Big Bang Cosmology and the Microwave Background

Salman Habib and Raymond J. Laflamme

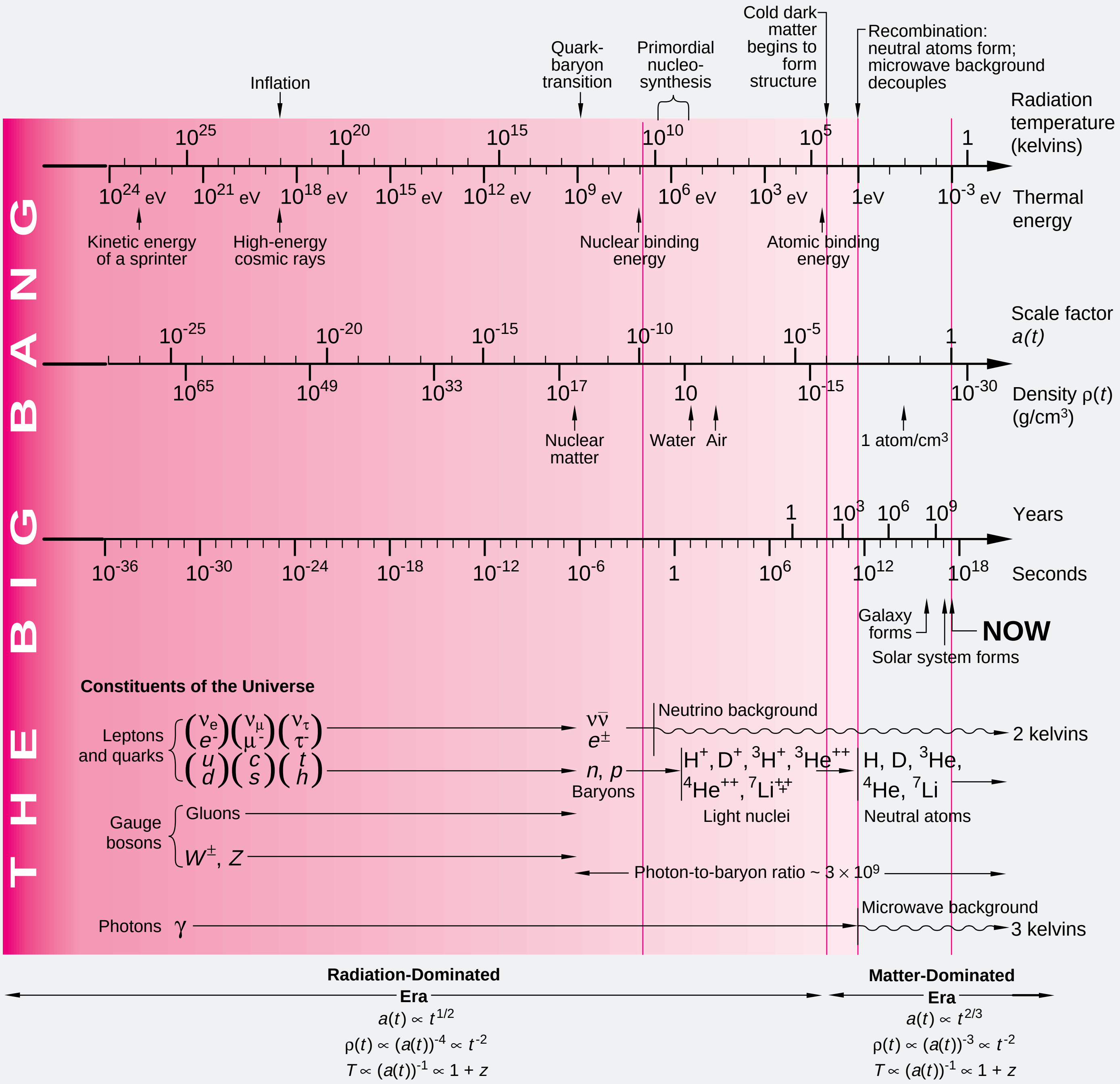
Big Bang cosmology implies a certain predictable course for the thermal history of the universe. Here we review that history, emphasizing the two classic pieces of evidence that support it: the relative abundances of the light elements produced by primordial nucleosynthesis and the existence, spectrum, and fantastic isotropy of the cosmic microwave background. We will also discuss the modifications that might be caused by the presence of cold dark matter.

