Supernovae and the Universe

Probing the Size, Shape, and Fate of the Universe with Supernovae

1. Our Expanding Universe

Distant galaxies, those so far away that, unlike the Magellanic Clouds, or our sister spiral Andromeda, we do not not sense their individual gravities, are moving away from us. Their speed is nearly proportional to their distance. One can get this effect by setting off a bomb. The faster fragments get further away in a given amount of time so, at a later instant, the faster fragments are further away with a distance that depends linearly on the speed. This, Einstein has taught us, is not how the Universe works. The bomb analogy requires their to be a pre-existing space, independent of the matter in the "bomb," into which the bomb explodes. Einstein has taught us, as we explored in Chapter 9, that space is a curving, dynamical entity that is shaped by the gravitating matter within it. Pre-existing empty space with a bomb in the center makes no sense mathematically or conceptually in Einstein's Universe.

Rather, Einstein taught us that space itself can expand, carrying the essentially motionless galaxies apart. In this manner, all distant galaxies, those that do not share an immediate gravitational grip, move away from all others. There is no center of the explosion. The fact that we see all distant galaxies moving away from us is an effect created by the uniform expansion of space. With some thought, you can convince yourself that the apparent speed with which galaxies recede depends linearly on the distance, just as observed.

We expected this expansion to be slowing down. This is because the Universe is filled with matter that exerts gravity. For seventy years or so, the challenge to cosmology was to determine whether the expected gravitational deceleration was enough to halt the expansion, or too little, so the Universe would continue to expand, but at an ever slower rate. One of the major glories of science is that with proper attention to Nature, preconceived notions as powerful as these can be overcome. It worked in this case!

2. The Shape of the Universe

To use supernovae or any other technique to measure cosmological distances requires some perspective on what we are trying to accomplish and how we are doing the task. Recall from Chapter 9 the various two-dimensional analogs we have employed to picture curved space. The two-dimensional space around a gravitating object is funnel-like when viewed from the perspective of three dimensions. The two-dimensional analog of the Universe itself, at one moment of time, can be represented as the surface of a sphere, an infinite flat plane, or a saddle extending upward to infinity fore and aft and downward to infinity sideways, as shown in Figure 12.1. These two-dimensional analogs are the embedding diagrams for the Universe. They help picture curvature in three dimensions. These two-dimensional surfaces have no two-dimensional centers, no two-dimensional edges, and no two-dimensional outsides. Likewise, for the most basic conceptions of our real three-dimensional Universe, there is no three-dimensional center, no three-dimensional edge, and no three-dimensional outside.

We have stressed that looking down on a two-dimensional embedding diagram from a higher, three-dimensional perspective is cheating in a sense because there is no way we can look down on our three-dimensional curved space from an "outside." That outside to our three-dimensional Universe, by analogy, would itself have to be a fourth spatial dimension. If there were an observer in that fourth spatial dimension, that observer could see the curvature of our Universe or that around the Earth or around a black hole in much the same way that we can see the curvature of the surface of a sphere. On a more direct and personal level, such an observer would also not be limited to viewing our surfaces, our skin, and our facial features as we do one another. An observer from a hypothetical fourth dimension would also be able simultaneously to see our volume, our guts, and our bones, much as we can see the interior of a circle inscribed on a sheet of paper. This is an amusing perspective, but it is not one of physics. Not until Chapter 14, at least.

Rather, the proper perspective is to recognize that a two-dimensional creature living in any of these curved two-dimensional spaces of Figure 12.1 could determine that the space curves and by how much by doing geometry, by carefully measuring distances and angles. That is now our task! We are three-dimensional supernova observers trapped in our three-dimensional Universe. We must determine the curvature of our three-dimensional space without stepping outside of three dimensions, something we simply cannot do. Fortunately, we do not need to step outside. We just have to be careful with our geometry and our astrophysics.

3. The Age of the Universe

The Universe we see around us began in what we call the big bang. There are still mysteries surrounding how the Universe came to be. We will touch on some of them in Chapter 14. There is, however, no doubt that the visible Universe arose in a very dense, hot state, and expanded outward. Although the first instants are murky, ordinary particles, protons and electrons formed very quickly, and the Universe was pure hydrogen for a while. The light elements – helium, lithium – formed when this expansion was a few minutes old. When it was a million years old, the matter got sufficiently dilute that the radiation from its heat could stream freely. We see that radiation as the *cosmic background radiation* that comes at us from all directions. This cosmic radiation is red-shifted by the expansion that pulls everything in the Universe away from everything else. We understand this process very well. Further expansion of the Universe brought the agglomeration of matter into galaxies, stars, and planets in ways we are still striving to understand. Continued expansion pulls all the distant galaxies apart. Understanding the expansion of the Universe allows us to measure its age.

