The History and Fate of the Universe:

*A guide to accompany the Contemporary Physics Education Project Cosmology Chart**

Lawrence M. Krauss, Case Western Reserve University, Cleveland, OH

Cosmology, the field that focuses on the origin and evolution of the universe on its largest scales, is undergoing what many have called a "golden era." New observations, on the ground, in the air, and in space, combined with exciting new theoretical insights, have, over the past decade or two, literally revolutionized our picture of the universe in which we live. Ideas that were essentially pure speculation 20 years ago now rest firmly on the bedrock of experiment. At the same time, many new questions have arisen, and some once firmly held notions about the future of the universe have been displaced. In this article I present a guide to our current understanding of the history and fate of the universe to parallel and supplement the overview presented in the Contemporary Physics Education Project Cosmology Chart.

The Big Bang

Let us begin at the beginning. The big bang, as it is now called, represents the initial configuration out of which the observable universe arose. We observe the universe to be expanding. Namely, the separation between distant galaxies has been increasing with time (galaxies are located about one million light-years apart on average throughout the visible universe, and there are about 4×10^{11} billion galaxies contained in the visible universe). According to Hubble's law, galaxies that are twice as far away from us are receding, on average, twice as fast.

Fig. 1. One can tell how far away each candle is both by the brightness of its flame and from its apparent size. A "standard candle" appears dimmer at greater distances.

Note: The Contemporary Physics Education Project urges readers to field test *The History and Fate of the Universe* chart in your classroom. Please download the evaluation form at http://cpepweb.org/fieldtest. The deadline for returning comments is July 7, 2003.

**The chart, entitled* The History and Fate of the Universe, *is included as a supplement to this issue.*

DOI: 10.1119/1.1557502

We determine the distant galaxies by, among other methods, using what are called *standard candles* (see Fig. 1). These are astrophysical objects whose intrinsic luminosity we believe we understand (because such objects are sometimes near to us, so that we can use independent distance measures to determine their luminosity). Then, from their apparent luminosities, we can also determine their distances, the light spreads out over a larger and larger sphere, and so its intensity decreases as R^{-2} (i.e., the inverse of the distance squared). Standard candles range from periodically variable stars, to exploding stars, called *supernovae*. We measure the relative velocity of distant objects, associated with the expansion of the universe, from their *redshift* (see Fig. 2). Using known spectral lines based on emission or absorption by various elements known to be in these systems, one can compare the observed frequency of these lines with the known frequencies as measured in the laboratory here on Earth. The ratio of these frequencies is related to the radial motion of these objects toward or away from us (as a fraction of the speed of light).

While the observations of galaxies moving away from us may make it seem as if we are at the center of the universe, it rather implies that space has been uniformly expanding in all directions. If we follow the evolution of the universe backward, using the presently observed expansion and the known laws of physics, the initial density of matter and radiation would have been infinite. This seems physically impossible, but at the very least as one approaches such a configuration, the laws of physics, as we now understand them, would break down. Thus, we do not have the tools to describe this initial configuration adequately, nor what caused the initial configuration to exist. It is reasonable to believe, however, because of the relationship between space and time, and the distribution of matter and energy in the universe that is governed by Einstein's equations of general relativity, that time itself may not have been a well-defined quantity at the incredibly high densities associated with the earliest moment of creation. Thus, asking what happened before the big bang may not be a well-defined question!

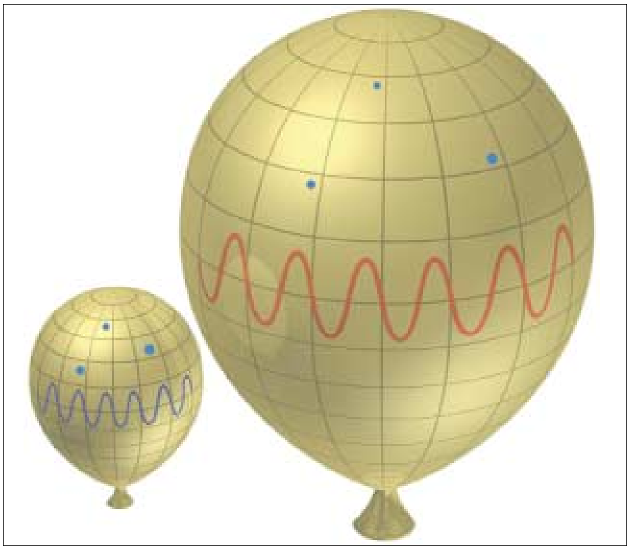


Fig. 2. As the universe expands, the dots, representing galaxies, move apart from each other and the wavelength of light increases (redshifts).