Figure 1 outlines the calendar of events, or major epochs, in the history. It can be divided into two main periods: the radiation-dominated period, which lasted for approximately ten thousand years (10^{11} seconds), and the matter-dominated period from ten thousand years after the Big Bang to the present. During the first fraction of a second after the Big Bang, the constituents of the universe undoubtedly included the particles of the standard model of particle physics: leptons, quarks, and the gauge bosons that mediate interactions among them, all moving at velocities so close to the velocity of light that from the point of view of thermodynamics, they behaved as particles of radiation (that is, massless particles). The physics during this time is highly speculative. It is assumed that thermal energies were initially at the Planck scale (10^{28} eV), where the exotic and poorly understood phenomenon of quantum gravity could have played a crucial role. Many particle physicists and astrophysicists believe that a bit later, as the temperature cooled to 10^{26} kelvins, there was a brief period of inflation. During the inflationary phase, the size of the universe grew exponentially by a factor of at least 10^{28}. Less speculative is the prediction that at a temperature around 10^{12} kelvins, a quark-hadron transition occurred, when all the quarks and gluons combined to form protons and neutrons. The particles of cold dark matter, if they exist, would also have been created sometime after the inflationary phase.

The radiation-dominated era. From about 10^{-2} second to the present, standard Big Bang cosmology presents an almost universally accepted picture, which is borne out by the COBE data and other recent observations. By that time the universe consisted of particles that were stable (or long-lived com-

Figure 1. The Thermal History of the Universe

The figure (adapted courtesy of Mike Turner) shows important epochs in the history of the universe according to standard Big Bang cosmology. It also shows the particles of matter and radiation constituting the universe during each epoch. The history is laid out on logarithmic scales to emphasize the early part. The scales are temperature, thermal energy kT, universal scale factor $a(t)$ (note that $a(t) = 1/(z + 1)$ where z is the redshift), density $\rho(t)$, and time. Temperature, thermal energy, redshift, and energy density (or the equivalent mass density) all decrease with time, whereas the scale of the universe $a(t) \equiv R(t)/R(t_0)$ increases. The temperature plotted is the temperature of the radiation, which until recombination is the temperature of the baryonic matter (neutrons and protons) as well. The formulae relating these variables are shown beneath the figure. They follow from conservation of energy and from the assumptions that matter and radiation were initially in thermal equilibrium and that the universe has expanded adiabatically since the hot Big Bang. The relationship between $\rho(t)$ and $a(t)$ changes depending on whether the universe is radiation-dominated or matter-dominated, that is, on whether most of the energy in the universe is in the form of radiation (photons, neutrinos, and other particles moving at speeds near c) or most of the energy is in the form of matter. The temperature and scale-factor axes are lined up with respect to each other by the observation that at present ($t = t_0$) the temperature of the cosmic background radiation is approximately 3 kelvins and that by definition $a(t_0) \equiv 1$. The time and density axes are lined up with the size axis by using a specific value for the Hubble constant and the assumption that the dimensionless density parameter Ω is equal to 1. Then one can place any event on all three axes if one knows its position on



one axis (usually temperature, sometimes redshift). Note that the scale-factor axis does not show the expansion due to inflation. In the lower portion of the figure, the line for neutrinos (and antineutrinos) turns into a wiggly line to indicate that at that time the neutrinos decoupled from matter and photons and subsequently expanded freely. The decoupling occurred when the universe became cool enough and dilute enough that neutrinos no longer interacted significantly with matter, at a temperature around 10^{10} kelvins. Similarly, at a temperature of about 3000 kelvins, the line for photons turns into a wiggly line to indicate their decoupling from matter due to recombination, the combining of the electrons with the positively charged nuclei to form neutral atoms. The subsequent free expansion of photons produced the cosmic microwave background radiation now observed.

pared to the age of the universe then): electrons, neutrinos, protons and neutrons, and photons, as well as the yet-unidentified constituents of dark matter. Unstable particles such as mu and tau leptons, mesons, and the gauge bosons of the weak force had disappeared through decay to stable particles. Figure 1 outlines the changes through time of the constituents of matter.