As emphasized in Section 1, it is important to realize that the big bang did not occur as an explosion in a preexisting space, like a bomb in outer space. Rather space itself expanded, carrying the matter with it. One popular analogy is the behavior of spots on the surface of an expanding balloon. The spots do not move with respect to the rubber surface as the balloon expands, but they become ever farther apart, as shown in Figure 12.2. A three-dimensional analogy is raisins in a rising loaf of bread. The raisins never drift in the dough, but again move ever farther apart until the loaf stops rising. The second analogy is limited and a little deceptive because the loaf of bread is finite. The three-dimensional loaf of bread is surrounded by ordinary three-dimensional space into which it expands, whereas the space of the Universe is all encompassing. The first analogy is limited because it is restricted to two dimensions, but it is more revealing in a way. One can see that the two-dimensional surface of the balloon has no two-dimensional outside, neither the outside as we understand it from our three-dimensional perspective nor what we regard as inside the balloon, which still requires going off into a third-dimensional "hyperspace" from the perspective of a two-dimensional creature inhabiting the two-dimensional surface. Likewise, the loaf of bread is perceived to have a center, whereas (ignoring the opening though which one blows) there is no two-dimensional center to the two-dimensional surface of a perfect sphere to which the balloon is an approximation. Unlike the loaf of bread, the balloon shows that if attention is restricted to the confines of the dimensions of the space, two for the surface of the balloon, three for our Universe as we perceive it, there is no center, there is no edge, and there is no outside. These are tricky and fascinating issues, and we will return to them in Chapter 14.

For our current purposes, it is sufficient to picture the expansion of the balloon and its dots or the bread and its raisins to understand how to measure the age of the Universe. The effect of the expansion of the Universe is still much the same as an explosion in preexisting space even if the concepts are radically different. If you can measure how far away something is from you, say a distant supernova, and determine how fast it is traveling away from you, by measuring its Doppler shift to the red, then you can tell how long it has been traveling to get as far as it has. You get the same answer for every supernova and every galaxy. The faster they move away from us, the more distant they are, but they took the same time to get there, drawn by the expansion of the underlying space.

The parameter that is measured in this way is called the *Hubble constant*, after Edwin Hubble who pioneered this sort of measurement of distances and determined the nature of the Universal expansion. The Hubble constant tells you how fast something will be moving away from you at a given distance. Techniques for measuring the distances to Type Ia supernovae outlined in Section 5, and other techniques as well, say that velocity will be about 65 kilometers per second for every million parsecs in distance. The age is related to the inverse of the Hubble constant. Obtaining the age of the Universe from the Hubble constant involves another subtlety because it depends on the curvature of space and the acceleration of the Universe. Neglecting that subtlety for the moment, the corresponding age of the expanding Universe is roughly just the inverse of the Hubble constant. If a supernova moving at 65 kilometers per second is 1 million parsecs away, it must have been moving away from us for about 10 billion to 15 billion years. If another supernova is moving away from us at 650 kilometers per second and is at 10 million parsecs, then the time for it to get there is just the same, 10 billion to 15 billion years. We get the same answer for every supernova, as we must because we are measuring the same age in every case, the age of the Universe.

The best current estimate is a remarkably precise 13.7 billion years based on measurement of the cosmic background radiation (Section 6??????). The age estimated in this way does not depend on a detailed determination of the shape of the Universe. Whether our Universe is closed and finite in space and time, or open and infinite, its current age is about 14 billion years.

4. The Fate of the Universe

The game is not over with the measurement of the Hubble constant. It is not enough to measure how old the Universe is. We want to know what will happen to it in the future. Since the days of Hubble, astronomers, particularly the subset known as cosmologists, have been engaged in a grand quest to determine the "fundamental parameters of the Universe." This quest was shaped by Einstein's theory of gravity. The first attempts to apply Einstein's theory to the whole Universe showed that there were three parameters that would describe the whole shebang: the Hubble constant, the overall curvature of the Universe, and the rate at which the Universe is changing its speed of expansion due to the gravitational pull of the matter and energy within it. The issue of curvature is whether the Universe is the three-dimensional analog of the surface of a sphere, a flat plane, or a saddle, as shown in Figures 12.1 and 12.2. Einstein's theory showed that it had to be one of the three. Furthermore, with a key, but reasonable, simplifying assumption that the Universe had the same content, on average, everywhere, the theory showed that the fate is tied to the geometry. If the Universe were sphere-like, it would have a finite life and re-contract to a singularity; if it were flat, it would expand forever, just reaching zero expansion rate at the end of time; and if it were saddle-like, it would expand forever at a finite velocity. We will see later in this chapter and in Chapter 14 that these three parameters may not tell the whole story, but they make up a critical part of it. Determining these parameters occupied cosmology for most of the twentieth century.

5. Dark Matter

There are various ways of going about measuring the other two parameters in addition to the Hubble constant. The underlying theory requires the constraint of two specific quantities. One is the mass density of the gravitating matter in the Universe at the current epoch. In its simplest guise, this means determining the total mass of all kinds of stuff that has a finite mass and does not move at the speed of light. This mass includes stars, planets, and dust, but it also means any component of the mysterious *dark matter* that consists of particles, no matter how exotic. The photons of light that permeate the Universe also count. They have a mass-equivalent energy ($E = mc^2$), but the gravitational affect of this energy alone is small. The other quantity to be constrained (and ultimately measured) is the value of what is called the *vacuum energy density*. Recall that even a vacuum has an energy associated with it. This energy underlies the emission of Hawking radiation from black holes. The vacuum may have even more subtle properties that would only be manifested when its effects are determined on the scale of the whole Universe.