Another aspect of the big bang that is highly nonintuitive is the fact that the big-bang explosion had no center! It happened every place at the same time. Thus, while galaxies today are receding from each other, they are not moving away from some central point (see Fig. 3). Indeed, each point in space can be considered the center. The standard analogy to use in thinking about this idea is to imagine blowing up a balloon covered with small dots. As the balloon increases in size, the dots move apart from one another, but on the surface of the balloon there is no center. This two-dimensional analogy is useful when thinking about the real universe, which has three spatial dimensions (that we know of) and one dimension of time.

Glossary, References, and Further Information

A glossary of terms in this article and on the accompanying chart, along with references, is available at the website below. During summer 2003, the website will be developed with many details about the history and fate of the universe for a high school audience. It will follow the style of the Particle Adventure (http://ParticleAdventure.org).

http://UniverseAdventure.org

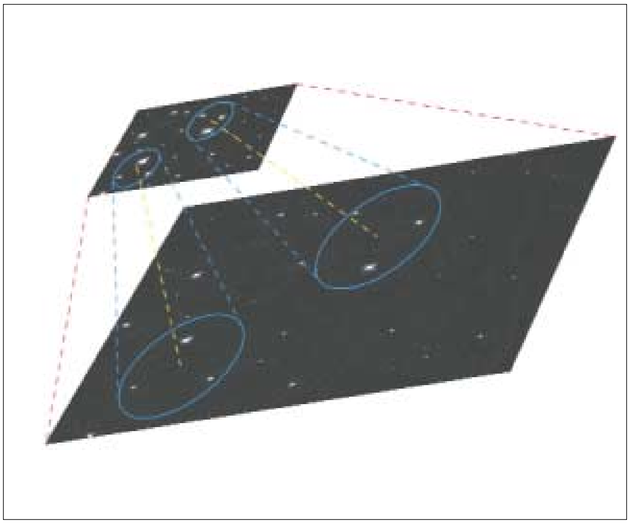


Fig. 3. Regardless of where you are located, the universe seems to be expanding away from you.

As we try to work back to understand the earliest moments after the big bang, we can attempt to recreate the conditions associated with matter and energy at early times. We do this with high-energy particle accelerators, in which elementary particles are accelerated to very high speeds, reaching energies comparable to those of the particles in the high-temperature (greater than 10^{11} K) expanding gas of the early universe. With our present accelerators, we have been able to directly probe the laws of physics associated with a time approximately 10^{-12} s after the big bang!

To attempt to probe earlier times in the history of the universe, we must at this time rely on theoretical calculations based on elementary particle physics. Since some of these calculations are based on ideas that have not yet been directly tested, we cannot consider them as firmly grounded as those appropriate to the universe after it achieved an age of 10^{-12} s.

Einstein's general theory of relativity gives the equations that govern the expansion of the universe. These relate the expansion rate to the energy density of the universe and to the curvature of space-time. The universe can exist in one of three different geometries: open, closed, or flat. If the only forms of energy in the universe are matter and radiation, then an open universe will expand forever, while a closed universe will eventually recollapse. A flat universe is the boundary between these two cases.

Inflation

One of the facets of the universe that has perplexed cosmologists ever since the discovery in 1929 that the universe is expanding is the fact that on its largest scales the universe is so uniform. As I shall describe, the distribution of both matter and radiation on scales larger than about 100 million light-years is remarkably uniform. What makes this perplexing is the fact that because no information can be transported from place to place faster than the speed of light, there are regions of the universe that are only now beginning to come into contact with other regions. Yet when we look at such disparate regions, they seem to have virtually identical properties, of density and temperature. How could this come about if they were never in "causal communication" (i.e., if their relative distances always exceed the distance light could travel since the big bang) or thermal contact?

In 1980, a resolution of this puzzle was proposed that has now become widely accepted because it resolves not only this puzzle, but also another fundamental problem in cosmology called the flatness problem (see section "The Present Time"). Moreover, it is based on ideas in particle physics that seem to be an essential part of our understanding of fundamental interactions. This proposal, called *inflation*, suggests that in the very early moments after the big bang, a very large amount of energy became stored in otherwise empty space as the universe underwent a phase transition (perhaps associated with the unification of forces in nature) from one configuration to another (much as energy gets stored in systems on Earth that undergo phase transitions, such as when ice melts into water).

