historical Development of Two Theories of the Universe*. Princeton University Press, Princeton, 1996.

McCrea, W. H. "Willem de Sitter, 1872–1934." *Journal of the British Astronomical Association* 82, 178 (1972).

Metz, W. D. "The decline of the Hubble constant: a new age for the universe." *Science* 178, 600 (1972).

Milne, E. A. "A Newtonian expanding universe." *Quarterly Journal of Mathematics*, 5, 64 (1934).

Milne, E. A. *Relativity, Gravitation and World Structure*. Oxford University Press, Oxford, 1935.

Murdoch, H. S. "Recession velocities greater than light." *Quarterly Journal of the Royal Astronomical Society* 18, 242 (1977).

Nietzsche, F. See Brush, S. G.

Robertson, H. P. "On relativistic cosmology." *Philosophical Magazine* 5, 835 (1928).

Robertson, H. P. "The expanding universe." *Science* 76, 221 (1932).

Sandage, A. R. "Observational cosmology." *Observatory* 88, 91 (1968).

Sandage, A. R. "Distances to galaxies: the Hubble constant, Friedmann time, and the edge of the world." *Quarterly Journal of the Royal Astronomical Society* 13, 282 (1972).

Sitter, W. de. "The expanding universe." *Scientia* 49, 1 (1931).

Sitter, W. de. *Kosmos*. Harvard University Press, Cambridge, Massachusetts, 1932.

Smith, R. W. "The origin of the velocity–distance law." *Journal for the History of Astronomy* 10, 133 (1979).

Tolman, R. C. "The age of the universe." *Reviews of Modern Physics* 54, 374 (1949).

Weyl, H. *Space, Time, Matter*. Methuen, London, 1922.

Whitrow, G. J. "Hubble, Edwin Powell." *Dictionary of Scientific Biography* 6, 528 (1967).