Dark matter is stuff that gravitates, but emits no detectable light. By detecting the gravitational effects of dark matter on the stars and gas that we can see, we have determined that there is four or five times more of this stuff in the Universe than of what we think of as ordinary matter composed of protons, neutrons, and electrons; that is to say, ordinary matter like stars, planets, and people. Most of the mass of this "ordinary" matter is in protons and neutrons, the low-mass electrons contribute little to the total, so this component is known generally as the baryonic (Chapter 1) component of the Universe. Baryonic matter gravitates, but also, in proper circumstances, shines. That is how we find it. The dark matter gravitates, that is how we detect it. On the other hand, it must not have an electrical charge, or it would create electromagnetic radiation, light. Nor can it react by means of the strong nuclear force or it would behave far differently. The best guess is that it is composed of some particle, like a neutrino, only different, that reacts only to gravity and the weak nuclear force. There are on-going experiments to try to detect a particle of dark matter, but there have been no unambiguous results.

One might wonder whether the dark matter could be black holes. The answer is no. The ratio of hydrogen to helium that emerged from the big bang depends on the amount of proton/neutron-like stuff, the amount of baryons (Chapter 1). The observed ratio of hydrogen to helium says that there never was enough baryonic matter to account for all the dark matter, whether or not some of the baryons later fell into black holes. The dark matter is something different and something special, and it is the truly "ordinary" matter in the Universe; stuff like us is rare to the point of insignificance when it comes to determining the gravitational heft of the Universe. On the other hand, baryons, arranged into people, can think about the Universe, and the dark matter, undoubtedly, cannot.

The Dark Matter has played an amazing role in the Universe, given that we cannot see it. The *Cosmic Background Explorer* (*COBE*) satellite launched in 1989 revealed that the cosmic background radiation left over from the big bang is of an exceedingly well-defined temperature, as expected. *COBE* also revealed faint irregularities in the temperature of the radiation from different parts of the sky. The *Wilkinson Microwave Anisotropy Probe*, or *WMAP*, launched in 2001, has provided the best measurement yet of those minute, but systematic fluctuations in the cosmic background radiation. These fluctuations were also expected and even inevitable, given our understanding of the big bang. The big bang grew out of a "singularity." That singularity must have been subject to quantum fluctuations in its properties that are imposed on the expansion of the Universe and hence on the density and temperature of the matter in the Universe (Chapter 14, Section 2). Detection of these irregularities at the level of one part in one hundred thousand was another major vindication of the big bang picture. The original explosion of the big bang left the same incredibly tiny quantum irregularities in the density of the dark matter, slight over-concentrations separated from pockets of ever so slight paucity.

As the Universe expanded, those density irregularities in the Dark Matter grew. When the Universe became transparent at the beginning of the Dark Ages (Chapter 11, Section 7) when it was only a million years old, these slight wrinkles in density deviated from the average by only one part in one hundred thousand. Yet those irregularities continued to grow and became large pockets of high and low density. Those rare protons, neutrons and electrons fell into the high density pockets of dark matter. The protons, neutrons, and electrons, in turn, formed the stars and galaxies we see scattered through the Universe. The whole structure of the Universe at which we can marvel now, and on which we depend for our existence, came from these initially tiny wrinkles in the dark matter that, in turn, trace back to the fluctuations of quantum uncertainty at the beginning. This is a truly amazing creation story, one backed by ever more detailed observational confirmation.

6. Vacuum Energy – Einstein's Blunder That Wasn't

There is also a story behind the vacuum energy. The vacuum energy is, in principle, related to the quantum properties of the vacuum, but something like it arises in Einstein's theory of gravity where it is called the *cosmological constant*. Astronomers who write the history of this subject tend to quote Einstein himself in this regard with great glee. Einstein called the cosmological constant "the greatest blunder of my life." The historian's glee and Einstein's self-criticism are probably unfair. The cosmological constant emerges from the mathematics of Einstein in a perfectly natural way (it appears as a constant of integration, for those who know calculus). It is not a question of whether it exists in this mathematical sense. It certainly does. The issue is whether it is zero or not, and whatever its value, including zero, what the physics is that determines that value.

The reason Einstein regarded his treatment of the cosmological constant to be a "blunder" is that his first mathematical models for the Universe showed that the Universe would contract or expand. Einstein's intuition told him that the Universe could not possibly do such a radical thing. To render the solution static, Einstein went back to the equations and realized that he had implicitly set the value of the cosmological constant to zero. If he assigned it just the right nonzero value, then the cosmological constant could serve as an extra effect to balance the tendency of the Universe to expand or contract. Shortly afterward, Hubble proved that the Universe *is* expanding. It appeared to Einstein that the cosmological constant was unnecessary, a blunder.

