Introduction to Astrophysics and Cosmology

The Cosmic Microwave Background Radiation

Based on: Chapter 29 of An Introduction to Modern Astrophysics

Helga Dénes 2025 S1 Yachay Tech

hdenes@yachaytech.edu.ec

Cosmic nucleosynthesis

In 1946 George Gamow realized that the newborn, **dense universe must have been hot enough for a burst of nuclear reactions to occur,** he proposed that a sequence of reactions in the very early universe could **explain the measured cosmic abundance curve.** Gamow, together with Ralph Alpher, published this idea two years later.

Still later, however, detailed calculations by Alpher and Robert Herman showed that Gamow's idea was flawed because there were roadblocks to assembling succeedingly heavier nuclei simply by adding protons or neutrons. **There are no stable nuclei with five or eight nucleons**, leaving ${}^{4}_{2}He$ as the heaviest element that can be formed as Gamow proposed. (A small amount of an isotope of lithium, ${}^{7}_{3}Li$, is also formed in the early universe by the nuclear fusion of ${}^{4}_{2}He$ with ${}^{3}_{1}H$ and the fusion of ${}^{4}_{2}He$ with ${}^{3}_{2}He$. The latter produces ${}^{7}_{4}Be$, which radioactively decays to ${}^{7}_{3}Li$.)

At that time, there was also a large problem with the idea of a hot, dense universe coming into existence approximately one Hubble time ago. Edwin Hubble's original value of his constant was $H_0 = 500$ km s^{-1} Mpc^{-1}, which corresponds to $t_H = 1/H_0 = 10^9$ yr for the age of the universe. This is only a fraction of Earth's age, which in 1928 had been radioactively dated as several billion years. It was certainly hard to understand how Earth could be older than the universe.

The steady state model

In 1946 Hermann Bondi, Thomas Gold, and Fred Hoyle attempted to find an alternative to Gamow's unpalatable Big Bang universe. In papers published in 1948 and 1949, they proposed their model of a **steady-state universe**. It extended the cosmological principle to include time, stating that, **in addition to the universe being isotropic and homogeneous, it also appears the same *at all times*.** A steady-state universe has no beginning and no end. It is infinitely old, and as it expands, a continuous creation of matter is required to maintain the average density of the universe at its present level. This changes the interpretation of the Hubble time; rather than the characteristic age of the universe, t_H becomes a characteristic time for the creation of matter.

If the universe roughly doubles in size in time t_H, then its volume becomes eight times greater, and so the rate of matter creation required to maintain the universe as it is today is approximately $8\rho_0/t_H = 8H_0\rho_0$. Just a few hydrogen atoms per cubic meter of space would need to be created every ten billion years, a rate far too small to be measured experimentally. **In the original steady-state models, the "when," "where," and "how" of the spontaneous appearance of new matter (in violation of the law of conservation of mass–energy) were questions left unanswered.** The appeal of the steady-state universe was its resolution of the timescale problem.

The steady state model

Hoyle sought an explanation in the nuclear reactions that took place inside stars. He joined forces with Geoffrey and Margaret Burbidge, and William Fowler. In 1957 they published their seminal paper, referred to as B^2FH, that laid out the theory of **stellar nucleosynthesis.**

The B^2FH analysis was a success, and its results were compatible with both the Big Bang and steady-state cosmologies. During the 1950s both theories had their supporters and detractors.

However, the steady-state theory had a serious problem explaining the large amount of helium observed in the universe. Astronomers had established that **about one-quarter (0.274 ± 0.016) of the baryonic mass of the universe is in the form of helium.** When compared to the cosmic abundances of the heavier elements, it was clear that **stellar nucleosynthesis could not account for the amount of helium observed,** especially considering that carbon, nitrogen, and oxygen are the results of exhaustively burning the star's helium core.

Gamow, Alpher, and Herman had shown that **the Big Bang could at least explain the abundance of helium, but where was the proof that such a violent event had ever occurred?**

The cooling of the Universe

The dense, early universe must have been very hot. In this hot, dense universe, the **mean free path of photons would have been short enough to maintain thermodynamic equilibrium**. Although an expanding universe cannot be precisely in equilibrium, this assumption of thermodynamic equilibrium is extremely good.

Under these conditions the **radiation field has a blackbody spectrum.** In 1948 Alpher and Herman published their description of how **this blackbody radiation would have cooled as the universe expanded,** and they predicted that the universe should now be filled with blackbody radiation at a **temperature of 5 K**.

The cooling of the blackbody radiation can be derived by considering its energy density $u = aT^4$. According to the fluid equation (Eq. 53) with $w_{rad} = 1/3$ for blackbody blackbody radiation and $R(t_0) = 1$,

$$R^{3(1+w_{\rm rad})} u_{\rm rad} = R^4 u_{\rm rad} = u_{\rm rad,0}. \tag{57}$$

The cooling of the Universe

The energy density today, $u_{rad,0}$, is smaller than the earlier value u_{rad} by a factor of R^4; a factor of R^3 is due to the fact that the volume of the universe has increased since then, and the other factor of R comes from the lesser energy of today's longer-wavelength photons ($E_{photon} = hc/\lambda$), a result of the cosmological redshift. Thus

$$R^4 a T^4 = a T_0^4,$$

and we find that the present blackbody temperature must be related to the temperature at an earlier time by

$$RT = T_0. \tag{58}$$

That is, the product of the scale factor and the blackbody temperature remains constant as the universe expands. When the universe was half as large, it was twice as hot.

The cooling of the Universe

An order-of-magnitude **estimate** of the present blackbody temperature of the universe may be calculated by **considering the temperature and baryonic mass density that must have prevailed in the early universe when helium was being formed**. The fusion of hydrogen nuclei requires roughly that T $\simeq 10^9$ K and $\rho_b \simeq 10^{-2}$ kg m^{-3}. If the temperature were any higher, the deuterium nuclei involved in the fusion chain would have undergone photodissociation due to the presence of energetic blackbody radiation, whereas a lower temperature would have made the Coulomb barrier between the nuclei too difficult to overcome. The quoted density is needed to produce the observed amounts of ${}^3_2He$ and other nuclei. From Eqs. (5) and (17), the value of the scale factor at the time of the helium formation was roughly

$$R \simeq \left(\frac{\rho_{b,0}}{\rho_b}\right)^{1/3} = 3.47 \times 10^{-9}.$$

At that time, the universe was only a few billionths of its present size.

The cooling of the Universe

Combining the scale factor with T (R) = 10^9 K, the present temperature of the blackbody radiation can be estimated from Eq. (58) as

$$T_0 = RT(R) \simeq 3.47 \text{ K},$$

similar to Alpher and Hermann's original estimate of 5 K in 1948.

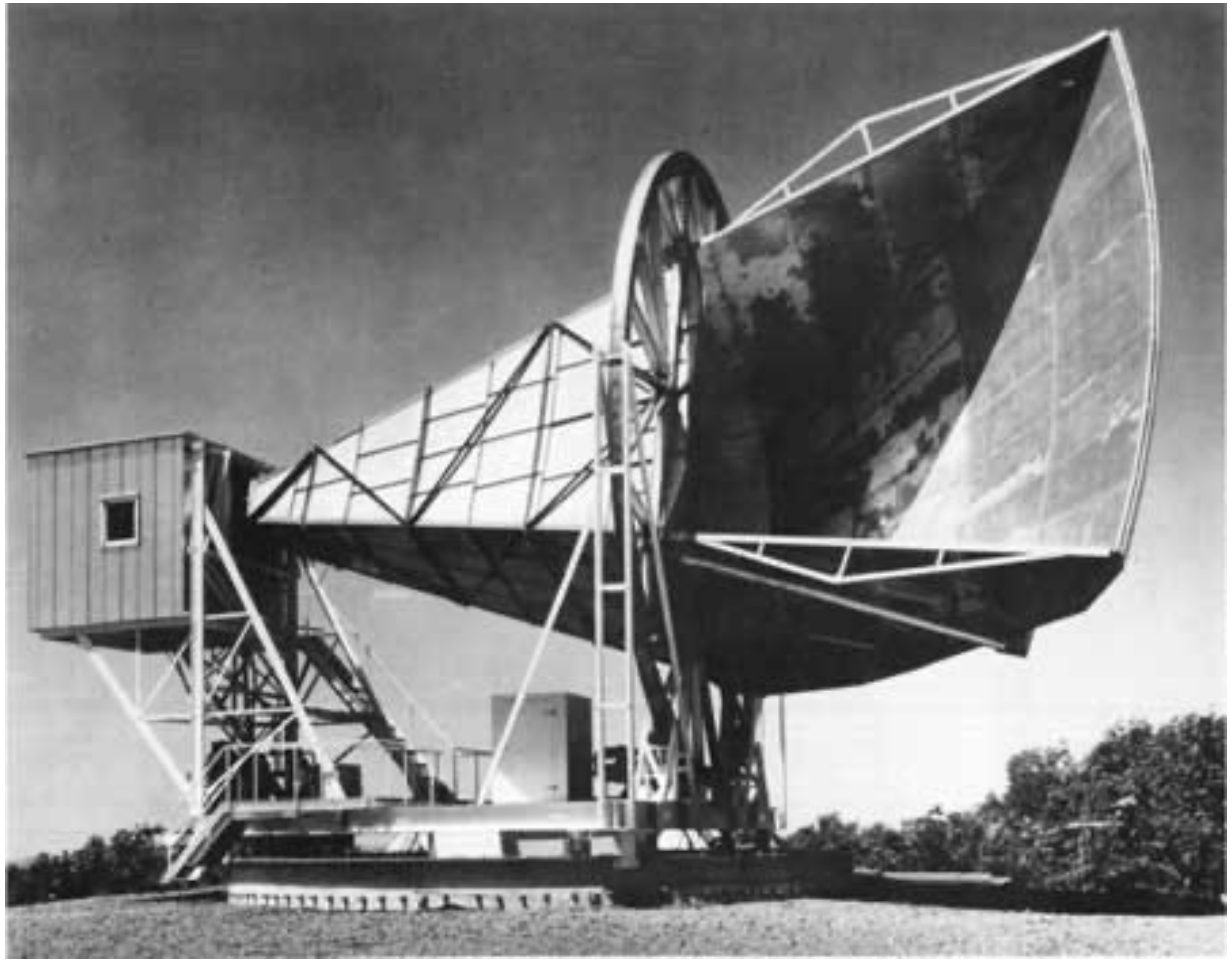
Wien's law then gives the peak wavelength of the blackbody spectrum as

$$\lambda_{\text{max}} = \frac{0.00290 \text{ m K}}{T_0} \simeq 8.36 \times 10^{-4} \text{ m}.$$

The discovery of the CMB

In 1964 Peebles calculated that the blackbody radiation left over from the Big Bang should have a temperature of about 10 K.

Arno Penzias and Robert Wilson were working with a huge horn reflector antenna that had been used to communicate with the new Telstar satellite. Despite a year of effort, the two men had been unable to get rid of **a persistent hiss in the signal.** The hiss came continually from all directions in the sky and remained even after Penzias and Wilson had scrubbed their antenna clean, taped over seams and rivets, and removed two pigeons that had nested inside the horn. They knew that a **3 K blackbody** would produce their interference but were unaware of any possible source until Penzias learned of Peeble's calculation of a 10K background.