At about $t = 10^{-2}$ second, the universe was a "primeval fireball" consisting of baryonic matter and radiation in thermal equilibrium (and perhaps dark matter). Moreover, most of the energy was in the form of radiation. That follows from simply extrapolating the present density of matter and radiation back through the presumed adiabatic expansion of the universe. The energy density therefore obeyed the formula for black-body radiation, that is, for radiation in thermal equilibrium with a totally absorbing body. Thus the energy density u of each relativistic species present—whether photons, massless neutrinos, or even electrons and positrons moving at nearly the speed of light—obeyed the Stefan-Boltzmann law:

$$u_{\text{radiation}} = \bar{\rho}_{\text{radiation}}\, c^2 = \sigma\, T^4,$$

where T is temperature and σ is the analogue of the Stefan-Boltzmann constant.

Black-body radiation (even with a small admixture of equilibrated matter) behaves as a dissipationless fluid, so the radiation-dominated universe expanded adiabatically, and the total entropy was conserved. As mentioned in the main article in the discussion of the redshift, the adiabatic expansion "stretched" the wavelengths of photons and all other massless (or effectively massless) species in proportion to the universal scale factor $a(t)$, or $\lambda(t) = \lambda_0 a(t)$. Since all wavelengths were stretched proportionally, the radiation continued to have a black-body spectrum, or Planck distribution, a fact that explains why the black-body spectrum of the cosmic background radiation is not a surprise. However, the temperature of the radiation, which is proportional to the average energy per quantum or inversely proportional to the average wavelength, decreased inversely with the scale factor, $T \propto (a(t))^{-1}$, or in proportion to the redshift ($T \propto z$). The energy density of the radiation therefore decreased due to both the stretching of the wavelengths of the quanta and the dilution of their number density; thus $\bar{\rho}_{\text{radiation}}$ decreased rapidly, as $(a(t))^{-4}$.

Equation 3 in the caption of Figure 3 in the main text is solved to relate the scale factor to time. In particular, when radiation was the major contributor to the energy density, $a(t)$ increased as $t^{1/2}$, and so the temperature of the universe decreased as $t^{-1/2}$. Thus during the radiation-dominated era, every decrease by a factor of 10 in temperature corresponds to an increase by a factor of 100 in time since the Big Bang. As the primeval fireball of matter and radiation expanded and cooled in thermal equilibrium, the decrease in temperature was like a cosmic clock, ticking off stages in the physics of the universe.

Primordial synthesis of the light elements. In the first few minutes after the Big Bang, the one major event that directly left observable traces was the primordial synthesis of the light elements through nuclear fusion reactions. In particular most of the helium-4 observed today in stars was almost assuredly synthesized then.

As the temperature fell to 10^9 kelvins ($t \approx 10^2$ seconds), conditions were right for nucleosynthesis to begin. Two quantities controlled the relative abundances of light elements that resulted. The first is the fraction of the baryons that were neutrons, a fraction that decreased with time as neutrons were converted to protons through the weak interactions. Calculations predict that the number of neutrons at the time of primordial nucleosynthesis is approximately 12 percent of the total number of baryons and that almost all those neutrons ended up in helium-4 nuclei, the most tightly bound of the light-element nuclei. In other words, Big Bang cosmology predicts that the cosmic abundance by weight, or mass fraction, of helium-4 nuclei relative to hydrogen nuclei (protons) is about 24 percent, in agreement with observation.