Einstein may have blundered in guessing that the Universe was static, and hence in the value to which he set the cosmological constant, but he did not blunder in introducing the idea. In the long run, it is the latter that is more important, and another tribute to the power of Einstein's theory. The blunder was much less than it is often made out to be. We now see that even the issue of whether the cosmological constant might be exactly zero is not a trivial one, but one that involves some of the deepest thinking about the Universe. More than that, there are hints that the cosmological constant is not zero, and that definitely raises profound issues of physics and cosmology.

7. Type Ia Supernovae as Calibrated Candles and Understood Candles

Apart from their intrinsic interest as star-destroying explosions, supernovae have other uses simply because they are so bright. Their great luminosity means that they are visible across the Universe. More specifically, supernovae are signposts that determine the distances to their host galaxies. Careful measurements of those distances allow astronomers to map out how fast the Universe is expanding and hence how old it is, the curvature of space, and clues to the fate of the Universe. The use of supernovae in this way has expanded extensively in the last decade and the results have been dramatic. Supernovae have provided clues that the Universe may expand forever and that it is even now in the grip of powerful repulsive forces that accelerate its outward rush.

The use of supernovae to measure distances is based on a simple principle: things farther away look dimmer. Turned around, how dim a supernova appears to be is a measure of how far away it is. The basis for this intuitively reasonable notion is that, when light spreads out from a central source equally in all directions, the locus of the photons emitted at a given time defines a larger and larger surface. The light falling on a detector of a given area, a human eyeball or a telescope, then captures a smaller and smaller fraction of the total the farther away the detector is from the source. The fraction decreases just as the total area into which the radiation floods increases and that goes like the distance squared (the area is $4\pi D^2$, where D is the distance; this turns out the be a profound and important statement, as we will explore in Chapter 14). This means that the apparent brightness of a source of a given total luminosity decreases like the inverse of the square of the distance. In simple terms, the fainter a given kind of object appears, whether it is a porch light, a star, or a supernova, the farther away it must be. If you know how bright the object really is, then you can tell from how bright it apparently is how far away it must be. This gives us a powerful tool for measuring distances. The key is to figure out how bright a given object really is.

Recall that Type Ia supernovae are generally the brightest of all the different types

(Chapter 6). This makes them especially good signposts for measuring large distances. If we knew exactly how bright they were, the task of measuring distances would be rather easy. We would just see how bright a supernova looked in a given telescope and read off the distance. The immediate problem is to determine the intrinsic brightness of a given supernova.

For a long time, there was some reason to believe that Type Ia supernovae were all equally bright. That would have made the task of measuring their distances particularly easy. The jargon for this is that such identical supernovae would represent a *standard candle*. The idea is that, if you have a set of "candles" of identical, known brightness, they can serve as a "standard" with which to compare other sources of luminosity and to measure distances. In the last decade, we have determined that Type Ia supernovae are not exactly the same, but that the differences are systematic. That allows astronomers to make allowances for the differences between individual Type Ia supernovae.

In particular, astronomers have found that the Type Ia supernovae that are intrinsically brighter decline in brightness more slowly than those that are intrinsically dimmer. We believe that we even have a basic understanding of why this is true. Some variation in the exploding white dwarf causes variation in the amount of radioactive nickel-56 produced in the explosion. The extra energy from radioactive decay does not just make the supernova brighter, it also keeps the expanding matter opaque longer. The radiation takes longer to leak out, giving the slower decay. The trend that relates the brightness of the supernova to the rate of decline from peak light gives the means to determine the brightness of the supernova. One just needs to see how fast the supernova declines, and that tells you how bright it really is. Comparison with how bright it seems in the telescope then gives the distance.

There are two ways of doing this comparison. One uses only the empirical data from the supernova with no attempt at a theoretical understanding. This method requires some comparison with other astronomical objects for which the distances are established in some other way. This calibration sets the overall scale of just how bright a Type Ia supernova with a given rate of decline really is. This must be done for as many supernovae as possible for which the distance is already known (beginning with a dozen or so, with the sample growing steadily). Then the brightness-decline relationship gives the intrinsic brightness and hence the distance from a measurement of the decline rate alone. This technique uses Type Ia not as standard candles but as light sources for which the brightness of each supernova can be calibrated compared with known sources, hence the phrase "calibrated candles."

The other technique to employ Type Ia supernovae to measure distances uses theoretical models of the explosions to determine how bright the supernova must be to produce a given light curve and spectrum. This technique thus attempts to employ "understanding" rather than "calibration" to provide the necessary information to turn the decline rate into a known intrinsic brightness. This technique thus uses Type Ia supernovae as "understood candles."