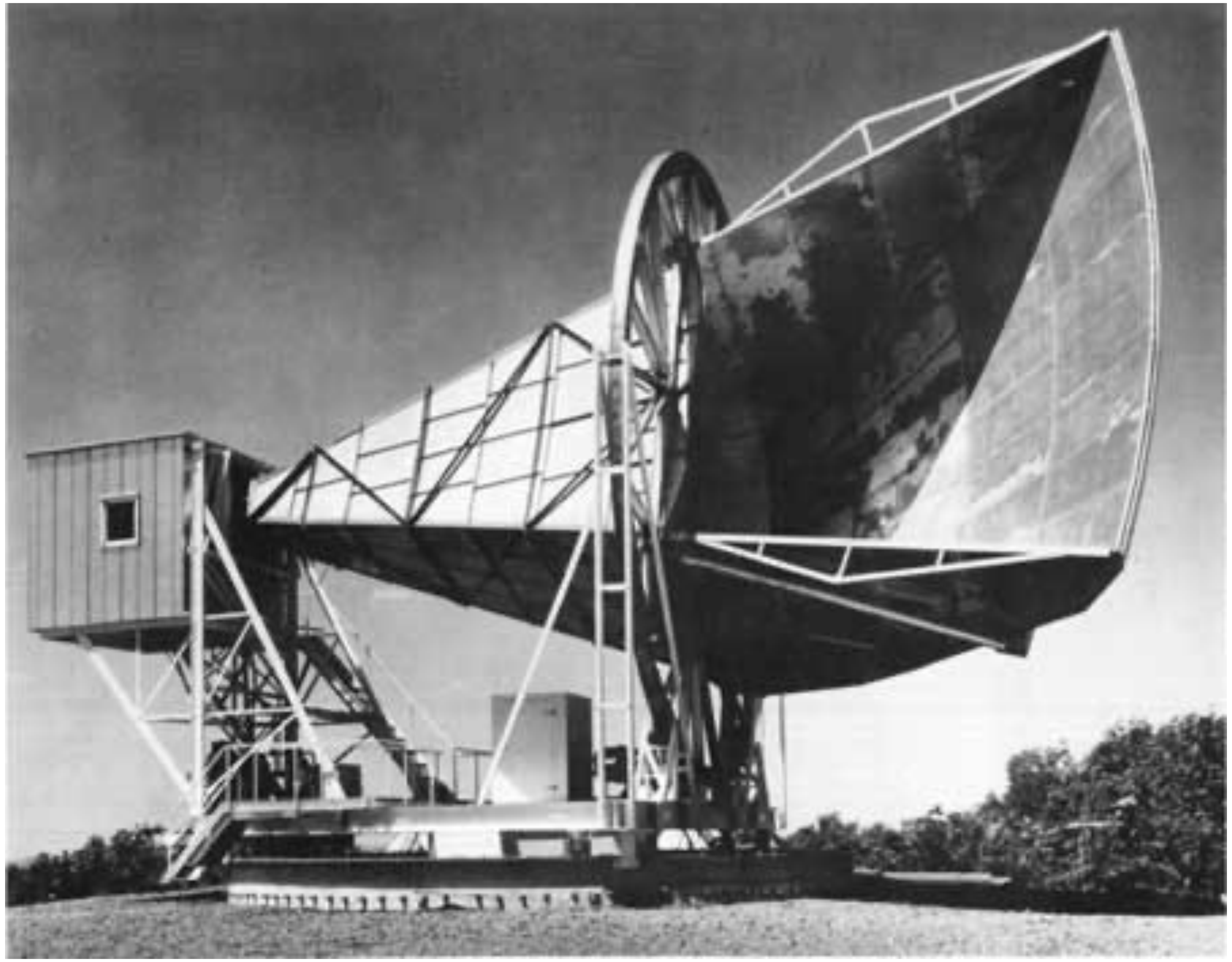
The **horn antenna** at Bell Labs, Holmdel, NJ used by Penzias and Wilson to discover the 3 K cosmic microwave background radiation in 1965

The discovery of the CMB

Penzias and Wilson had detected the blackbody radiation that fills the universe, with a peak wavelength of λ_{max} = 1.06 mm in the microwave region of the electromagnetic spectrum.

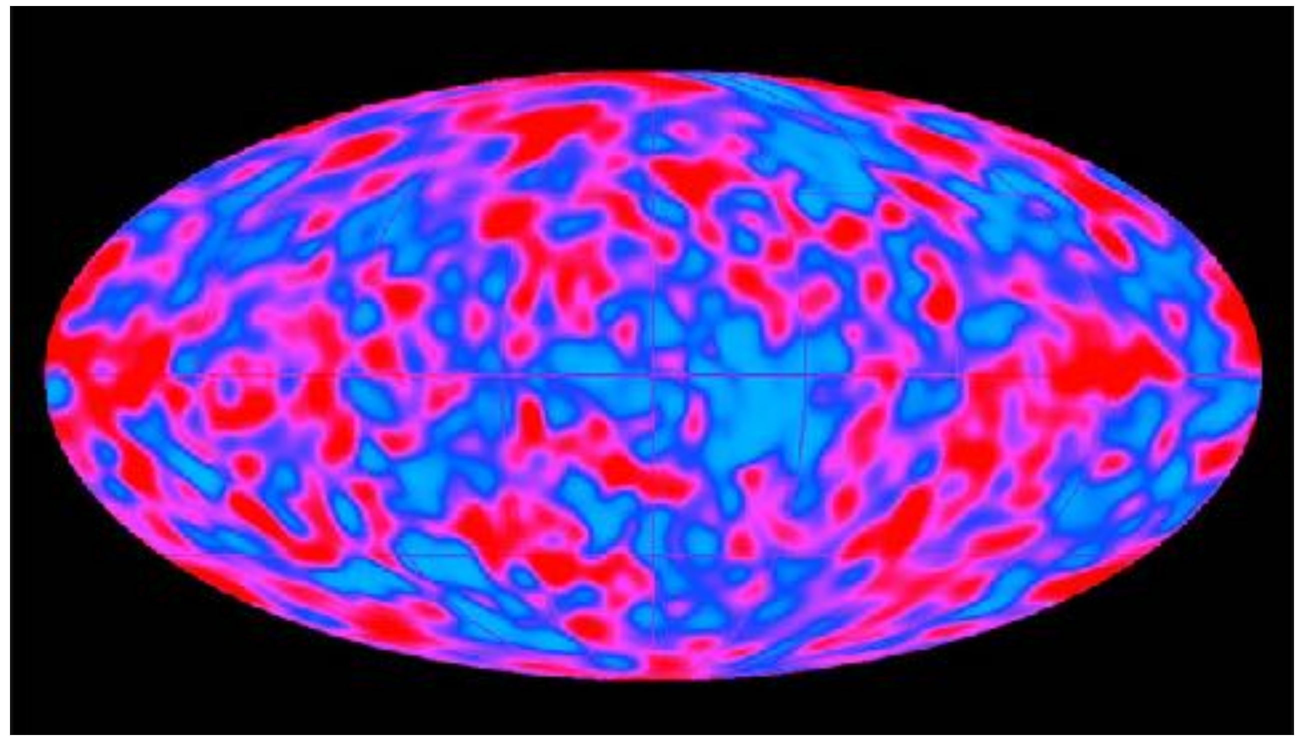
This afterglow of the Big Bang is now known as the **cosmic microwave background**, often abbreviated as the **CMB**.

The discovery of the CMB was the end of the steady-state cosmology. Further measurements at other wavelengths confirmed that the **shape of the CMB spectrum was that of a blackbody.**



The **horn antenna** at Bell Labs, Holmdel, NJ used by Penzias and Wilson to discover the 3 K cosmic microwave background radiation in 1965

The discovery of the CMB



COBE image of the CMB

In 1991 a striking measurement of the cosmic microwave background was obtained by the COBE satellite. The COBE measurement of the spectrum of the CMB is shown in the Figure 9. The data points fall almost perfectly on the theoretical spectrum of a **2.725-K blackbody**. The Planck function $B_\nu(T)$ in Fig. 9 peaks at a frequency of 160 GHz, corresponding to the frequency version of Wien's law,

$$\frac{\nu_{\max}}{T} = 5.88 \times 10^{10}\ \mathrm{Hz\ K^{-1}}. \tag{59}$$

The WMAP value for the CMB is

$$\boxed{[T_0]_{\mathrm{WMAP}} = 2.725 \pm 0.002\ \mathrm{K},} \tag{60}$$

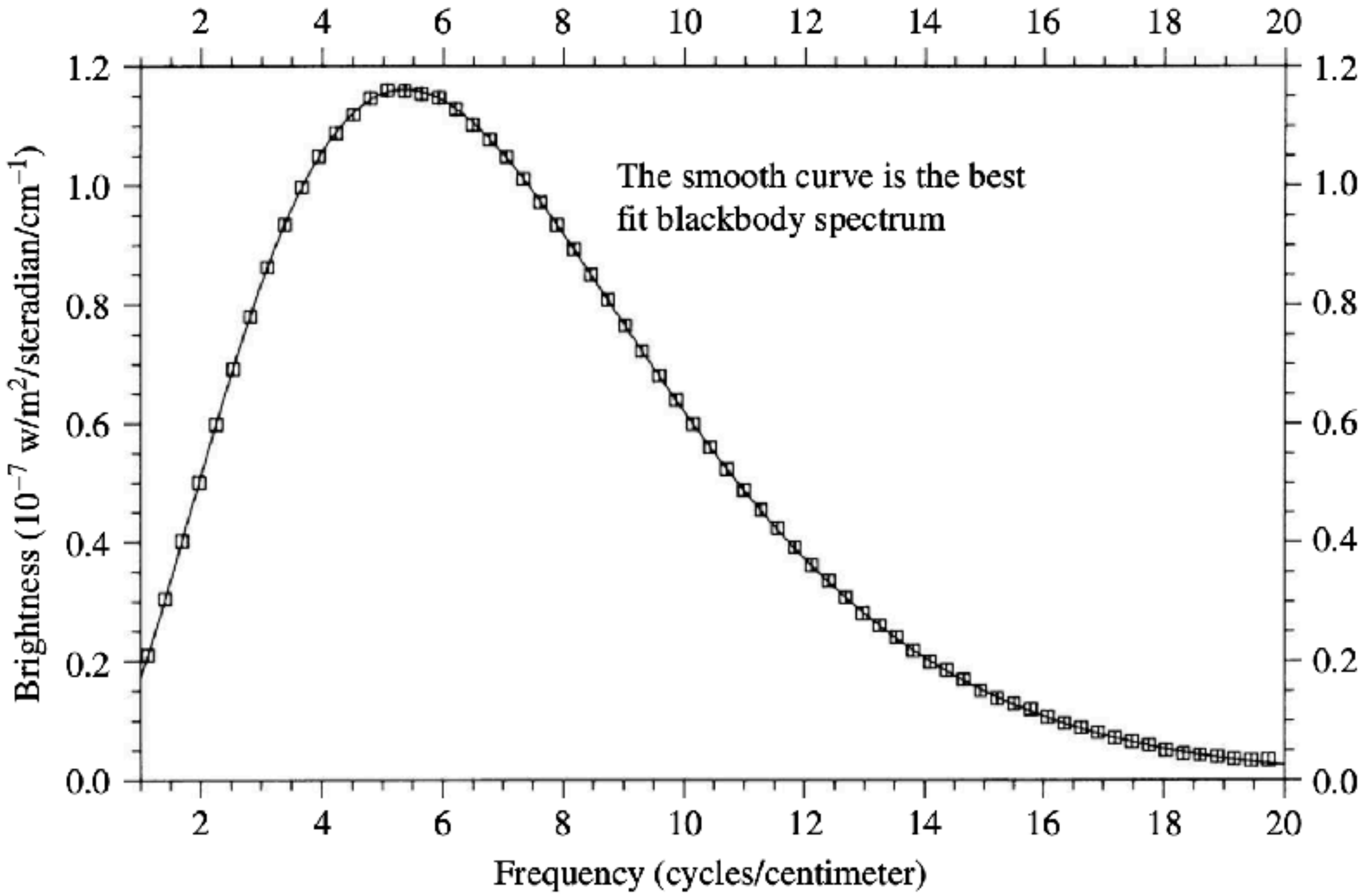
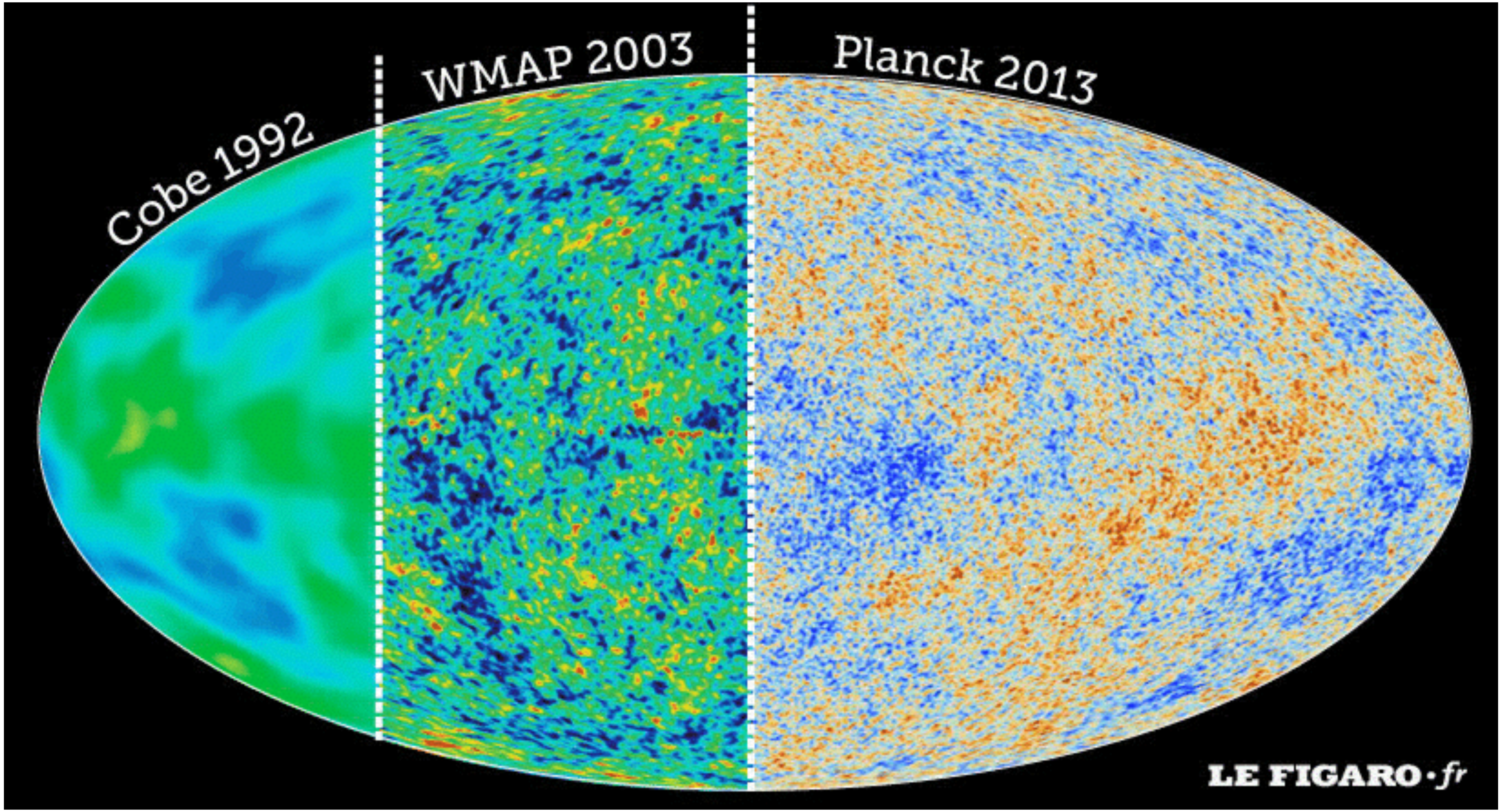