The energy associated with the phase transition could cause the expansion of the universe to speed up. Very early on, if such an acceleration occurred for some short time interval (as short as 10^{-30} s), the universe could have increased in size by a factor of 10^{27} or so. In this case, regions that are today separated by large distances were once much closer together than they would otherwise have been. Indeed, before inflation happened they would have been close enough to be in thermal contact, and thus their temperature and density could have come into equilibrium. Moreover, during inflation any initial density differences on small patches would have been stretched to extremely large physical scales, so that on observable scales the universe would be essentially uniform.

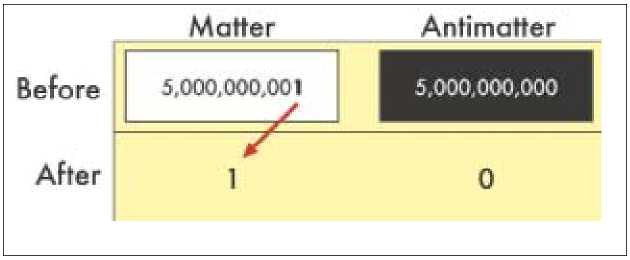
Finally, it turns out that because the physics that governs inflation involves very small-length scales, the laws of quantum mechanics must be taken into account. Remarkably, one finds that fluctuations induced by the random nature of quantum processes during inflation would have caused very small fluctuations in the density of matter and radiation to exist on large scales after inflation. As we shall see, these fluctuations could have grown, due to self-gravitation as the universe expanded, to form the large-scale structures that we see today.

For these reasons inflation is a very good candidate for describing what the early universe may have been like. It must be stressed, however, that at the present time we have no direct evidence that unambiguously tests or confirms any specific inflationary model.

The Radiation Era

As the universe expands, the densities of both matter and radiation decrease as the volume increases. Radiation is defined as material with velocities at or close to the speed of light, i.e., whose rest mass is much less than the thermal energy available. Today this includes photons and perhaps neutrinos. Radiation, which exerts a pressure, does work during the expansion, and for this reason the energy density of radiation decreases not as R^{-3} but rather as R^{-4}. This implies that the ratio of the energy density of radiation to that of matter decreases linearly with the size of the universe. Thus, as we work our way backward, even if the energy density of matter exceeds that of radiation today, at some sufficiently early time, radiation would have dominated. At the present time, the energy density of radiation is about 10^{-4} that of matter. Therefore, up until the universe was about 10^{-4} times its present size, it would have been dominated by a radiation gas.

Today, there is about one baryon (i.e., proton or neutron) for every 10^{9} or so photons in the universe. This 10^{9}-to-1 ratio was very puzzling as cosmologists began to ponder very early times. Also, because the laws of quantum mechanics combined with those of relativity imply that every particle can have an antiparticle of equal mass, the question arose as to why the universe seemed to contain only matter, and not equal amounts of matter and antimatter. After all, as long as the available thermal energy exceeded the masses of the elementary particle species, the interconversion of mass and energy would produce equal abundances of matter and antimatter in the radiation gas.



	Matter	Antimatter
Before	5,000,000,001	5,000,000,000
After	1	0

Fig. 4. During the first microsecond of the universe, there were both particles and antiparticles in abundance, being created and destroyed at a rapid rate. Tiny, subtle effects led to about one more particle being created out of each 5 × 10^9. Matter and antimatter met and annihilated, leaving only one out of each 5 × 10^9 particles.

The ratio of baryons to photons today implies that at some stage early on in the history of the universe there were about equal numbers of baryons and photons, but not quite an equal number of particles and antiparticles. As the universe cooled, particles and antiparticles annihilated into pure radiation, but because of the initial asymmetry not all of the particles could annihilate with antiparticles (see Fig. 4). If one extra baryon existed for every 10^9or so pairs of baryons and antibaryons, then ultimately we would end up with the situation we observe today. Were it not for that asymmetry, we would live in a universe without matter — with no galaxies, stars, planets, or people — consisting essentially of pure radiation.