The first technique, using the Type Ia supernovae as calibrated candles, is only as good as the calibration and the implicit assumptions that underlie the empirical relation between peak brightness and the rate of decline. A key assumption is that the brightness-decline relation is unique. Two supernovae with identical decline rates are assumed to have the same intrinsic peak brightness. The second technique, using Type Ia as understood candles, is only as good as the rather complex underlying theory of the explosion and of the production of luminosity. This method can, in principle, allow for cases where, because of more subtle circumstances, other variables enter and two supernovae with the same decline rate do not have the same peak brightness. The two methods agree rather well. They both give the same age of the Universe (Section 3).

8. Supernovae and Cosmology

Using supernovae to determine the other fundamental parameters of the Universe has been a dream for decades. Many people have worked for a long time to bring it to pass. One of the pioneers, Stirling Colgate of the Los Alamos National Laboratory, estimated that to get the job done when he started working on an automated supernova search telescope in the early 1970s, he would have had to invent seven or eight brand new technologies. These included digital control of the telescope and its instrumentation, electronic detectors to replace photographic plates (Colgate called all this “dig-as” for digital astronomy; the tide of the digital revolution has fully enveloped astronomy by now, but the term never caught on), thin lightweight mirrors, time-sharing computers necessary for many people to work cooperatively on the complex computer code required to control the telescope and scan images, and cheap microwave links to allow remote control of the telescope from a distant site. The telephone company wanted $3 million for a microwave link from his telescope to the headquarters in Socorro, New Mexico. Colgate had only $3,000 for the job. He invented a simple method of error checking and installed the link with the funds and equipment he had.

In the 1990s, the technical capability, the development of critical techniques, and the willingness to devote a great deal of hard work have come together to bring this dream to fruition if not in quite the fully automated way Stirling Colgate envisioned. A key development has been the construction of large new telescopes and the special electronic detectors to record faint images over relatively large patches of the sky. Another was the launch, repair, and updating of the *Hubble Space Telescope*.

A team of astronomers at the Lawrence Berkeley Lab of the University of California, now headed by Saul Perlmutter, pioneered the breakthrough in technique. One of the inhibitions of research on supernovae is that their eruption is always a surprise. This means astronomers have to scramble to get data when an explosion occurs. Telescopes are often in the wrong configuration with the wrong instrumentation, the Moon is too bright to see the faint supernova light, or the weather is poor. The result is that we still do not get adequate information on most supernovae.

The LBL team realized that in certain circumstances they could discover supernovae “on cue.” They could then schedule procedures in advance to follow them up. These techniques work in precisely the context where one can use the resulting discoveries to do cosmology with supernovae. The trick is that if one looks out to very large distances, a given image obtained with a telescope spans a huge volume containing a huge number of galaxies. It is impossible to predict which of the many galaxies will produce a supernova, but if enough galaxies are in the image, one can be confident that some supernovae will erupt. It turns out that one does not even have to know which specific galaxies are there in advance. If one looks distant enough, there will always be plenty of galaxies and plenty of supernovae. The distances involved, billions of light years, are also just the distances astronomers needed to probe to learn about cosmology.

More particularly, the technique developed by the Berkeley team is to schedule time on a large telescope when the Moon is not up and the sky is dark. They obtain a first image of the sky. They then return and take another image of the same patch of sky 2 or 3 weeks later after the Moon has passed through its bright phase and is no longer a problem. They compare the second image to the first and look for any new lights in the faint images. This is not trivial because both the galaxies and the supernovae are very faint. Many person-decades have been invested in the computer codes that can automate this process and detect and eliminate flashes of man-made light, cosmic rays that strike the detector, asteroids, and other things that are just a nuisance for this project.

Nothing can be done about bum weather, but these procedures have brought the other factors under control. In addition to the LBL group, another group sprang up in competition led by Brian Schmidt of Mt. Stromlo Observatory near Canberra and comprising astronomers in Chile, at Harvard, and elsewhere. The results were striking. The two groups of astronomers guaranteed the discovery of roughly a dozen very distant supernovae each time they returned to take the second image. Because they knew far in advance when they would take the second image, they could coordinate the prior scheduling of other telescopes. In this way, they were prepared to get critical spectral and photometric information as soon as they determine the precise location of the new discoveries. Rapid global communication, including the Internet, also played a key role here. Both teams also used the *Hubble Space Telescope* to closely examine the host galaxies after the supernovae have faded. This is a critical step because one must subtract off the light of the host galaxy to get a pure signal from the supernova. Determining the light of the galaxy alone can be done efficiently after the supernova has faded, but not when the supernova first goes off and the light is a complex admixture of supernova and galaxy emission. This technique requires patience. Several months must pass before the supernova has faded sufficiently, and many more months are required for careful calibration and analysis. Using these techniques, the number of supernovae discovered per year has shot up to around 100, most of them at distances that span a good fraction of the observable Universe.

Recall from Section 7 that for a given intrinsic luminosity, the apparent brightness of a supernova declines as the inverse of the distance squared. This result, like the ratio of the circumference to the radius of a circle and the sum of the interior angles of a triangle, depends on the curvature of the underlying space. The power of the method of using supernovae is that they can, in principle, give such precise measurements of the distance at such great distances that the effects of the curvature of the space can be gleaned; whether the Universe has a curvature that is the analog of a sphere, a flat plane, or a very big Pringle. The results of these efforts shocked the worlds of astronomy, cosmology, and physics.