FIGURE 9 The COBE measurement of the spectrum of the cosmic microwave background, which is that of a blackbody with a temperature of 2.725 K. The horizontal axis (Frequency) is actually $1/\lambda$ (cm^{-1}); the spectrum peaks at a frequency of 160 GHz (5.35 cycles per centimeter). (Figure adapted from Mather et al., *Ap. J. Lett.*, *354*, L37, 1990. Courtesy of NASA/GSFC and the COBE Science Working Group.)

The discovery of the CMB



CMB measurements overt the years.

The dipole anisotropy of the CMB

The cosmic microwave background suffuses the entire universe. It does not emanate from any object but, rather, originated in the early universe right after the Big Bang. For this reason, **all observers at rest with respect to the Hubble flow** (no peculiar velocity) **see the same spectrum for the CMB, with the same intensity in all directions** (the CMB is isotropic). In particular, two observers in different galaxies that are being carried apart by the Hubble flow see the same blackbody spectrum.

However, there is a **Doppler shift of the CMB caused by an observer's peculiar velocity through space**, relative to the Hubble flow. Using Wien's law, **a shift in wavelength can be expressed as a change in the temperature of the blackbody radiation**. For example, a slight blueshift (smaller λ_{max}) would correspond to a slightly higher temperature.

Suppose an observer at rest relative to the Hubble flow determines that the cosmic microwave background has a temperature T_{rest}. Then, the temperature measured by an observer with a peculiar velocity v relative to the Hubble flow is

$$T_{\text{moving}} = \frac{T_{\text{rest}}\sqrt{1 - v^2/c^2}}{1 - (v/c)\cos\theta}, \tag{61}$$

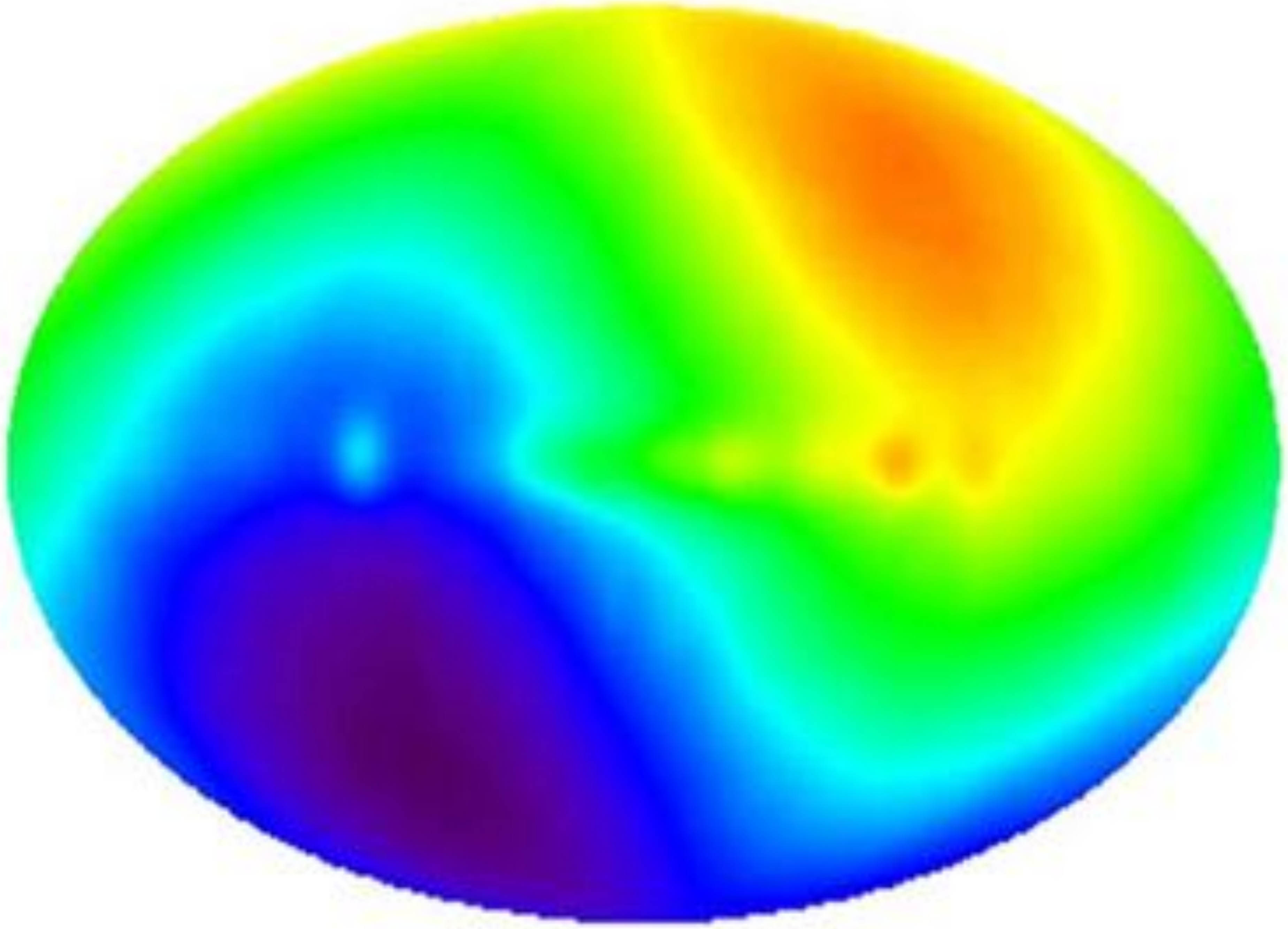
where θ is the angle between the direction of observation and the direction of motion.

The dipole anisotropy of the CMB

Both observers see a blackbody spectrum, but the **moving observer measures a slightly hotter temperature in the forward direction (θ = 0) and a slightly cooler temperature in the opposite direction**. If the peculiar velocity is v ≪ c, then

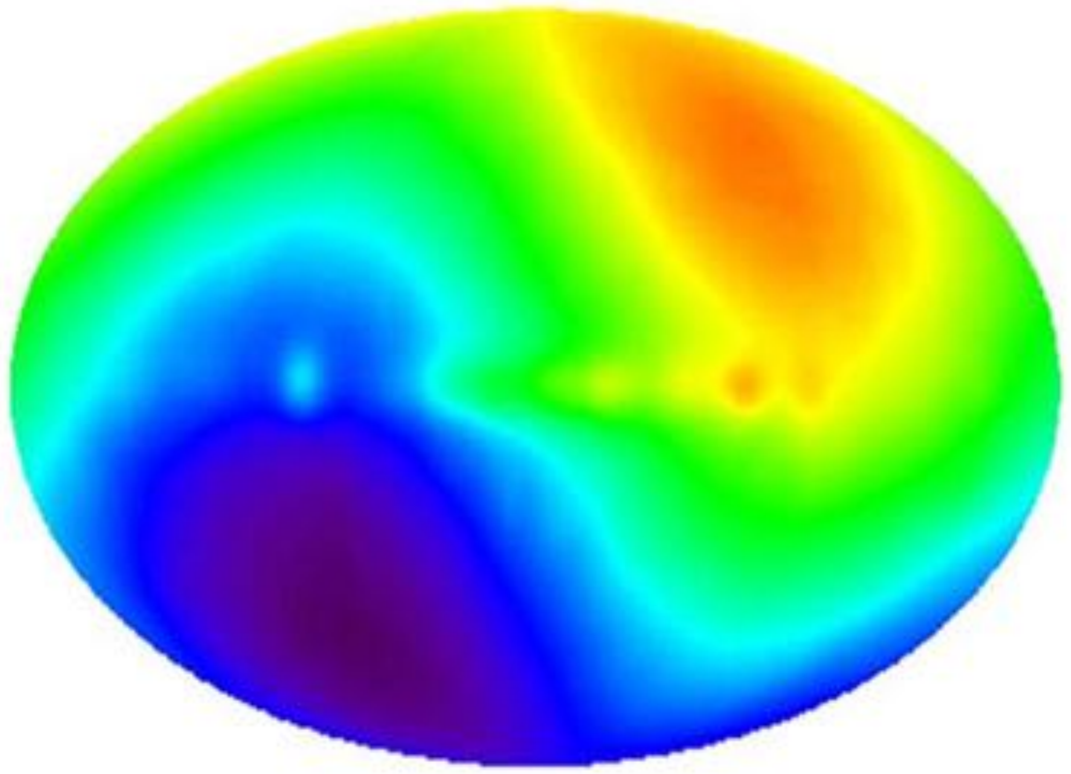
$$T_{\text{moving}} \simeq T_{\text{rest}} \left(1 + \frac{v}{c} \cos\theta\right) \tag{62}$$

The second term on the right-hand side, called the **dipole anisotropy** of the CMB, has been detected and measured; see Figure.



The dipole anisotropy if the CMB.

The dipole anisotropy of the CMB



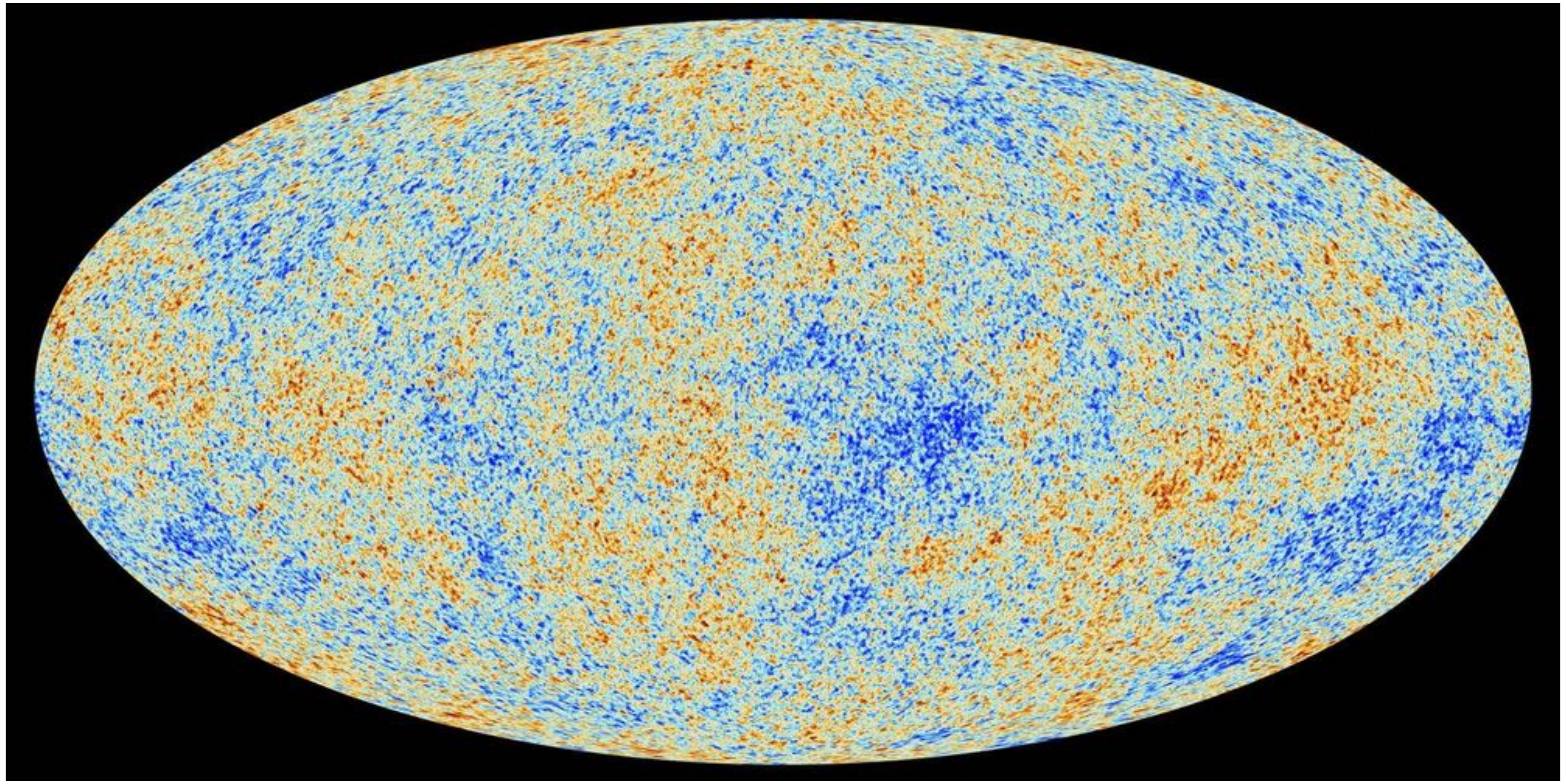
The temperature variation indicates that **the peculiar velocity of the Sun relative to the Hubble flow** is 370.6 ± 0.4 km s^{-1} in the direction $(\alpha, \delta) = (11.2^h, -7^\circ)$, between the constellations of Leo and Crater.