We do not currently know exactly why this strange initial condition, with such a small asymmetry between matter and antimatter, arose in the very early universe. It could be that it was fixed by the same physics that set up the initial big-bang explosion. However, over the past 30 years, it has been recognized that natural physical processes, which can be understood in terms of the symmetries of elementary particle physics, could in principle produce such an asymmetry very early on, even if the initial universe were completely matter-antimatter symmetric. These processes would have occurred when the universe was perhaps 10^{-30} s old. While we do not yet know the

The Formation of Matter

Nucleons Form: 10^{-4} seconds — The very early universe was a plasma or "soup" of fundamental particles. As the plasma cooled, the "strong" force bound the quarks into nucleons (protons and neutrons) and into mesons.

Nuclei Form: 10^{2} seconds — In the first few minutes, protons and neutrons fused to form helium and other low-mass nuclei (^{2}H, ^{3}He, ^{7}Li). The relative abundances of these elements measured today agree precisely with abundances that can be calculated from the CMB (cosmic microwave background) temperature and the expansion of the universe.

Atoms Form: 3×10^{5} years — The CMB radiation finally cooled enough so that nuclei (charged ions) could hold onto electrons and form neutral atoms. Photons and neutral atoms then evolved independently. Under gravity, matter gradually clustered together, while photons traveled freely. With precision experiments we can observe this light from when atoms formed 14 Gyr ago.

First Stars and Galaxies Form: 3×10^{8} years — Small differences in the density of matter were amplified as denser regions gravitationally attracted even more matter. Over time, larger and larger structures grew, from galaxies to clusters of galaxies to superclusters. These began as small differences in the density of matter, but gravitational attraction made more and more matter clump together. Scientists using telescopes and computer simulations can trace this formation and evolution.

precise mechanism by which the observed asymmetry came about, some possibilities can be tested at accelerators.

At a time of approximately one-millionth of a second after the big bang, the radiation gas consisted — in addition to photons — of protons and neutrons, electrons, muons, neutrinos, and their antiparticles. Before this time the temperature was sufficiently high so that individual baryons had not yet condensed from their constituent quarks. As the temperature of the universe continued to fall below the equivalent mass of each particle species, particles and antiparticles in this species annihilated into lower-mass particles that continued to populate the radiation gas. During this period, the temperature of the radiation gas in the universe fell roughly as the inverse square root of time, i.e., approximately as $(1\ \mathrm{s}/t)^{1/2}\ (10^{9}\ \mathrm{K})$.

When the universe cooled to a temperature of about 10^{10} K (i.e., $t \approx 1$ s), the temperature of the radiation gas was sufficiently low so that the average available energy was lower than that required to convert protons into the slightly more massive neutrons. At this point, free neutrons could decay (with a lifetime of about 10 minutes), but their number density would no longer be replenished. Had this process continued, there would have been essentially no neutrons left in the universe after several hours. However, at this time, the weak interactions between protons and neutrons were such that neutrons and protons could begin to bind together to form deuterium, and from there nuclear reactions could produce helium and even lithium. We can use nuclear reaction rates as measured in the laboratory to estimate the efficiency of these processes in the early universe.

Matter made of protons and neutrons is called *baryonic* matter or ordinary matter. It is one of the most remarkable and robust predictions of the big-bang picture that roughly 25% of the baryonic mass in the universe should end up as helium and 75% as hydrogen, while merely one part in 10^{10} of the nuclei in the universe after it was about five minutes old or so should have been lithium. Essentially nothing higher on the periodic table than lithium would have been created in the big bang. And when we attempt to determine the primordial abundance of light elements on the basis of astrophysical observations today, we find about 25% as much helium as hydrogen by mass, and one part in 10^{10} lithium by number!

We can use the agreement between theory and observation, in fact, to constrain the abundance of protons and neutrons in the universe. However, we find that only about 5% of the total density required to result in a flat universe exists today in the form of baryonic matter. This includes all galaxies, stars, and dust we observe today, plus probably at least as much baryonic matter that is not in a luminous form.

The Matter Era

Following the synthesis of light nuclei in the first few minutes of the big bang, the universe continued to cool until it achieved an age of roughly 10,000 years. At this time the density of radiation and the density of matter became roughly equal. Following this, up until close to the present time, matter dominated the expansion of the universe.