9. Acceleration!

As mentioned earlier, the amount of gravitating mass of all kinds in the Universe affects the curvature of the Universe and tends to slow down the expansion because of the mutual self-gravity of all the mass-energy. If the Universe is slowing down, then it was expanding more rapidly in the past. This means that, when we look at supernovae long, long ago and far, far away, with a given Doppler red shift, they will be a little closer and a little brighter than if the Universe had just been coasting at a constant speed, as shown in Figure 12.3. The Universe will also be younger than one would estimate from a given value of the Hubble constant and the assumption that the Universe had always expanded at the current rate.

That is all there is to it if the value of the vacuum energy density is zero. In the language of Einstein’s cosmological constant, if the cosmological constant is not zero, then the effect depends on whether the cosmological constant is positive or negative. If it were negative, the energy of the vacuum would add to the gravity of the matter and slow the expansion even more. If the cosmological constant were positive, the vacuum energy has the effect of an anti-gravitating, repulsive force causing the Universe to fly apart ever faster as it ages. That sounds like a strange effect, but it is possible within the framework of Einstein's theory and another measure of why the introduction of the cosmological constant was not a blunder but a very fascinating step. The mass density in the Universe must be positive, but the value of the cosmological constant could be positive or negative or zero and must be determined by observation or theory. If the cosmological constant were positive, it would act in the opposite way to the mass density. A positive cosmological constant would tend to make the Universe accelerate rather than decelerate, as shown in Figure 12.3. This means that a supernova at a given red shift will be a little farther away and a little dimmer than if the Universe had expanded at a constant rate. Likewise, the Universe would be a little older than one would estimate for a given Hubble constant and the assumption of a constant rate of expansion.

Because the effect of the positive mass density and of a positive vacuum energy density work in opposite directions to determine the dynamics of the Universe, the measurement of distances to supernovae tends to constrain the difference between the two effects. Using supernovae alone, the effects cannot be easily separated. Careful measurement of the apparent brightness and red shift of Type Ia supernovae of a given rate of decline and hence intrinsic brightness can, however, constrain the values of the mass density and the vacuum energy. From those constraints and a knowledge of the Hubble constant, the curvature of space and the rate of change of the speed of expansion of the Universe can also be estimated.

The measurement of distant supernovae gave two surprises. One was that there does not seem to be enough gravitating matter, mostly dark matter, to close the Universe. Other astronomical techniques give the same result. They all need to be further refined and considered, but astronomers have basically accepted the result. If this were all there were to it, the suggestion would be that the Universe did not contain enough stuff to close it, and hence it would expand forever.

The other result was even more surprising. Compared to the local sample of supernovae on which the calibration is done and compared to a Universe for which the vacuum energy is zero, the distant supernovae were a bit too dim. If this effect is caused purely by cosmological dynamics, then the implication is that the supernovae are a bit farther away for a given red shift. This effect, in turn, can only be explained if the Universe were not decelerating, nor even coasting, but *accelerating* its expansion! This was a striking and unexpected result. It was as if one tossed a ball in the air and rather than having it fall back into your hand, it raced every faster up into the sky! This expansion demands a finite and positive cosmological constant or an equivalent anti-gravitating effect of the vacuum energy density.

This result was so unexpected and dramatic, that there was an immediate frenzy to question the rather subtle results of the supernova work, not the least by the two teams among themselves and in the spirit of heated competition. The result has been a failure to impeach the result in any appreciable way. The distant supernovae are not materially different than nearby ones; there is no otherwise unexplained dust that could make them appear dimmer.

Even more important, other complementary, but completely independent, techniques have measured the same effect. The most significant is the careful measurement of the tiny fluctuations in the cosmic background radiation when the Universe became transparent at an age of a million years that were also imprinted in the Dark Matter (Section 5). Careful measurement of the fluctuations in the background radiation also constrain the matter density and the vacuum energy density. This technique is a critical complement to the research based on supernovae. The mass density and the cosmological constant tend to work in concert to make the fluctuations grow in amplitude as the Universe ages. The larger the mass density, the stronger the gravity and the faster the fluctuations will tend to grow. On the other hand, if there is a finite and positive vacuum energy density so that the Universe tends to accelerate, then the Universe will be a little older than it otherwise would be, other things being the same, and this gives the fluctuations more time to grow, again making them larger. The result is that the measurement of the fluctuations in the cosmic background will tend to measure the sum of the mass density and the effect of the vacuum energy density, whereas the supernova technique measures the difference between these quantities. Neither technique by itself gives the full picture. If, however, we have independent measures of both the sum and the difference of the mass density and the vacuum energy density, then, in an algebraic sense, we can solve for both unknowns. The incredible characterization of the fluctuations of the temperature of the cosmic background radiation by *WMAP* has provided a critical source of complementary information. The precise pattern of radiation fluctuations on the sky gives a measurement of the age of the Universe, the amount of gravitating Dark Matter, and a measure of this anti-gravitating effect.