The second determinant of relative abundances produced by primordial nucleosynthesis is the ratio of the number of photons to the number of baryons (or equivalently the photon entropy per baryon). That ratio does not change with time and therefore, as we will see below, places a constraint on the baryon density today. It also controlled the rates of various competing nuclear reactions at the time of nucleosynthesis. For example, nucleosynthesis could not really begin until the density of energetic photons became low enough that deuterons (deuterium nuclei) formed by the fusion of neutrons and protons (neutron + proton → deuteron + photon) were unlikely to be blasted apart by the reverse reaction. That condition depends on the temperature, which must be about 10^9 kelvins, but also depends strongly on the ratio of photons to baryons. Once the deuteron abundance increased, nuclear reactions involving one or two deuterons could build on one another to form primordial abundances of all isotopes of hydrogen and helium: deuterium, tritium, helium-3, and helium-4 as well as very small amounts of lithium-7.

Primordial nucleosynthesis manufactured all those nuclei in abundances consistent with current data provided the photon-to-baryon ratio was between 2.5×10^9 and 3.6×10^9. The agreement between nucleosynthesis calculations and observations is considered convincing evidence that in the beginning the universe was hot, radiation-dominated, and in thermal equilibrium—that the hot Big Bang did indeed occur!

Further, since the ratio of photons to baryons during primordial nucleosynthesis was preserved to the present, it can be combined with the known number density of photons in the cosmic microwave background to predict the baryon density in the universe today. The bounds on the ratio yield bounds on the baryon density of $(2.5 \pm 0.5) \times 10^{-31}$ gram/centimeter^3, and bounds on the baryonic contribution to Ω of $0.01\ h^{-2} < \Omega_{\text{baryon}} < 0.015\ h^{-2}$. If we assume that $h = 1/2$, then $\Omega_{\text{baryon}} \approx 0.05$, which is much larger than Ω_{luminous}. This prediction for Ω_{baryon} places a constraint on all models of structure formation.

The transition to a matter-dominated universe. Following nucleosynthesis, the cosmic soup consisted of photons in thermal equilibrium with hydrogen and helium nuclei and a number of electrons, left over from electron-positron annihilation, sufficient to balance the charge of the baryonic matter. The radiation-dominated universe continued to expand adiabatically, and entropy conservation continued to ensure that the temperature decreased inversely with the size of the universe, $T \propto (a(t))^{-1}$. Thus $\bar{\rho}_{\text{radiation}}$, the equivalent mass density of the radiation, continued to decrease as $(a(t))^{-4}$, whereas $\bar{\rho}_{\text{matter}}$, which was dominated by the matter's rest mass rather than its kinetic energy, decreased inversely as the volume, or as $(a(t))^{-3}$.

The matter-dominated era. Since the energy density of the radiation decreased more rapidly than that of the matter, at some point the energy density of the matter became larger, and from then on the universe has been matter-dominated. The temperature (or time, t_{eq}) at which the switch from a radiation-dominated to a matter-dominated universe occurred depends on the total matter density in the universe—baryons plus any species of dark matter that might be out there. If baryons make up all the matter and the total matter density today is 2.5×10^{-31} gram/centimeter^3 (that is, $\Omega_{\text{baryon}} = 0.05$), then the switchover occurred at a temperature of about 1000 kelvins ($t_{\text{eq}} \approx 10^6$ years), whereas if $\Omega = 1$ and baryons make up only 5 percent of the matter, as assumed by the standard CDM model, then the universe became matter-dominated when the cosmic scale was a factor of 20 smaller or the temperature was a factor of 20 higher, at about 20,000 kelvins ($t_{\text{eq}} \approx 10^4$ years). Figure 1 was constructed with the assumption that $\Omega = 1$.