As far as the growth of structure in the universe is concerned, this time is pivotal. Where radiation, with its high pressure, dominated the expansion, small density fluctuations in matter could not collapse inward due to gravity, because the outward pressure of radiation combating the collapse would be too great. In fact, such density fluctuations in ordinary (or baryonic) matter (which was coupled to radiation at that time due to electromagnetic interactions) would actually oscillate, as pressure and gravity opposed each other. During such oscillations, energy could be dissipated, causing the fluctuations themselves to eventually disappear. Small density fluctuations in any material not coupled to the surrounding radiation bath would not have dissipated, but in a radiation-dominated universe they would not grow either.

As a result, as I later describe, without some material that did not interact with electromagnetic radiation, the initial seeds of structure formation would have dissipated before structure could form. This is one of the many reasons we believe that "dark matter," which appears to dominate the density of matter on galactic scales and larger, is indeed composed of some new type of elementary particle that does not couple to electromagnetic radiation. This idea is further reinforced by the fact that the inferred net density of dark matter in the universe exceeds, by a factor of five to six, the maximum density of baryonic matter inferred from calculations of the light element abundance, as described earlier.

The Cosmic Microwave Background

Nature has provided us a unique window through which we can see a "snapshot" of the universe shortly after matter began to dominate the expansion. Because of the finite speed of light, as we look farther and farther out into the universe, the objects we observe are seen as they looked earlier and earlier in time. In principle, if we looked out far enough, we could actually see the big bang itself. However, we cannot do this, at least with electromagnetic radiation, because for all times earlier than roughly 300,000 years after the big bang, the universe was opaque to radiation; that is, the radiation was thoroughly scattered. The dominant material in the universe, hydrogen, can exist in neutral atomic form at temperatures lower than 3000 K. But prior to 3×10^5 yr after the big bang, the temperature exceeded this value and so the radiation gas was sufficiently energetic to ionize hydrogen. Ionized gas (plasma) strongly scatters electromagnetic radiation.

Thus, if we attempt to look back arbitrarily far, we encounter a "wall" representing a time when the universe was opaque to radiation (see Fig. 5). However, we will see radiation emitted from the "last-scattering surface"— the location that represents, relative to us, the time when the temperature dropped below 3000 K and matter became neutral. The radiation emitted at that time could travel through the universe

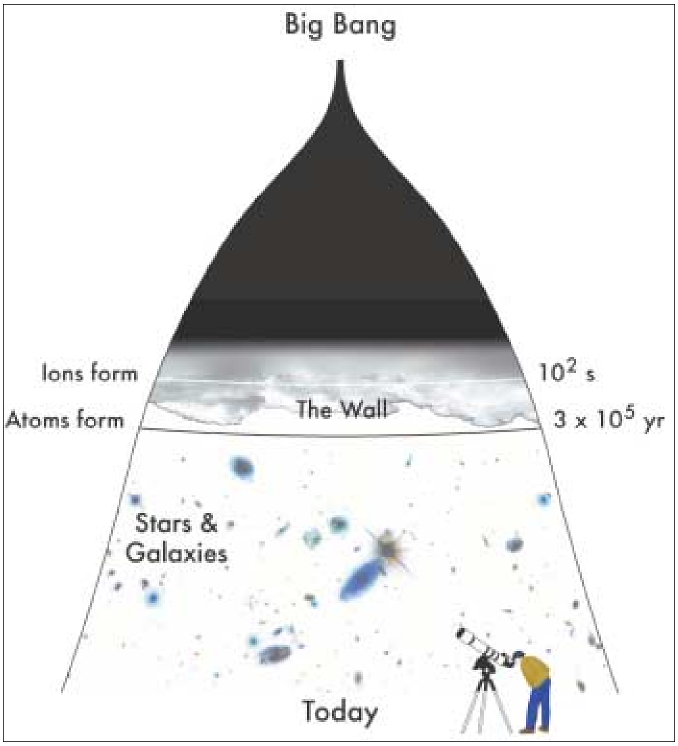


Fig. 5. Light interacts with charged ions and electrons but is unaffected by neutral atoms. Before the universe became neutral at 3 × 10^5 yr, charged particles scattered all light. When light scatters it changes direction, and with more scattering it loses its way. This is just what happens in clouds, where small water droplets diffuse the light so that you cannot see light from objects past the "wall" of the cloud.

unimpeded, reaching our telescopes without significantly interacting with the intervening matter. This radiation would have continued cooling as the universe expanded by a factor of about 1000, up until the present. It would reach us as a uniform radiation bath, coming at us in all directions, with a temperature today of about 3 K, so that the dominant wavelengths received would be in the microwave band. For this reason the radiation is called the Cosmic Microwave Background (CMB).