Of course, **the Sun is orbiting the Galaxy, and the Milky Way is moving within the Local Group of galaxies**. When these motions are accounted for, **the peculiar motion of the Local Group relative to the Hubble flow is** about 627 km s^{-1} toward $(\alpha, \delta) = (11.1^h, -27^\circ)$, in the middle of the constellation Hydra.

From this observation and measurements of the velocities of other galaxies and clusters of galaxies, astronomers have discovered **a large-scale streaming motion of thousands of galaxies** at ~ 600 km s^{-1} in the direction of $(\alpha, \delta) = (13.3^h, -44^\circ)$ in the constellation Centaurus. The Hydra–Centaurus supercluster is also being carried along in this riverlike perturbation of the Hubble flow.

The dipole anisotropy of the CMB

After the dipole anisotropy has been subtracted from the CMB, the remaining **radiation is incredibly isotropic**, having nearly equal intensity in all directions. Sensitive instruments, however, have revealed that **the CMB does have hotter and cooler areas.** The CMB appears as a patchwork of small regions, **about 1° degree or less in diameter,** where the temperature departs from the average value (T_0) by about one part in 10^5. Careful observations and analyses of these regions by WMAP and various ground-based and balloon-borne experiments have produced the first precision measurement of cosmological parameters.



The Planck measurements provide the latest, highest precision measurements for cosmology.

The Sunyaev Zel'dovich effect

It should be emphasized that an observer in a galaxy being carried along with the Hubble flow (no peculiar velocity) does not measure a Doppler shift of the CMB. An observer in a distant galaxy receding from us at an appreciable fraction of the speed of light sees the same CMB spectrum that we do.

Evidence of this is produced when **low-energy photons of the CMB pass through the hot ($\simeq 10^8$ K) ionized intracluster gas in a cluster of galaxies**. A small fraction (typically 10^{-3} to 10^{-2}) of the photons are scattered to higher energies by the high-energy electrons in the gas. This **inverse Compton scattering** increases the frequency of a scattered photon by an average amount $\Delta\nu$ of

$$\frac{\overline{\Delta\nu}}{\nu} = 4\frac{kT_e}{m_e c^2}, \qquad (63)$$

where T_e is the temperature of the electron gas.

The resulting distortion of the CMB spectrum, shown in Fig. 11, is called the thermal **Sunyaev–Zel'dovich effect**.

$$\frac{\Delta T}{T_0} \simeq -2\frac{kT_e}{m_e c^2}\tau \qquad (64)$$

The Sunyaev Zel’dovich effect

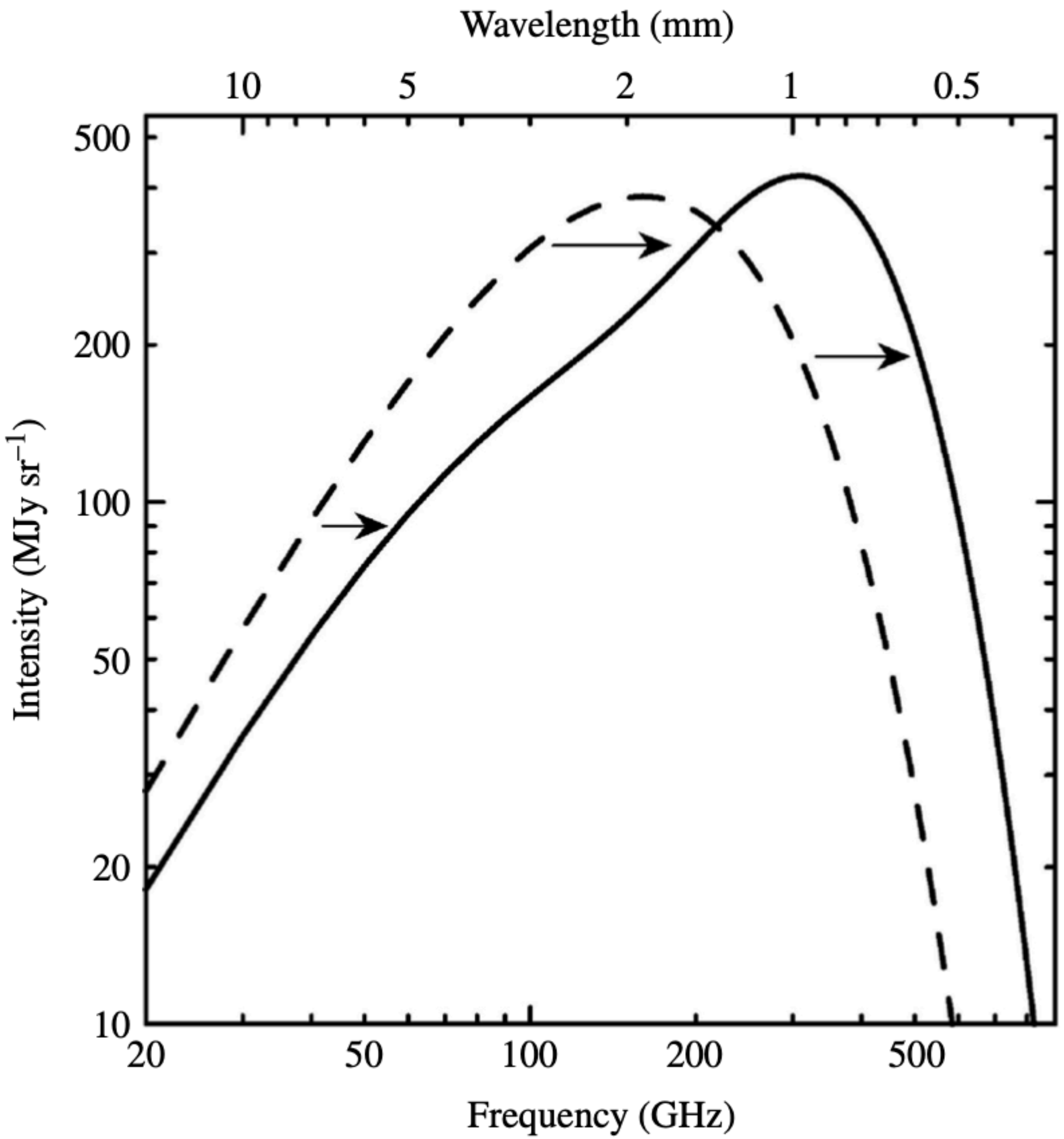


FIGURE 11 The undistorted CMB spectrum (dashed line) and the spectrum distorted by the Sunyaev–Zel’dovich effect (solid line). In a rich cluster of galaxies, CMB photons may be scattered to higher frequencies by colliding with the electrons in the hot intracluster gas. For frequencies less than the peak frequency, more photons are scattered out of a frequency interval than into it, so the intensity at that frequency decreases. Similarly, for frequencies greater than the peak frequency, fewer photons are scattered out of a frequency interval than into it, so the intensity at that frequency increases. The net result is a shift of the CMB spectrum to higher frequencies. The calculated distortion has been exaggerated by employing a fictional cluster 1000 times more massive than a typical rich cluster of galaxies. (Figure adapted from Carlstrom, Holder, and Reese, *Annu. Rev. Astron. Astrophys.*, *40*, 646, 2002. Reproduced with permission from the *Annual Review of Astronomy and Astrophysics*, Volume 40, ©2002 by Annual Reviews Inc.)

The Sunyaev Zel'dovich effect

Hot gas in a galaxy cluster. Purple - X-ray image.

Although the spectrum no longer has the precise shape of a blackbody, its translation to higher frequencies may be used to define an effective decrease ΔT in the temperature T_0 of the CMB of approximately

$$\frac{\Delta T}{T_0} \simeq -2 \frac{kT_e}{m_e c^2} \tau \qquad (64)$$

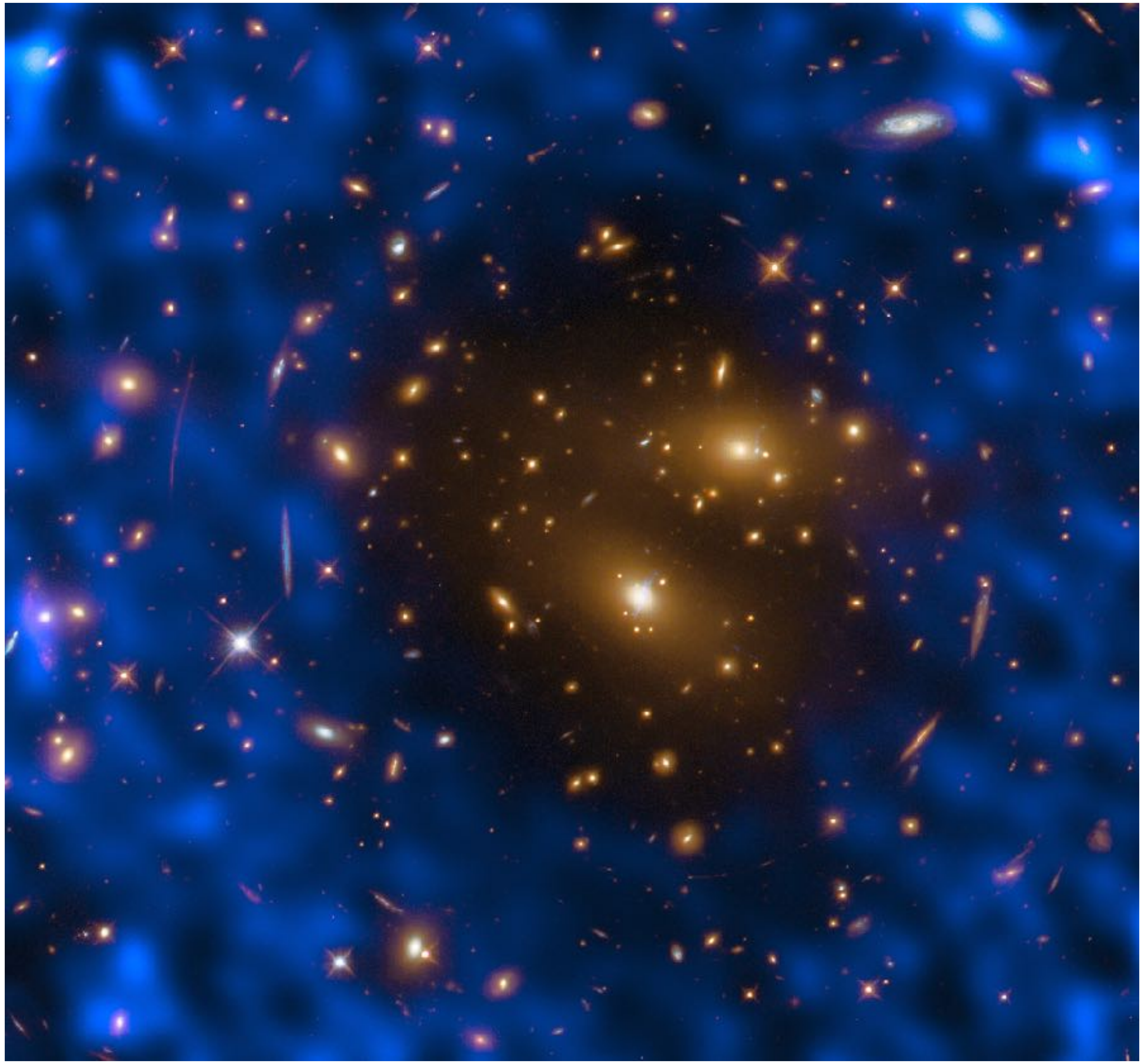
where τ is the optical depth of the intracluster gas along the line of sight.