The beginning of the matter-dominated era marks the time when the mutual gravitational attraction of the matter became stronger than the gravitational pull of the background sea of radiation. Consequently matter could, in principle, begin to clump together under the influence of gravity. However, the baryons (primarily in the form of hydrogen and helium nuclei) and the electrons are electrically charged and therefore remained coupled to the photons through electromagnetic interactions, primarily the scattering of electrons and photons (Thomson scattering). Thus, their clumping was impeded by radiation pressure that was much stronger than gravitational forces in the nearly uniform mass distribution—the photon-to-baryon ratio is at least a billion to one—so density inhomogeneities could not grow in baryonic matter.

In contrast, non-baryonic dark matter, which by definition does not participate in electromagnetic interactions, began to develop density inhomogeneities when the universe became matter-dominated, provided the dark-matter temperature, or average thermal energy, was low enough and the masses of the dark matter particles were high enough that the particles were moving at nonrelativistic speeds. (Otherwise, small-scale fluctuations would have been washed out by the free streaming of dark-matter particles.) That is, the dark matter must have been "cold" for structure to have begun to form on all scales at t_{eq}, when the energy in matter was equal to the energy in radiation.

Recombination. When the universe cooled to about 3000 kelvins, the situation changed dramatically. That temperature is far enough below the ionization temperature of helium and hydrogen atoms that almost all the electrons combine with the helium and hydrogen nuclei to form neutral atoms. Cosmologists call that event "recombination" (a somewhat misleading term since the electrons and nuclei were combining for the first time). After recombination, charged particles were rare, and so the scattering of photons by matter was also rare. Thus the time of recombination, t_{recomb}, marks the end of the tight coupling between baryonic matter and photons; thermal equilibrium between matter and radiation is no longer maintained, and the universe is said to be transparent to radiation since photons can travel through the universe with little chance of being scattered.

Once the radiation pressure ceased to have an effect on baryonic matter, that matter too could begin to feel the pull of gravity. In fact, if dark matter were not present, recombination would have to mark the beginning of the growth of density fluctuations due to gravitational forces. However, as stated in the introduction to the main article, recombination apparently occurred too late (and the density of baryonic matter was too low) for the growth of gravitational instabilities to explain the observed large-scale structure. Thus the majority of cosmologists believe dark matter must have led that development.

The discovery of the cosmic background radiation. According to Big Bang cosmology, the radiation that existed at the time of recombination was thermal and consequently had a black-body spectrum at a temperature of about 3000 kelvins; in other words, its energy density as a function of wavelength had the Planck distribution. At recombination, the universe became transparent to radiation, so the photons survived unchanged except for their redshift due to cosmic expansion. As the wavelength of each photon increased in proportion to the scale factor $a(t)$, the form of the spectrum was preserved; that is, the radiation now has a black-body spectrum but at a temperature lower by a factor $a(t_0)/a(t_{\text{recomb}})$ than the temperature at the time of recombination. That spectrum is shown in Figure 2. Indeed, Big Bang cosmology predicts that low-temperature black-body radiation should now be observable from all empty parts of the sky—this remnant of recombination forms a cosmic background.