In 1965 the first detection of this background radiation was made. In 1989, the cosmic background explorer (COBE) satellite was launched to "image" this background. This satellite was able to probe the temperature of radiation coming from all directions, with a resolution in temperature of almost one part in 100,000 and an angular resolution on the sky of about 10°. It produced an "all sky" map of the last- scattering surface and demonstrated the existence of small temperature (and hence density) fluctuations on that surface that were the precursor "seeds" for large-scale structure formation (see Fig. 6).

The relative size of these temperature fluctuations, about one part in 100,000, was so small that if normal matter accounted for all mass in the universe, gravity would not have had sufficient time after the CMB era to cause the fluctuations to grow and then to condense to form the structures we see today. This is because normal matter interacts strongly with radiation until the CMB last-scattering surface, and thus density fluctuations in matter before that time could not have grown on small (i.e., galaxy-size) scales. Only if a "dark" component of matter that did not interact with electromagnetic radiation had existed could small-scale density fluctuations have been preserved and grown sufficiently so that they might ultimately form galaxies and the like. As I have stressed earlier, this is just one of many different arguments that confirm the need for some type of nonbaryonic dark matter in the universe.

A decade later, several ground-based microwave experiments were able to probe the CMB on a much finer angular scale. This allowed another remarkable discovery to be made. Since the CMB last-scattering surface existed when the universe was approximately 300,000 years old, 300,000 light-years represents a special length scale on that surface. This is the scale over which different regions of the last-scattering surface could have been in causal contact with each other. Regions separated by larger distances could not have communicated, and no physical process could have responded to density fluctuations on scales larger than this.

For this reason, this length scale should be imprinted on the temperature fluctuations we see in the CMB. Today that length scale — 300,000 light-years — corresponds to an angular size of about one degree. However, the exact angular scale would depend upon the geometry of the universe in which we live. In a flat universe, light travels in straight lines. However, in a closed universe, light rays converge as one looks back in time, and in an open universe they diverge. As a result, in a closed universe the angular scale associated with a fixed distance across the last-scattering surface would be larger, while in an open universe it would be smaller than it would be in a flat universe.

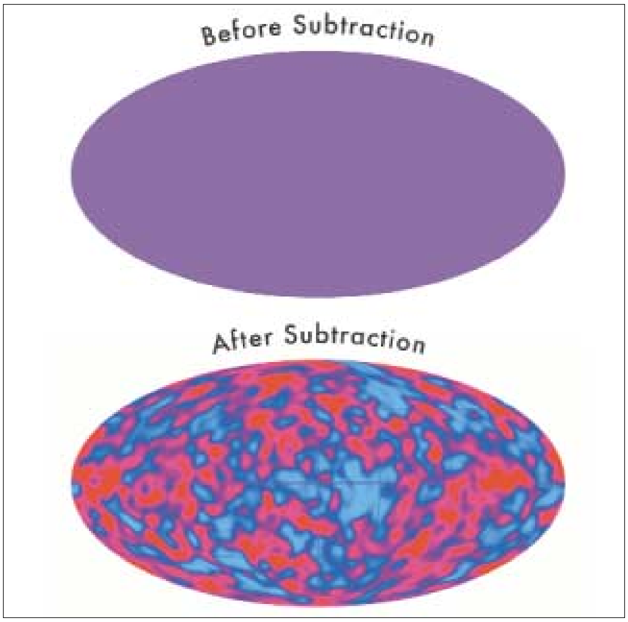


Fig. 6. Images of the universe as illuminated by the CMB from the time that atoms first formed. Both show an image of the entire spherical sky mapped into an oval. The top image shows the remarkable uniformity of the whole sky and early universe. The lower image has the uniform background and emission from our galaxy removed, and shows the tiny variations at the level of one part in 100,000.

In 1999, scientists who conducted several CMB detection experiments announced that they had definitively detected fluctuations in the CMB on the requisite angular scales (which correspond to scales that today contain super clusters of galaxies, the largest gravitationally bound objects in the universe). From their data, we now have very good evidence that to high accuracy, our universe is spatially flat on its largest scales.