Combining supernovae, *WMAP*, and other results, has given rise to a new *concordance model* of the Universe; a Universe composed of about 1/3 Dark Matter, about 2/3 of this anti-gravitating influence, and a small smattering of stuff like us for garnish.

10. The Shape of the Universe Revisited

Although the Dark Matter gravitates and the Dark Energy anti-gravitates, they contribute in similar ways to determine the total energy density that in turn determines the curvature of the Universe. What we have learned is that there is enough Dark Matter and baryonic matter to give about 1/3 of that needed to render the Universe flat. Now we have learned that there is enough Dark Energy to give about 2/3 of that needed to render the Universe flat. The total 1/3 + 2/3 = 1. Within current observational uncertainties, the best guess is that our Universe is flat, but accelerating!

This does not mean that our Universe is flat at absolutely each and every point. There are still real stars and black holes and galaxies that curve the space around them. This result means, rather, that when averaged over large volumes containing huge numbers of stars and galaxies, the average curvature is flat in three dimensions, the analogy of a flat plane, a space in which, on average, two initially parallel laser beams will always remain parallel and, if we could do the measurement, all very large triangles would always have their interior angles sum to 180 degrees.

Given the remaining uncertainties, we cannot rule out that the Universe is barely open or barely closed. There is an argument that it must be truly flat to extraordinary accuracy. This argument is based on the inflationary model of the Universe, that very early in its expansion, the Universe underwent a huge expansion in size, stretching all of space to a huge degree. The implication is that, whatever the shape might have been of the Universe before this, curved or flat, the final result would be essentially flat. This is equivalent to saying, for a two-dimensional surface, that if it were sufficiently large, we could not distinguish the curvature of a very large sphere or a very large saddle from a truly flat plane. The Earth seems flat to casual observation because it is so large compared to the human scale. Imagine the surface of the Earth blown up to the size of the observable Universe. If we entertain that the Universe might be just the teensiest open or the teensiest closed, its fate is even more uncertain, as we shall see below.

11. Dark Energy

In a very deep way, we do not know what this anti-gravitating influence is that is causing the Universe to accelerate. It has been given the name *Dark Energy*, a term that has caught on broadly, but is just a mask to hide our ignorance of what is going on. What we do know is some things the Dark Energy is not. It cannot be composed of any "normal" particle like protons, neutrons, and electrons, nor even the yet unknown particles of Dark Matter, because those all gravitate. We also know that the Dark Energy cannot be accounted for by any currently known theory of physics. The Dark Energy was not just a surprise to astronomers and cosmologists; it represents a challenge to fundamental physics. That got the attention of physicists and, among other things, has profoundly changed the nature of supernova research. Supernovae are no longer just the play thing of astronomers. Physicists now think supernovae, at least Type Ia, are their experiment with Nature.

The current guesses are that the Dark Energy is some sort of force field that permeates the vacuum and pushes or anti-gravitates. It is perhaps useful to picture the force field that arises when you try to push two magnetic north poles together. There is no magnetic "substance" in the space between the poles (this experiment would work perfectly well in the vacuum of outer space), but the repulsive force is palpable. The Dark Energy is not a magnetic field, but this example serves to illustrate that there could be some repulsive field permeating empty space. An important aspect of the Dark Energy pictured in this way is that, since it is a property of empty space, it does not get diluted as the Universe expands; the expansion just yields more space, more volume, and hence more Dark Energy. The amount of Dark Energy per cubic centimeter of empty space could be the same as the Universe expands, whereas the gravitating Dark Matter would be diluted by the expansion and its gravitating effect would be ever less.

Given this perspective and the assumption that the vacuum energy density is roughly constant, the prediction is that in the young Universe, the density of matter would dominate and the Universe would be decelerated by the gravity of that matter. The anti-gravitating effects of the Dark Energy would also be there, but too small in proportion to have much affect. At the Universe expands, however, the matter is diluted and its gravity becomes weaker. The Dark Energy remains undiluted, since it is a property of the empty space itself, and eventually there comes an epoch where the effect of the matter becomes less than that of the Dark Energy and the Universe begins to accelerate under that now dominating influence. Remarkably, Adam Riess of the Space Telescope Science Institute and his collaborators have used the *Hubble Space Telescope* to measure just such an effect. Even more distant supernovae, observed when the Universe was even younger, show the effects of deceleration. The acceleration of the Dark Energy took over about 5 billion years ago, when the Universe was about 2/3 of its present age, coincidently about the time our Sun was born. Why the Dark Energy should be of the value that its effects would be revealed about "now" in cosmological terms, is one of the mysteries associated with the Dark Energy.