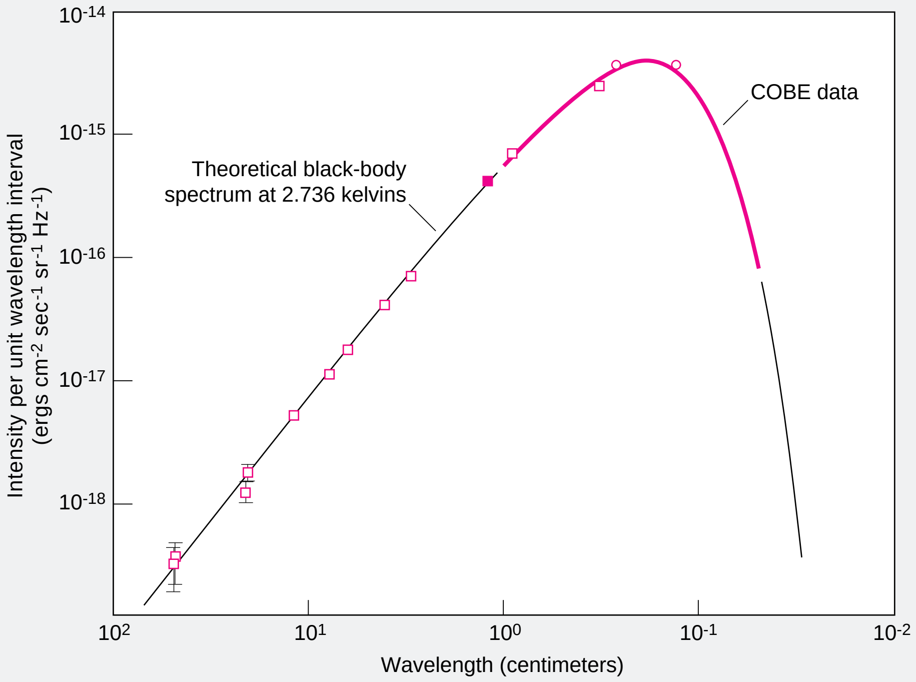


Figure 2. The Planck Distribution and the Cosmic Background Radiation

The figure shows the Planck distribution, or the energy density per unit wavelength of radiation in thermal equilibrium with a perfectly absorbing (black) body, at a temperature of 2.736 kelvins. The red data points represent measurements by instruments other than COBE of the energy density per unit wavelength for the cosmic background radiation. The precision of most of the data points is so high that error bars would be no bigger than the symbols representing the points. The red part of the curve represents the measurements made by the COBE satellite; its error bars and deviations from the Planck distribution are too small to appear. Thus the background radiation appears to have a nearly perfect "black-body" spectrum, indicating that matter and radiation were in thermal equilibrium at the time of recombination, as predicted by standard Big Bang cosmology. (Figure adapted by courtesy of D. T. Wilkinson.)

The existence of a cosmic remnant left over from the hot Big Bang was first predicted by George Gamow, R. A. Alpher, and R. C. Herman in 1948 on the basis of their analysis of the conditions needed for primordial nucleosynthesis. In the early sixties R. H. Dicke and P. J. E. Peebles took this notion seriously enough to plan experiments to detect its presence. Ironically, in 1965, while Dicke, Peebles, D. T. Wilkinson, and P. G. Roll were building their equipment, the cosmic background radiation was detected inadvertently by Arno Penzias and Robert W. Wilson at Bell Laboratories. Penzias and Wilson were investigating the internal noise properties of an antenna designed to detect radio waves emitted by our own galaxy. Quite unexpectedly they discovered extra noise at a wavelength of 7.35 centimeters whose intensity was independent of direction. The intensity also did not vary with the time of day or the time of year. If that radiation was assumed to be one point of a black-body spectrum, then the inferred temperature of the spectrum would be between 2.5 and 4.5 kelvins. Peebles, Dicke, Wilkinson, and Roll in a separate paper suggested that this highly unusual radiation was the predicted remnant of the Big Bang.

Since 1965 astronomers have measured the intensity of the cosmic background radiation at many wavelengths. In 1989 the COBE satellite was launched specifically to study various aspects of the radiation. The COBRA rocket-based experiments and one of the instruments on COBE confirmed to very high precision that the spectrum of the cosmic background is indeed a black-body spectrum. COBE made a particularly precise measurement of the temperature: 2.736 kelvins plus or minus a few millikelvins. A black-body spectrum at that temperature, shown in Figure 2, has a peak at the wavelength of 2 millimeters, which lies in the microwave region of the electromagnetic spectrum. The ratio of the present temperature of the cosmic microwave background to the photon temperature at recombination, $T(t_0)/T(t_{\rm rec})$, is about 10^{-3}, which means that the background radiation has been redshifted by a factor of about 1000 since the time of recombination.