The Present Time

Between the last-scattering surface and today, the universe continued to cool and matter continued to cluster gravitationally, ultimately forming stars, galaxies, and clusters of galaxies. It took almost 10^9 years for the first stars and galaxies to form, and with our largest telescopes we are able to observe those early galaxies.

Observations of the clustering of matter on large scales allow us to estimate the total density of matter in the universe today. We consistently find, on scales ranging from galaxies to clusters of galaxies, that almost 10 times as much mass in the universe is nonlu-

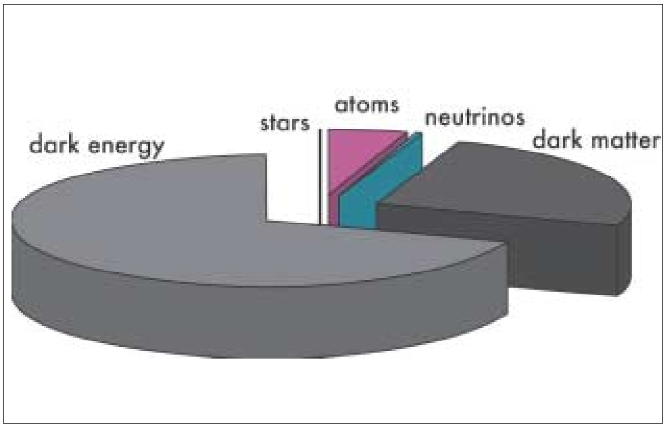


Fig. 7. The stars we see in the night sky, even with the most powerful telescopes, are only a very small part of the universe. More is in atoms contained in gas and dust, as well as in neutrinos and the still unidentified dark matter. But almost two-thirds of all energy is the mysterious dark energy, whose nature is one of the great questions facing cosmology.

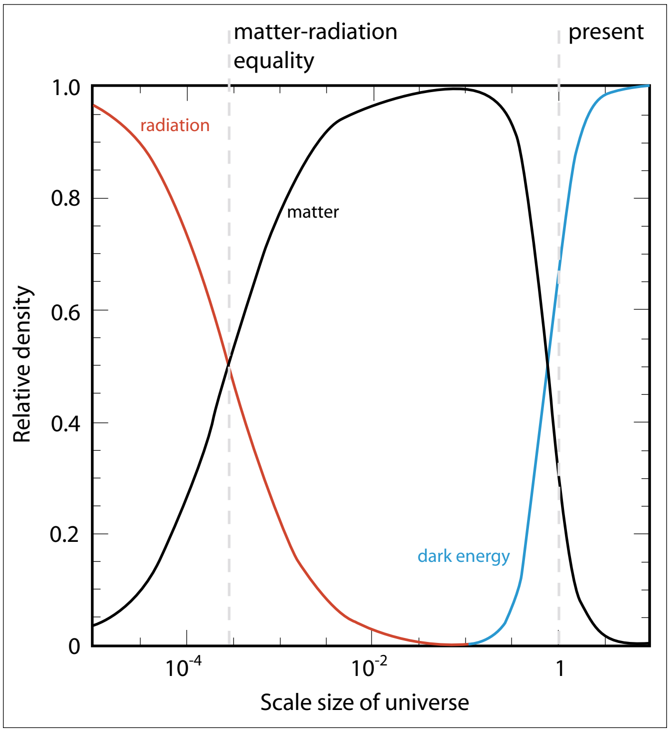


Fig. 8. The relative energy density of various components of the universe as a function of the scale factor of the universe (and thus of time). The sum of densities is set to unity. The curve labeled "dark energy" is representative and is computed for a cosmological constant corresponding to a flat universe.

minous as is luminous (see Figs. 7 and 8). Moreover, as described earlier, by comparing this total density to the density of baryons required for the predictions of the formation of light elements after the big bang to agree with observations, we find that most of this dark matter must be nonbaryonic. Experiments are under way at various locations around the world that are sensitive to different candidates for this dark matter.

The estimates of the total density of matter in the universe today now definitively establish the fact that there is only about 30% of the density required to result in a flat universe. When this result is combined with the recent CMB observations that imply a flat universe, this suggests that up to 70% of the energy of the universe may not reside in the form of matter at all, but perhaps in some exotic form of "dark energy" that permeates the space between galaxies and clusters.