Typical values of $\Delta T / T_0$ are a few times 10^{-4}. Observations of the Sunyaev–Zel'dovich effect for many clusters of galaxies confirm that **it is independent of the cluster's redshift**, as expected if the CMB spectrum observed at a cluster is not affected by the cluster's recessional velocity. Figure 12 shows the Sunyaev–Zel'dovich effect surrounding two clusters of galaxies.

In addition to confirming the cosmological nature of the CMB, the Sunyaev–Zel'dovich effect is a promising probe of the properties and evolution of clusters of galaxies in the early universe.

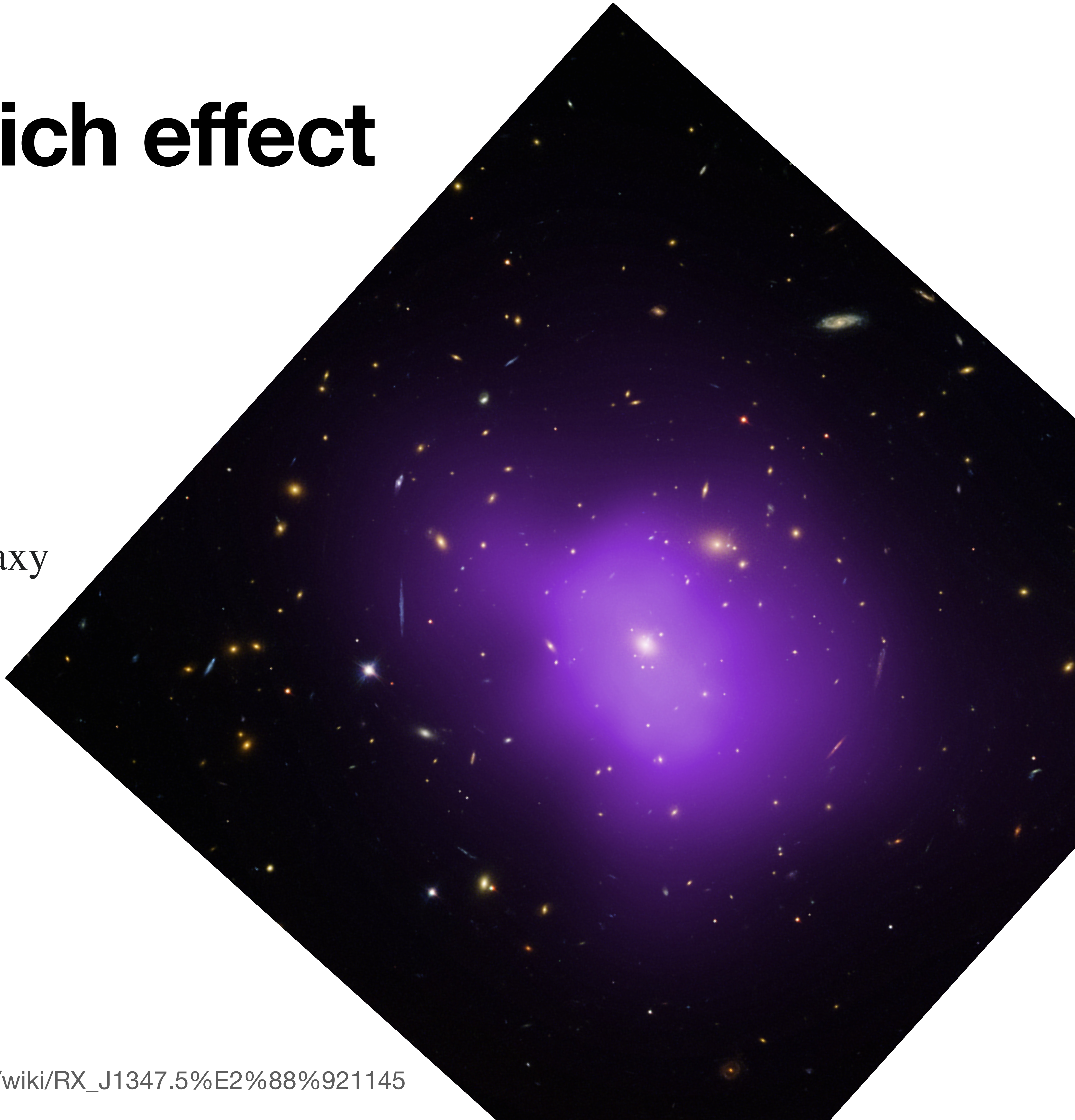
The Sunyaev-Zeldovich effect

This image shows the first measurements of the thermal Sunyaev-Zeldovich effect from the Atacama Large Millimeter/submillimeter Array (ALMA) in Chile (in blue). The target was **one of the most massive known galaxy clusters**, RX J1347.5–1145, the centre of which **shows up here in the dark "hole" in the ALMA observations.** The energy distribution of the CMB photons shifts and **appears as a temperature decrease at the wavelength observed by ALMA,** hence a dark patch is observed in this image at the location of the cluster.



https://en.wikipedia.org/wiki/RX_J1347.5%E2%88%921145

The Sunyaev-Zeldovich effect



The same galaxy cluster RX J1347.5–1145, here the purple ice shows X-ray emission from the cluster.
X-ray emission shows the hot ionised gas in the galaxy cluster.

https://en.wikipedia.org/wiki/RX_J1347.5%E2%88%921145

A Two-component model of the Universe

Previously we considered the expansion of a universe with a single component, pressureless dust, that was slowing down due to its own self-gravity. According to the relativistic equivalence of mass and energy, however, **the effect of the cosmic microwave background on the expansion must also be included.** We now know that **the gravitational effect of the CMB photons dominated the dynamics of the early universe**, although their effect is **completely negligible in the present universe.**

To incorporate this new feature, we introduce a **two-component model of the universe**,

- one that includes both the **total density of matter (baryonic and dark), ρ_m**,
- and the equivalent mass density of **relativistic particles (such as neutrinos and CMB photons), ρ_{rel}.**

It is the equation of state $\mathrm{P} = w\mathrm{u}$ (Eq. 52) that determines whether we count a particle as matter ($w_m = 0$), a relativistic particle (a photon or neutrino, for which $w_{rel} = 1/3$), or dark energy ($w_\Lambda = -1$). (The gravitational effect of the neutrinos' mass clearly persists; however, we will neglect the neutrinos' contribution to the value of $\Omega_{m,0}$ of roughly 0.003.)

A Two-component model of the Universe

$$R^{3(1+w)}\rho = \text{constant} = \rho_0, \tag{53}$$

Equation (53) shows that particles belonging to different categories are diluted differently by the expansion of the universe.

Of course, at earlier epochs when the universe was much hotter, even massive particles were relativistic. For example, an electron gas at $T > 6 \times 10^9$ K has $kT > m_e c^2$, implying that the electron gas is relativistic and its equation of state is described by $w_{rel} = 1/3$. However, we will ignore such complications for the remainder of this chapter and will consider photons and neutrinos as our only relativistic particles.

With both matter and relativistic particles included, Eq. (10) becomes

$$\left[\left(\frac{1}{R}\frac{dR}{dt}\right)^2 - \frac{8}{3}\pi G\left(\rho_m + \rho_{\text{rel}}\right)\right]R^2 = -kc^2. \tag{65}$$

The equivalent mass density of the CMB photons comes from the energy density of black-body radiation,

$$u_{\text{rad}} = aT^4, \tag{66}$$

A Two-component model of the Universe

where a is the radiation constant. We will rewrite this in the form

$$u_{\mathrm{rad}} = \frac{1}{2}\, g_{\mathrm{rad}} T^4, \tag{67}$$

where g_{rad} is the number of **degrees of freedom** of a photon. The value of g reflects the number of spin states n_{spin} and the possible existence of an antiparticle (n_{anti} = 1 or 2). A photon is its own antiparticle (n_{anti} =1) and can exist in n_{spin} =2 spin states, corresponding to its two possible polarizations with its spin parallel or antiparallel to its motion. Thus

$$g_{\mathrm{rad}} = 2 \tag{68}$$

for photons, as expected.

Neutrino decoupling

We will neglect the small mass of the other relativistic particle we are considering, the neutrino, and treat it as a massless particle. The **very early universe was sufficiently dense that neutrinos attained thermal equilibrium, with a spectrum very similar to that of blackbody radiation**, except the "−1" in the denominator of that equation becomes a "+1" for neutrinos. This occurs because photons are bosons, described by Bose–Einstein statistics, while **neutrinos are fermions, described by Fermi–Dirac statistics.**

Although the cosmic neutrino background has yet to be observed, we have confidence that it exists.

Recall that there are **three types (or flavors) of neutrinos**—ν_e, ν_μ and ν_τ—and that each neutrino has a corresponding antineutrino. The **total energy density of all three flavors** is given by

$$u_\nu = 3 \times \frac{7}{8} \times aT_\nu^4 = 2.625\, aT_\nu^4, \qquad (69)$$

where the 7/8 derives from the "+1" in the expression for Fermi–Dirac statistics, and T_ν is the temperature of the neutrinos.

Neutrino decoupling

As before, we write this as

$$u_\nu = \frac{1}{2}\left(\frac{7}{8}\right) g_\nu T_\nu^4, \tag{70}$$

where

$$g_\nu = 6 \tag{71}$$

In general,

$$g = (\#\ \text{types}) n_{\text{anti}} n_{\text{spin}}. \tag{72}$$

There is an antineutrino for each of the three types of neutrino so $n_{\text{anti}} = 2$, and neutrinos have one spin state (all neutrinos are left-handed) so $n_{\text{spin}} = 1$. We therefore recover $g_\nu = 3 \times 2 \times 1 = 6$.

Neutrino decoupling

The usual T in cosmology is always taken to be the temperature of the blackbody photons. However, the T_ν in Eq. (70) is the temperature of the neutrinos. For $T > 3.5 \times 10^{10}$ K, these temperatures are the same, and $T = T_\nu$. However, **as the temperature dropped below about 3.5×10^{10} K, the expansion of the universe diluted the number density of neutrinos, and the neutrinos ceased to interact with other particles.** Essentially, the cosmos expanded faster than the neutrino interaction rate, and the neutrinos decoupled from the other constituents of the universe.

Since the time of neutrino decoupling, the neutrinos have expanded and cooled at their own rate, independently of the CMB.

The energy density of relativistic particles

Because the **annihilation of electrons with positrons** continued to supply energy to the photons (via $e^- + e^+ \rightarrow \gamma + \gamma$) but not to the neutrinos, **the neutrino temperature is somewhat less than the temperature of the CMB photons**. Although it is beyond the scope of this book, it can be shown that T_ν is related to the temperature T of the CMB photons by

$$T_\nu = \left(\frac{4}{11}\right)^{1/3} T. \tag{73}$$

The total neutrino energy density is therefore

$$u_\nu = \frac{1}{2}\left(\frac{7}{8}\right) g_\nu \left(\frac{4}{11}\right)^{4/3} aT^4 = 0.681aT^4. \tag{74}$$

The energy density of relativistic particles

Thus the energy density for relativistic particles, both photons and neutrinos, is

$$u_{\rm rel} = \frac{1}{2} g_* T^4, \tag{75}$$

Where

$$g_* = g_{\rm rad} + \left(\frac{7}{8}\right) g_\nu \left(\frac{4}{11}\right)^{4/3} = 3.363 \tag{76}$$

is the **effective number of degrees of freedom** of the relativistic particles. We also define the equivalent mass density of relativistic particles as

$$\boxed{\rho_{\rm rel} = \frac{u_{\rm rel}}{c^2} = \frac{g_* T^4}{2c^2}. \qquad (77)}$$

This value of **g_* is valid back to the end of electron–positron annihilation, at about t = 1.3 s.** For the higher temperatures of the very early universe (t < 1 s), however, we will encounter a greater number of relativistic particles, and the value of g_* will grow accordingly.

The energy density of relativistic particles

Employing Eq. (8), Eq. (65) becomes

$$\boxed{H^2\left[1-(\Omega_m+\Omega_{\rm rel})\right]R^2=-kc^2, \qquad (78)}$$

Where

$$\Omega_m=\frac{\rho_m}{\rho_c}=\frac{8\pi G\rho_m}{3H^2} \tag{79}$$

is the density parameter for matter (both baryonic and dark), and

$$\Omega_{\rm rel}=\frac{\rho_{\rm rel}}{\rho_c}=\frac{8\pi G\rho_{\rm rel}}{3H^2}=\frac{4\pi G g_* a T^4}{3H^2c^2} \tag{80}$$

is the density parameter for relativistic particles (both **photons and neutrinos**).