Anisotropy of the cosmic background radiation and implications for models of large-scale structure. Information about the isotropy of the cosmic background radiation is crucial to studies of structure formation because it allows us to infer the size of density fluctuations at the recombination time, when the cosmic background radiation decoupled from baryonic matter. If the mass distribution at the time of decoupling had been totally uniform, the cosmic background radiation would be isotropic, or the same in all directions. However, in 1992 another instrument on COBE detected inhomogeneities; the temperature of the background radiation varies depending on the direction the radiation comes from. The temperature differences are quite small: $\delta T/T \approx 6 \times 10^{-6}$. Those differences are believed to reflect density inhomogeneities present at the time of recombination; photons emanating from a high-density region must have lost energy in "climbing out" of the region and thus suffered a redshift over and above the cosmic redshift. Fluctuations in the cosmic background radiation were seen at all angular separations larger than about 10 degrees of arc (the resolution of the relevant instrument on COBE).

Ten degrees of arc corresponds to hundreds of megaparsecs now and to hundreds of kiloparsecs at the time of recombination, a distance larger than the Hubble radius, $R_{\rm H} \equiv c/H(t)$, at that time. Thus the fluctuations measured by COBE had wavelengths larger than the Hubble radius (or the horizon) when they were imprinted on the cosmic background radiation. They therefore were apparently not connected causally at the time of the imprinting. In that case the near anisotropy of the background radiation is very difficult to understand. This paradox is an aspect of the "horizon problem" pointed out by Misner in the 1960s. The problem is solved by postulating a period of inflation, or exponential expansion of the very early universe. In inflationary scenarios, regions that now appear causally disconnected were in fact causally connected prior to inflation. The point of this discussion is that the fact that the scale of the observed temperature fluctuations is larger than the Hubble radius at $t_{\rm recomb}$ implies that these large-scale fluctuations could not have been affected by small-scale physics after inflation and must indeed carry information about primordial fluctuations.

Since the shape of the CDM spectrum of fluctuations at the time of recombination can be calculated from the assumptions of the CDM model, the observed amplitudes determine the normalization (amplitude) of the spectrum on all scales. Had the measured fluctuations been much smaller in amplitude than 6×10^{-6}, not only the CDM scenario, but also the idea that gravity is the decisive force responsible for structure formation, would have been in jeopardy. Not only are the COBE temperature-fluctuation measurements thus consistent with the CDM model, but as described in the section "Cold Dark Matter, Large Scales, and COBE" in the main text, they also fix the value of the only undetermined parameter of the model. □

Further Reading

E. Bertschinger. 1994. Cosmic structure formation. *Physica D*, in press.

H. P. Gush, M. Halpern, and E. H. Wishnow. 1990. Rocket measurement of the cosmic-background-radiation mm-wave spectrum. *Physical Review Letters* 65: 537–540.

J. C. Mather et al. 1994. Measurement of the cosmic microwave background spectrum by the *COBE* FIRAS instrument. *The Astrophysical Journal* 420: 439–444.

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

Michael S. Turner. 1987. A cosmologist's tour through the new particle zoo (candy shop?). In *Dark Matter in the Universe*, edited by J. Kormendy and G. R. Knapp. Reidel.

Terry P. Walker, Gary Steigman, David N. Schramm, Keith A. Olive, and Ho-Shik Kang. 1991. Primordial nucleosynthesis redux. *The Astrophysical Journal* 376: 51–69.

Steven Weinberg. 1972. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. John Wiley and Sons.

Steven Weinberg. 1977. *The First Three Minutes: A Modern View of the Origin of the Universe*. Basic Books.

Salman Habib performed his undergraduate work at the Indian Institute of Technology, Delhi, and his graduate work at the University of Maryland, College Park. He now has a joint postdoctoral fellowship in the Theoretical Astrophysics and the Elementary Particles and Field Theory groups at the Laboratory. His research interests are quantum field theory, statistical mechanics, nonlinear dynamics, cosmology, and astrophysics.

Raymond J. Laflamme received his bachelor's degree from the Université Laval in his native Canada and his Ph.D. from Cambridge University under the direction of Stephen Hawking. He is now an Oppenheimer fellow in the Theoretical Astrophysics Group, studying cosmology, the early universe, and quantum mechanics.