This remarkable conclusion was in fact anticipated by the earlier surprising discovery (in 1998) that the expansion of the universe is in fact accelerating, as was predicted to occur by Einstein's equations if empty space carries a nonzero energy density. This evidence is based on observations of distant supernovae — exploding stars that shine as brightly as an entire galaxy for a period of weeks (see Fig. 9). By comparing the brightness of distant to nearby supernovae, observers measure both the distance to these distant objects and their recession velocities. They have confirmed both that the Hubble expansion appears to be accelerating once again and that this acceleration is consistent with approximately 70% of the total energy density of the universe residing in empty space.

One way to further confirm this strange result is to attempt to determine the age of the oldest stars in our galaxy. For a fixed measured expansion rate today, a universe dominated by dark energy should be older than a universe dominated by matter. Recent observations of old globular clusters — compact groups of up to a million stars — located in the halo of the Milky Way galaxy confirm that the age of our galaxy is at least 10.5 Gyr, suggesting that the age of the universe is at least 11 Gyr, with a best-fit estimate of 13.5 Gyr. The maximum age for a flat matter-dominated universe with the measured Hubble expansion rate is about 10^{10} years. This result thus provides further support for the existence of dark energy.

We currently have no theoretical understanding of the nature and origin of this dark energy that seems to dominate the universe. Unraveling these issues is perhaps the most significant outstanding problem in physics and cosmology.

Geometry and Destiny

Once we allow for the existence of dark energy, our notions about the future evolution of the universe completely change. For example, it is no longer true that a closed universe must recollapse or an open universe must expand forever. Indeed, until we determine the nature of this dark energy (i.e., until we determine whether it represents a fixed energy associated with empty space that will never vary in the future, or whether it corresponds to some changeable energy density stored in some slowly varying cosmic field), we cannot say for certain what the ultimate fate of the universe will be. However, all signs currently correspond to an expansion of the universe that continues forever within a universe that may be infinite in spatial extent.

Current ideas in particle physics also open up the possibility that there are many causally disconnected universes within a grand "multiverse," or perhaps even that there are large extra dimensions in which our four-dimensional universe is embedded. All of these ideas are quite speculative at the present time, but a number of them are actually amenable to testing in the laboratory. Given the exciting surprises of the last decade in observational cosmology, it is clear that the universe comes up with possibilities that often exceed the imagination of astronomers.

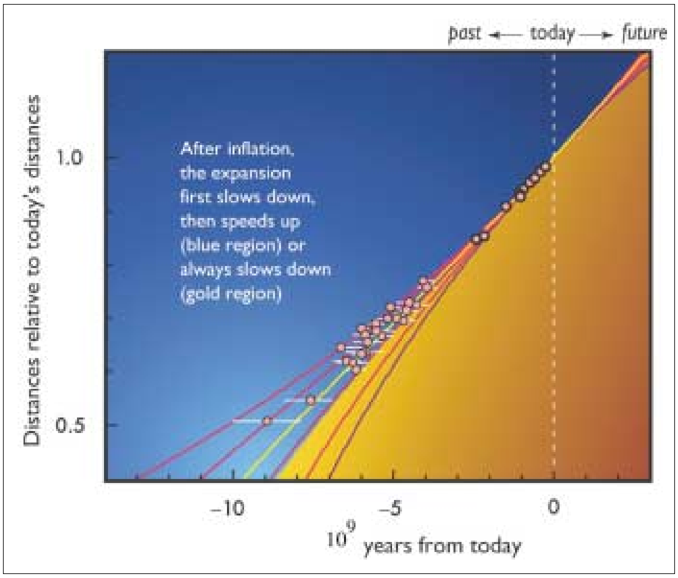


Fig. 9. Using supernovae to measure distances, scientists discovered that the data lie along a curve indicating an accelerating universe. The figure is courtesy of Saul Perlmutter of Lawrence Berkeley National Laboratory.

PACS codes: 98.01A, 95.30

Lawrence M. Krauss *is Ambrose Swasey Professor of Physics, professor of astronomy, and chair of the physics department at Case Western Reserve University. His newest book,* Atom: A Single Oxygen Atom's Odyssey From the Big Bang to Life on Earth and Beyond, *presents the cosmic history of the universe as witnessed by a single oxygen atom that is located in a glass of water today. For this book, he was awarded the American Institute of Physics Science Writing Award for 2002.*

Department of Physics and Astronomy, Case Western Reserve University, 10900 Euclid Ave., Cleveland, OH 44106-7079