The energy density of relativistic particles

Note that Eq. (78) implies that for a flat (k = 0) two-component universe, $\Omega_m + \Omega_{rel} = 1$. Inserting T_0 = 2.725 K, we find that

$$\Omega_{\mathrm{rel},0} = 8.24 \times 10^{-5},$$

which is very small compared with $[\Omega_{m,0}]$WMAP = 0.27.

Radiation dominated to matter dominated

Recalling that w_{rel}= 1/3 for relativistic particles, Eq. (53) yields

$$R^4 \rho_{\rm rel} = \rho_{\rm rel,0}, \quad (81)$$

which shows **how the equivalent mass density of relativistic particles varies with the scale factor R.** By comparing this with Eq. (5),

$$R^3 \rho_m = \rho_{m,0}, \quad (82)$$

for massive particles, we notice that ρ_{rel} increases more rapidly than the mass density ρ_m as the scale factor becomes smaller. As R → 0 **in the early universe, therefore, there must have been an early era when the radiation** (i.e., all relativistic particles, not just γ, ν) **dominated and governed the expansion of the universe.**

Radiation dominated to matter dominated

The transition from this **radiation era** to the present **matter era** occurred when the scale factor satisfied $\rho_{rel} = \rho_m$, or $\Omega_{rel} = \Omega_m$. From Eqs.(79–82), the equality of Ω_{rel} and Ω_m occurred when the scale factor was

$$R_{r,m} = \frac{\Omega_{\rm rel}}{\Omega_{m,0}} = 4.16 \times 10^{-5} \Omega_{m,0}^{-1} h^{-2},$$

with a WMAP value of

$$R_{r,m} = 3.05 \times 10^{-4}.$$

This corresponds to a redshift (Eq. 4) of

$$z_{r,m} = \frac{1}{R_{r,m}} - 1 = 2.41 \times 10^4 \Omega_{m,0} h^2,$$

Radiation dominated to matter dominated

which for WMAP values is

$$z_{r,m} = 3270.$$

This is in very good agreement with the WMAP result,

$$[z_{r,m}]_{\text{WMAP}} = 3233\ ^{+194}_{-210},$$

for the redshift when the the universe passed from being radiation-dominated to being matter-dominated. Using RT = T_0(Eq. 58), the temperature at this transition was

$$T_{r,m} = \frac{T_0}{R_{r,m}} = 6.56 \times 10^4 \Omega_{m,0} h^2 \text{ K}, \qquad T_{r,m} = 8920 \text{ K}$$

using WMAP values. Thus, when the universe had cooled to 8920 K and typical separations were some 4 $\times$ 10^{-5} of their present extent, relativistic particles ceased to govern the cosmic expansion, and matter assumed a dominant role.

Expansion in the two component model

To determine how the early universe expanded with time means **how the scale factor, R, behaved during the radiation era,** we begin by substituting Eqs. (81) and (82) into Eq. (65) to find

$$\left[\left(\frac{dR}{dt}\right)^2 - \frac{8}{3}\pi G\left(\frac{\rho_{m,0}}{R} + \frac{\rho_{\mathrm{rel},0}}{R^2}\right)\right] = -kc^2. \tag{83}$$

Because the early universe was essentially flat, we can set k = 0 and use a bit of algebra to obtain

$$\int_0^R \frac{R'\,dR'}{\sqrt{\rho_{m,0}R' + \rho_{\mathrm{rel},0}}} = \sqrt{\frac{8\pi G}{3}}\int_0^t dt'.$$

Integrating this eventually yields an expression for the age of the universe as a function of the scale factor R:

$$t(R) = \frac{2}{3}\frac{R_{r,m}^{3/2}}{H_0\sqrt{\Omega_{m,0}}}\left[2 + \left(\frac{R}{R_{r,m}} - 2\right)\sqrt{\frac{R}{R_{r,m}} + 1}\,\right], \tag{84}$$

Expansion in the two component model

where
$$\frac{2}{3}\frac{R_{r,m}^{3/2}}{H_0\sqrt{\Omega_{m,0}}} = 5.51 \times 10^{10} h^{-4}\Omega_{m,0}^{-2}\ \mathrm{s} = 1.75 \times 10^{3} h^{-4}\Omega_{m,0}^{-2}\ \mathrm{yr}.$$

The time $t_{r,m}$ of the transition from a radiation-dominated to a matter-dominated universe may be found by setting $R/R_{r,m} = 1$ and using WMAP values of $h = 0.71$ and $\Omega_{m,0} = 0.27$ to obtain

$$t_{r,m} = 1.74 \times 10^{12}\ \mathrm{s} = 5.52 \times 10^{4}\ \mathrm{yr}. \qquad (85)$$

The form of Eq. (84) becomes simpler deep in the radiation era, when $R \ll R_{r,m}$.

In this limit the factors of $\Omega_{m,0}$ cancel, resulting in

$$R(t) = \left(\frac{16\pi G g_* a}{3c^2}\right)^{1/4} T_0\, t^{1/2} \qquad (86)$$

$$= (1.51 \times 10^{-10}\ \mathrm{s}^{-1/2}) g_*^{1/4} t^{1/2}. \qquad (87)$$

Expansion in the two component model

This shows that during the radiation era, $R \propto t^{1/2}$. Using $T = T_0/R$ quickly reveals the temperature deep in the radiation era:

$$T(t) = \left(\frac{3c^2}{16\pi G g_* a}\right)^{1/4} t^{-1/2} \tag{88}$$

$$= (1.81 \times 10^{10}\ \mathrm{K\,s}^{1/2}) g_*^{-1/4} t^{-1/2}. \tag{89}$$

At the other extreme, for $R \gg R_{r,m}$, Eq. (84) becomes

$$t(R) = \frac{2}{3} \frac{R^{3/2}}{H_0\sqrt{\Omega_{m,0}}} \tag{90}$$

So

$$R(t) = \left(\frac{3}{2} H_0 t \sqrt{\Omega_{m,0}}\right)^{2/3} = \left(\frac{3\sqrt{\Omega_{m,0}}}{2}\right)^{2/3} \left(\frac{t}{t_H}\right)^{2/3}, \tag{91}$$

Expansion in the two component model

using $t_H = 1/H_0$ for the Hubble time. As expected, this displays the $R \propto t^{2/3}$ dependence we found earlier in Eq. (30) for a flat universe of pressureless dust.

Equation (90) can be expressed in terms of z using R = 1/(1 + z) to obtain

$$\frac{t(z)}{t_H} = \frac{2}{3}\frac{1}{(1+z)^{3/2}\sqrt{\Omega_{m,0}}}, \tag{92}$$

which may be compared with Eq. (40) for a flat universe of pressureless dust. Evaluating these for R = 1 (z = 0) and using WMAP values **gives the age of the universe as 12.5 billion years**, a billion years more than the mean age of the oldest globular clusters.
However, as we will see later, this estimate of the age of the universe is **still about one billion years too short,** as determined by a full analysis of the WMAP results.

Big Bang nucleosynthesis

The process that manufactured the lightest elements in the early universe is known as **Big Bang nucleosynthesis**.

Why is approximately one-quarter of the mass of the universe in the form of helium?

The **temperature at time t during the radiation era** is given by Eq. (89). At a temperature just below 10^{12} K (t ~ 10^{-4} s), the universe contained a mixture of photons (γ), electron–positron pairs, and electron and muon neutrinos and their antiparticles ($\nu_e, \nu_\mu, \bar{\nu}_e,\ \bar{\nu}_\mu$). There were also a smaller number of protons and neutrons, about five for every 10^{10} photons, that were constantly being transformed into each other via the reactions

$$n \rightleftharpoons p^+ + e^- + \overline{\nu}_e \tag{93}$$

$$n + e^+ \rightleftharpoons p^+ + \overline{\nu}_e \tag{94}$$

$$n + \nu_e \rightleftharpoons p^+ + e^- . \tag{95}$$

These constant conversions were easily accomplished because the mass difference between a proton and a neutron is only

$$(m_p - m_n)c^2 = 1.293 \text{ MeV},$$

Big Bang nucleosynthesis

while the characteristic thermal energy of particles at 10^{12} K is kT ≃ 86 MeV. The Boltzmann equation gives the equilibrium ratio of the number density of neutrons, n_n, to the number density of protons, n_p, as

$$\frac{n_n}{n_p} = e^{-(m_p - m_n)c^2/kT} . \tag{96}$$

At 10^{12} K, this ratio is 0.985. The **numbers of neutrons and protons were nearly equal** because the mass difference between the protons and neutrons is negligible at such a high temperature.

As the universe expanded and the temperature fell, the ratio of the **number densities continued** to be given by Eq. (96) **as long as reactions (93– 95) proceeded fast enough to reach equilibrium**.

When the temperature had declined to about **10^{10} K**, the timescale for these reactions exceeded the characteristic timescale of the expansion given by 1/H(t) = 2t.

Big Bang nucleosynthesis

At a bit above 10^{10} K, the reaction rates decreased significantly, for two reasons.

First, **the expansion had reduced the energy of the neutrinos until they were unable to participate in reactions** (93–95).

Also, shortly thereafter, **the characteristic thermal energy of the photons, kT , fell below the 1.022 MeV** threshold for creating electron–positron pairs via the pair-production process $\gamma \rightarrow e^- + e^+$. As a result, the **electrons and positrons annihilated each other without being replaced**, leaving only a **small remainder of excess electrons**. For these reasons, the neutrons could not be replenished as fast as they were destroyed, and there was not enough time for these reactions to reach equilibrium. **In a sense, the creation of new neutrons could not keep up with the rate of expansion of the universe.** The ratio of the number densities then became “frozen” at its value of $n_n/n_p = 0.223$ when $T \simeq 10^{10}$ K.

At this point, there were 223 neutrons for every 1000 protons (or 446 neutrons for every 2000 protons), and essentially no more neutrons were being created. The **beta decay reaction** continued to operate, however, **converting neutrons into protons with a half-life of $\tau_{1/2}$ = 614 s = 10.2 min**.

Big Bang nucleosynthesis

It was not yet possible for the protons and neutrons to combine to form deuterium nuclei (^2H) via

$$p^+ + n \rightleftharpoons {}^2_1\mathrm{H} + \gamma,$$

because at temperatures exceeding 10^9 K, the energetic radiation quickly dissociated the nuclei. As a result, the **neutrons and protons remained separated until the temperature had dropped from 10^{10} K to 10^9 K.** According to Eq. (89) with $g_* = 3.363$, this took approximately

$$t(10^9\ \mathrm{K}) - t(10^{10}\ \mathrm{K}) = 178\ \mathrm{s} - 1.78\ \mathrm{s} \approx 176\ \mathrm{s}.$$

From the law of radioactive decay, in this amount of time the 446 neutrons mentioned previously declined to 366, and the number of protons rose to 2080.

Below **10^9 K the neutrons and protons readily combined to form as many deuterium** nuclei as possible. A number of reactions then led to the **formation of ^4He,** the most tightly bound nucleus involved in Big Bang nucleosynthesis.

Big Bang nucleosynthesis

The most efficient reactions leading to ^4He include

$$^2_1\mathrm{H} + ^2_1\mathrm{H} \rightleftharpoons ^3_1\mathrm{H} + ^1_1\mathrm{H}$$
$$^3_1\mathrm{H} + ^2_1\mathrm{H} \rightleftharpoons ^4_2\mathrm{He} + n$$

$$^2_1\mathrm{H} + ^2_1\mathrm{H} \rightleftharpoons ^3_2\mathrm{He} + n$$
$$^3_2\mathrm{He} + ^2_1\mathrm{H} \rightleftharpoons ^4_2\mathrm{He} + ^1_1\mathrm{H}.$$

[Note that these reactions differ from those of the pp chain, which produce helium in the cores of stars.]

No other nuclei were formed with abundances approaching that of ^4He, although there were traces of ^2H, ^3He, and ^7Li (from the reaction $^4He + ^3H \rightarrow ^7Li + \gamma$).