By its mathematical appearance in Einstein's equations, the cosmological constant has a strictly imposed behavior. Inasmuch as the vacuum energy density is positive, the pressure associated with it must be negative and vary in exact proportion to the vacuum energy density. One can think of the negative pressure as the rough equivalent of a tension that pulls inward rather than a normal pressure that pushes outward. The latter gravitates; the former tends to anti-gravitate. The exact linearity between the pressure and the density if the Dark Energy is Einstein's cosmological constant gives a precise predicted behavior to the acceleration of the Universe. As far as we can tell from current observations, the Universe is behaving in exactly this way, as if the Dark Energy were exactly the same at all times and in all places in the Universe.

Even if it proves true that the Universe is behaving as if in the grip of exactly Einstein's cosmological constant, physicists will still want to know why the cosmological constant has the value it does in terms of fundamental quantum fields and forces. Physicists can estimate what the vacuum energy density should be based on the ideas of the vacuum energy associated with particle creation and annihilation in the vacuum, as invoked to understand Hawking radiation (Chapter 10, Section 6). Doing so gives an answer that is wrong by a factor of 10^{120}. My colleague Steve Weinberg calls this "the biggest mistake ever made by physicists." Physicists faced with this dilemma had long speculated that on the cosmological scale there was some cancellation of the local vacuum energy by some other force field that yielded exactly zero when applied to the whole Universe. Now they are faced with the dilemma that there must be some cancellation that is nearly perfect, but not quite. That is, conceptually, a much more challenging problem, but the one Nature has apparently delivered.

The Dark Energy thus raises profound questions about what the nature of the vacuum must be that it contains a quantum property that acts as a repulsive, anti-gravitating force. In the inflationary model of the Universe (Section 10), when the Universe was first born, it had a vacuum energy that did act as a repulsive force, an antigravity, that caused a piece of the Universe to expand vastly and rapidly to form the Universe we see today. According to the theory, this energy of the vacuum should have decayed away to zero by now. If the vacuum still has some of this repulsive energy, new theories of the vacuum will have to be developed.

The suggested constancy of the Dark Energy, though consistent with Einstein's cosmological constant, is itself a deep challenge to physics. In the most general terms, forces in physics have the feature that they will vary in time and space. One early version of such a theory, based on some of the tenets of string theory that we will explore in Chapter 14, was called *quintessence* by Paul Steinhardt of Princeton and his collaborators. The name came from the ancient Greek notion of a "fifth essence" (after earth, air, fire, and water), but in this case, it represented the possible behavior of a quantum field theory of the vacuum energy that would manifestly be varible in space and time.

The next big push to understand the Dark Energy will be to attempt to determine if, despite current indications, it does vary in time and space. Whatever the case, Dark Energy is neither predicted nor described by current theories of physics. Understanding Dark Energy is one of the great challenges to modern physics, a challenge the emerged from simply wondering just how far away we might see Type Ia supernovae.

12. The Fate of the Universe Revisited

This discovery of Dark Energy has also upset the cosmological game plan to discover the fate of the Universe by measuring the three fundamental parameters of cosmology as described in Section 4. It remains true that determining, directly or indirectly, the Hubble constant, the matter density, and the vacuum energy density, one can determine the shape of the Universe, open, closed, or flat. With a vacuum energy density, however, that information alone may not reveal the fate of the Universe.

If the Universe has a low gravitating matter density and finite, positive anti-gravitating vacuum energy density, as current results suggest, so that the tendency to coast outward is even accelerated, then infinite expansion is certainly suggested. In principle, however, a positive cosmological constant could continue to push the Universe into infinite expansion even if there were enough matter to close it, which there does not appear to be. If this were the fate of the Universe, the current "best guess," the Universe is doomed to expand into a dark oblivion. Galaxies would get so far apart that inhabitants of one could not see another. Stars would die out. Black holes would eventually evaporate by Hawking radiation. Current theories suggest that baryons and leptons, and probably the dark matter, would all decay to photons. The Universe would finally be this accelerating void filled with dim, dilute flashes of light.

If the acceleration of the Universe were slightly stronger than seems the case today, if the anti-gravitating effect were slightly more sensitive to the vacuum energy density than strictly proportional, then the acceleration itself would accelerate. This might suggest that the Universe would reach its dark oblivion even faster, but the implications are even more dire. If the Dark Energy behaves in this way, the prediction of Robert Caldwell of Dartmouth and his colleagues is the Universe would be subjected to a *Big Rip*, in which the growing acceleration would overcome the grip of gravity, pulling galaxies apart, then overcome electromagnetic forces, pulling molecules and atoms apart (ouch!), then overcome the strong nuclear force pulling nuclei apart, and then, finally, pulling space-time itself apart. Most physicists consider this possibility so repugnant, that they do not take it seriously.

On the other hand, given that the existence of a vacuum energy density raises issues of its origin that we clearly do not know how to answer, we cannot be sure that the cosmological constant is "constant." If this vacuum energy should switch signs and the effective cosmological constant become negative, then, again in principle, the Universe could be doomed to recollapse in a *Big Crunch*, even though it did not contain enough gravitating matter to accomplish that feat on its own. These results have opened up new, if misty, vistas in both cosmology and physics; and this is before we peer into hyperspace.