Big Bang nucleosynthesis

Figure 13 shows the network of reactions involved in Big Bang nucleosynthesis. Our sample of 366 neutrons and 2080 protons could form 183 ^4He nuclei, with 1714 protons (^{1} H) left over. Because a ^4He nucleus is four times more massive than a ^1H nucleus, the preceding analysis shows that the mass fraction of ^4He in the universe should have been about

$$\frac{4(183)}{1714 + 4(183)} = 0.299.$$

This rough estimate is consistent with the **primordial percentage of helium inferred from observations, between 23% and 24%.**

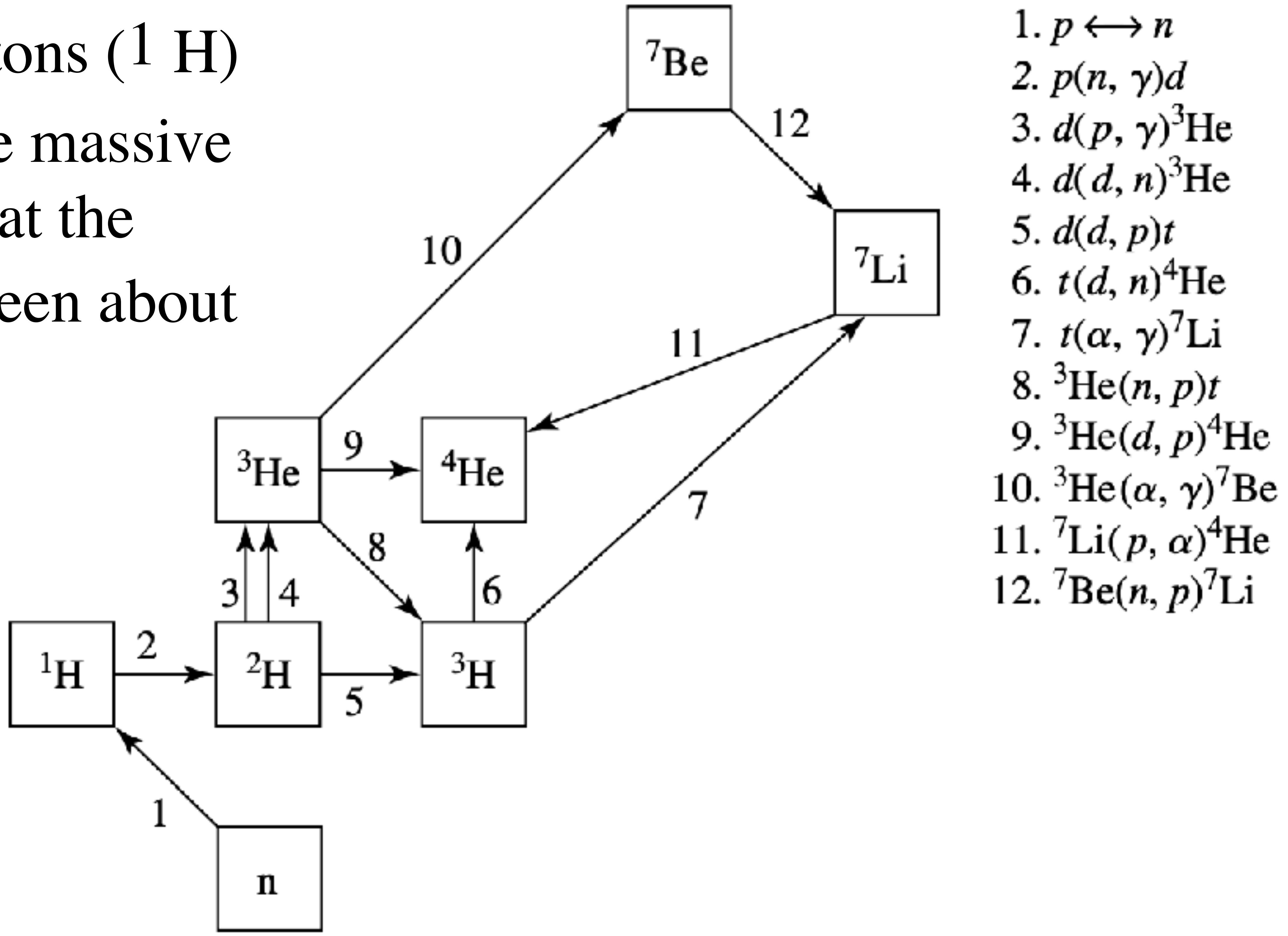
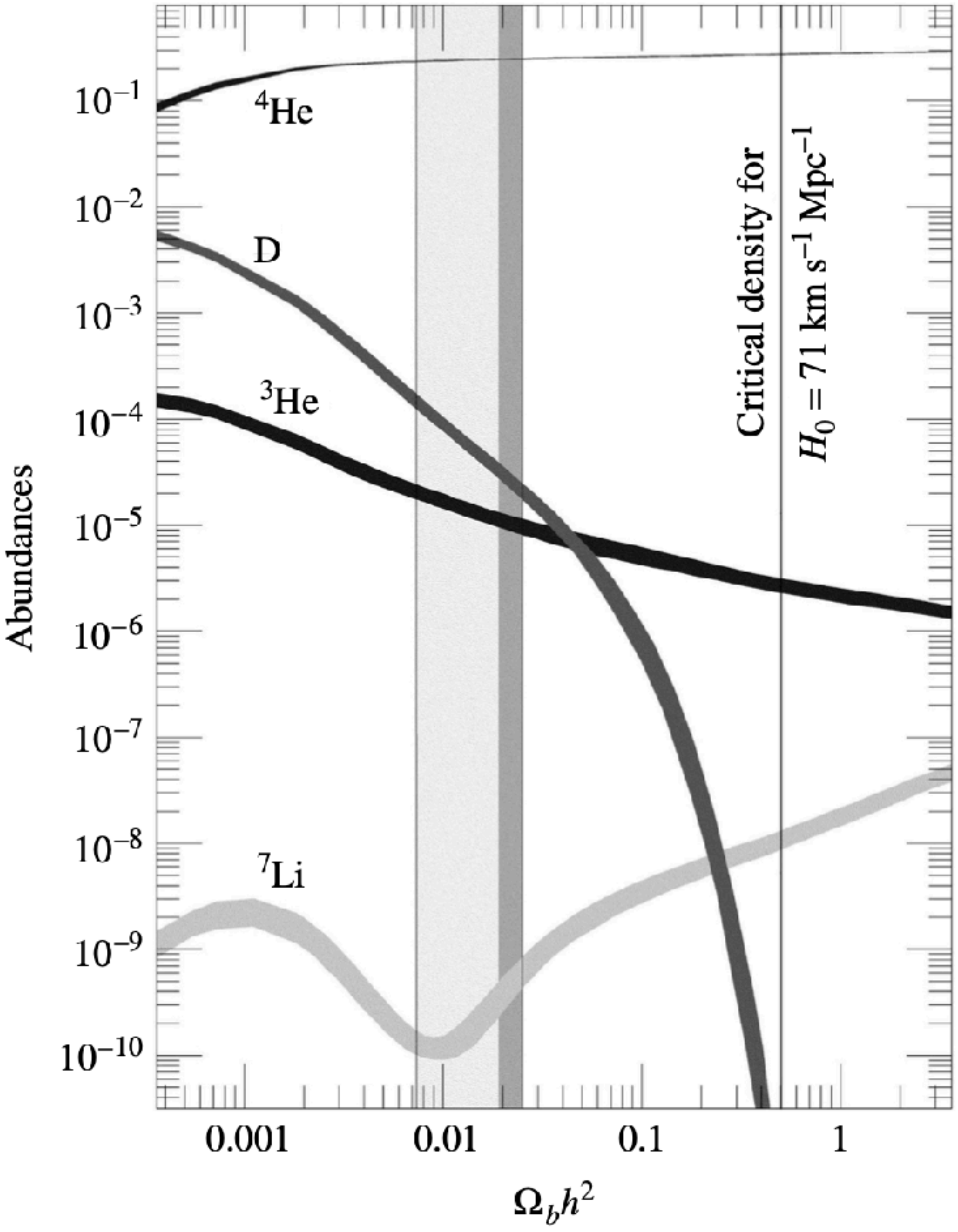


FIGURE 13 The reaction network that is responsible for Big Bang nucleosynthesis. The letter "*d*" stands for deuterium, and "*t*" stands for tritium. (Figure adapted from Nollett and Burles, *Phys. Rev. D*, *61*, 123505, 2000.)

Big Bang nucleosynthesis

Because essentially all of the available neutrons were incorporated in the ^4He nuclei, the abundance of ^4He was insensitive to the density of the universe at the time. However, **the amounts of ^2H, ^3He, and ^7Limanufactured in this way depend sensitively on the density of ordinary matter at the time of the reactions.**

Figure 14 shows the abundances of these nuclei as a function of the prevailing *present* density of baryonic matter. Comparing the theoretical curves with the observations makes it apparent that the present density of baryonic matter probably lies between 2 and 5 × 10^{-28} kg m^{-3}, only a few percent of the critical density of 1.88 × $10^{-26}h^2$ kg m^{-3}. This **explanation of the abundances of the light elements that were not manufactured by stars is one of the greatest achievements of the Big Bang theory.**



Big Bang nucleosynthesis

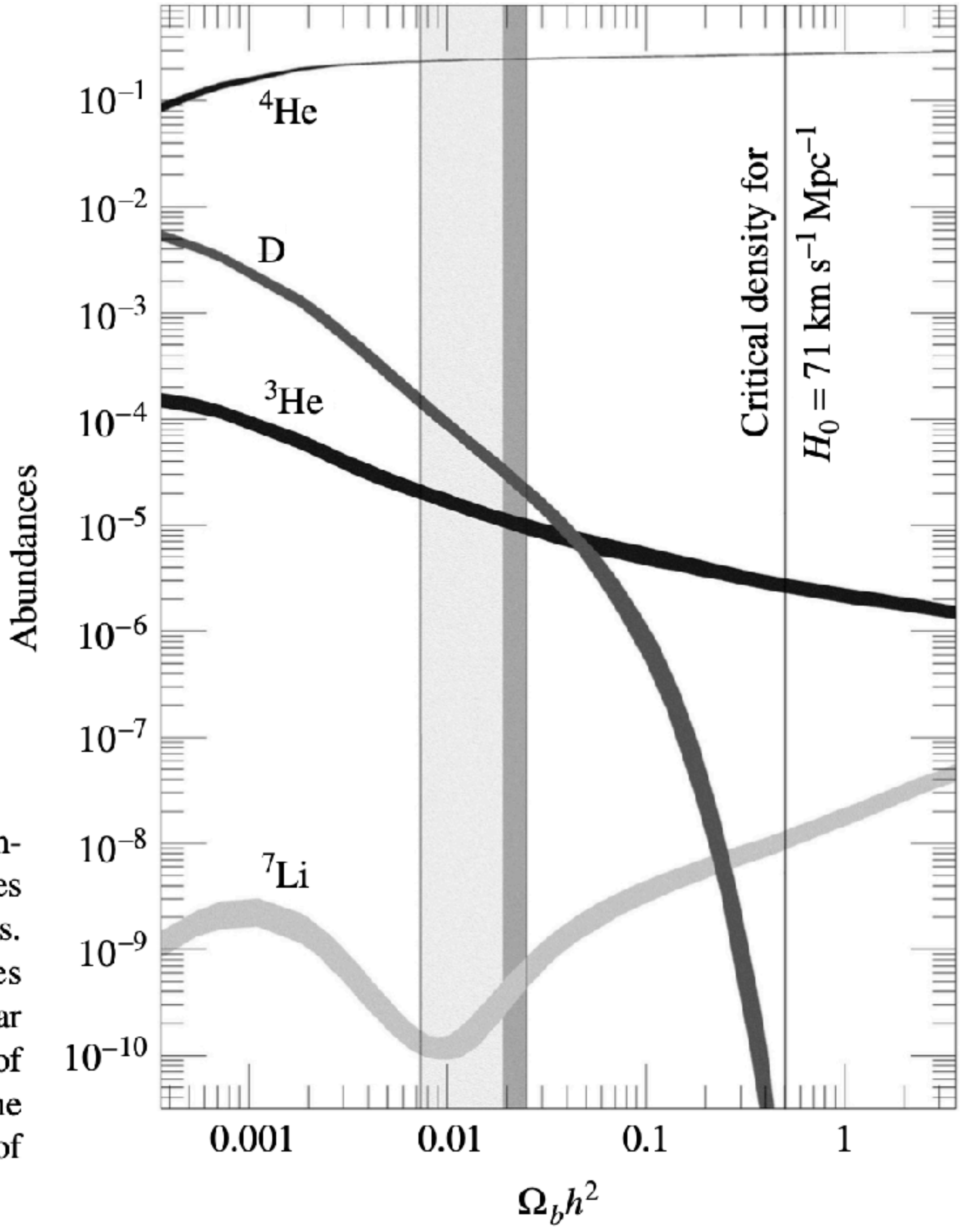


FIGURE 14 The calculated mass abundances of helium-4, deuterium, helium-3, and lithium-7 as a function of the present density of baryonic matter in the universe. The wide bar delineates the consistency interval, the range of values of $\Omega_{b,0}h^2$ that agree with the observed abundances. The narrow dark stripe at the right edge of the consistency interval corresponds to the abundances of primeval deuterium measured using the Lyman-α forest of absorption lines in high-z molecular clouds observed in front of quasars. The WMAP value of $\Omega_{b,0}h^2 = 0.0224$ runs down the center of the dark stripe, and the WMAP value of the critical density ($\Omega_{b,0}h^2 = 1h^2 = 0.504$) is shown at the right. Note that the agreement between the theoretical and observed abundances spans nine orders of magnitude. (Figure adapted from Schramm and Turner, *Rev. Mod. Phys.*, *70*, 303, 1998.)

The origin of the CMB

Now that we have described the nature of the universal expansion, let's return to the question of the origin of the CMB. **When we observe the cosmic microwave background, what are we actually viewing?**

The copious electrons in the hot environment of the very early universe obstructed the photons of the cosmic microwave background, allowing them to travel only relatively short distances before being scattered. **The scattering of photons by free electrons kept the electrons and photons in thermal equilibrium, meaning that they had the same temperature**.

However, as the **expansion of the universe** diluted the number density of free electrons, **the average time between scatterings** of a photon by an electron gradually **approached the characteristic timescale of the universal expansion,**

$$\tau_{\exp}(t) \equiv \left(\frac{1}{R(t)} \frac{dR(t)}{dt} \right)^{-1} = \frac{1}{H(t)}.$$

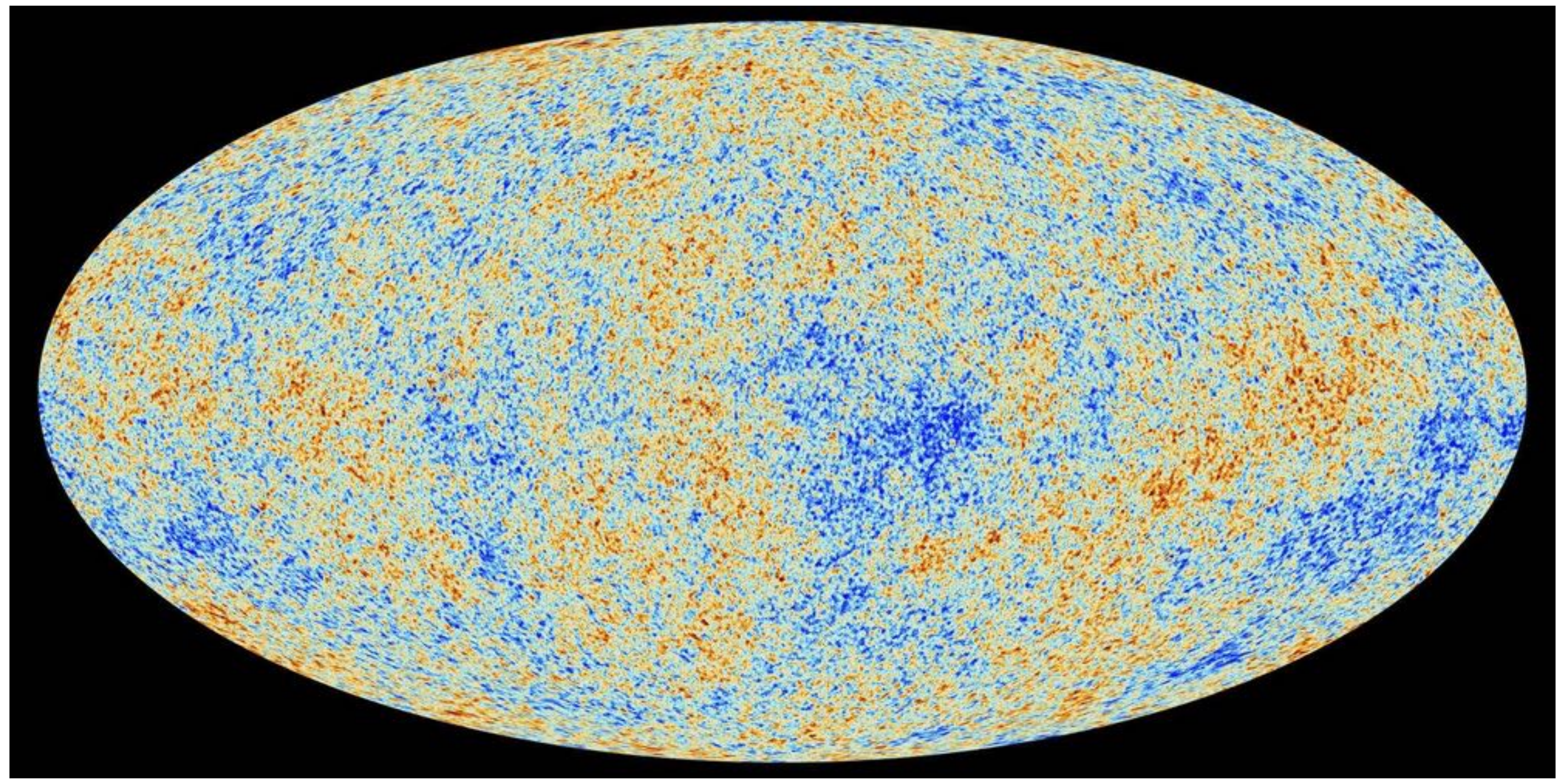
This expression is analogous to that of the pressure scale height. As the time of **decoupling** approached, the photons became increasingly disengaged from the electrons.

The origin of the CMB

If the electrons had remained free, decoupling would have occurred when the universe was about 20 million years old. However, when the universe was only some one million years old (10^{13} s), another event altered the opacity of the universe and rendered it transparent.

The independent evolution of radiation and matter began **when the temperature had cooled sufficiently to allow the free electrons to combine with nuclei of hydrogen and helium**. This formation of neutral atoms is sometimes referred to as **recombination.**

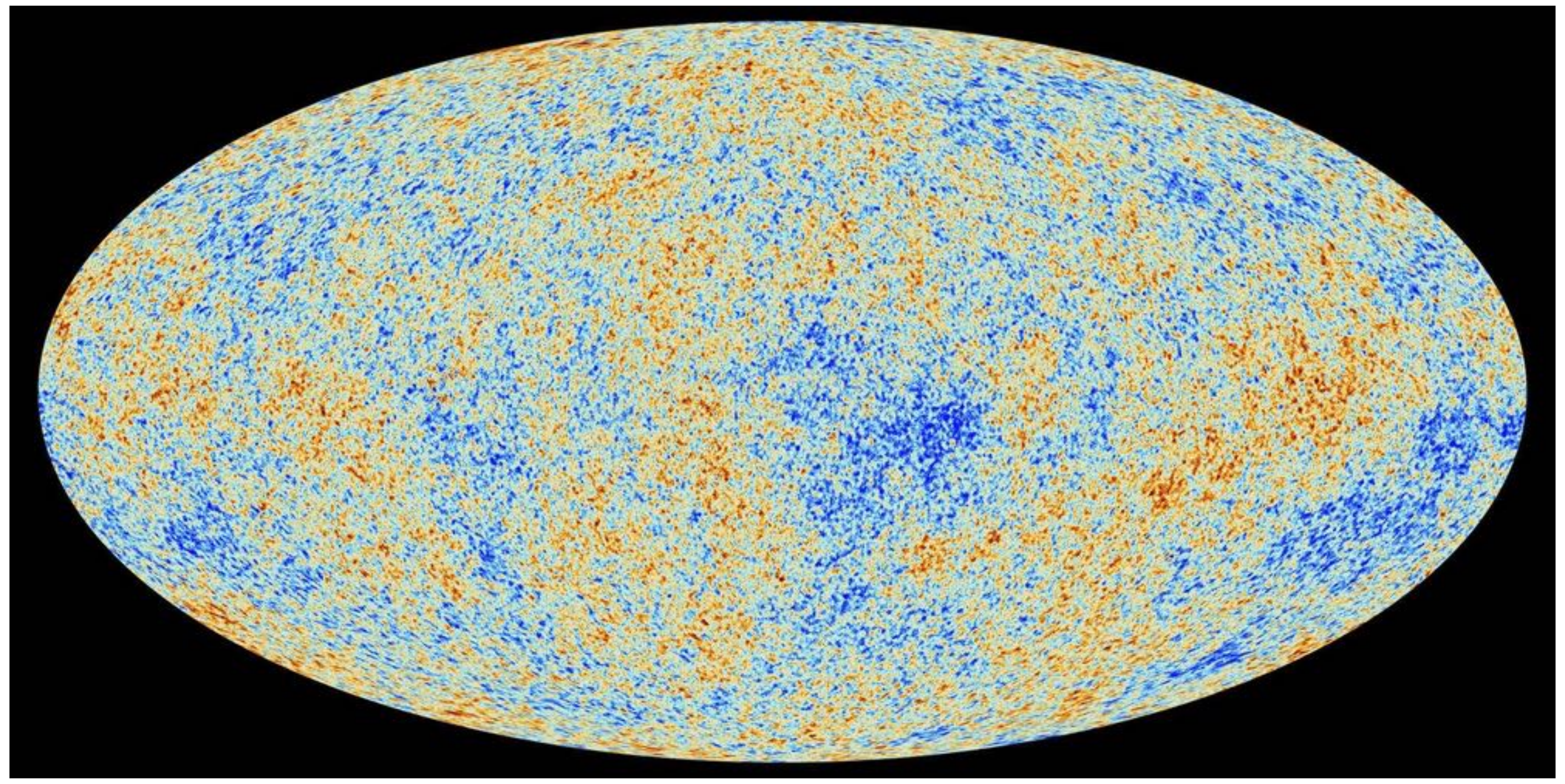
The loss of free electrons and the resulting drop in opacity completed the decoupling of radiation and matter, freeing the photons to roam unhindered throughout a newly transparent universe. **The photons of the cosmic microwave background that we observe today were last scattered during the time of recombination.**



The surface of last scattering

We define the surface of last scattering as a s**pherical surface, centered on the Earth, from which the CMB photons just now arriving at Earth were last scattered before beginning their unimpeded journey to us**.

The surface of last scattering **is the farthest redshift we can possibly observe at this moment in time**. More accurately, because recombination did not happen all at once, the surface of last scattering has a thickness Δz. Just as the light from the Sun was last scattered from somewhere within its photosphere, the **CMB photons originated within a layer, the "surface" of last scattering**. The surface of last scattering can therefore be thought of as a curtain that screens everything prior to decoupling from the direct view of astronomers. The earliest moments of the universe are hidden behind this veil and must be investigated with other methods than electromagnetic radiation (e.g. gravitational waves).



The conditions of recombination

The temperature at recombination can be estimated through use of the **Saha equation**,

$$\frac{N_{\mathrm{II}}}{N_{\mathrm{I}}} = \frac{2Z_{\mathrm{II}}}{n_e Z_{\mathrm{I}}} \left(\frac{2\pi m_e kT}{h^2} \right)^{3/2} e^{-\chi_{\mathrm{I}}/kT}.$$

Assuming (incorrectly) a composition of pure hydrogen for simplicity, we use $Z_{\mathrm{I}} = 2$ and $Z_{\mathrm{II}} = 1$. It is useful to define f to be the fraction of hydrogen atoms that are ionized, so

$$f = \frac{N_{\mathrm{II}}}{N_{\mathrm{I}} + N_{\mathrm{II}}} = \frac{N_{\mathrm{II}}/N_{\mathrm{I}}}{1 + N_{\mathrm{II}}/N_{\mathrm{I}}}, \qquad (97)$$

Or

$$\frac{N_{\mathrm{II}}}{N_{\mathrm{I}}} = \frac{f}{1 - f}. \qquad (98)$$

The conditions of recombination

For ionized hydrogen there is one free electron for every proton, $n_e = n_p$, so the number density of free electrons depends on *f* as

$$n_e = n_p = f(n_p + n_H) = \frac{f\rho_b}{m_H}, \qquad (99)$$

where ρ_b is the density of baryonic matter. Note that in obtaining Eq. (99) from Eq. (97), N_I corresponds to n_H, the number density of neutral hydrogen atoms, and N_{II} corresponds to n_p, the number density of protons (ionized hydrogen atoms). Using Eq. (82), we can write this as

$$n_e(R) = \frac{f\rho_{b,0}}{m_H R^3}, \qquad (100)$$

Substituting Eqs. (98) and (100) into the Saha equation, together with Eq. (58) for the blackbody temperature, we find

$$\frac{f}{1-f} = \frac{m_H R^3}{f\rho_{b,0}} \left(\frac{2\pi m_e k T_0}{h^2 R}\right)^{3/2} e^{-\chi_I R/kT_0}, \qquad (101)$$

where T0 = 2.725 K and χ_I = 13.6 eV.

The conditions of recombination

This can be solved numerically to find that the universe had cooled sufficiently for **one-half of its electrons and protons to combine to form atomic hydrogen (f = 0.5) when** the value of the scale factor was approximately $R \approx 7.25 \times 10^{-4}$ ($z \approx 1380$), corresponding to a temperature of about **3760 K**.

More precisely, the WMAP value for the redshift at the time of decoupling (i.e., the surface of last scattering) is

$$[z_{\mathrm{dec}}]_{\mathrm{WMAP}} = 1089 \pm 1.$$

We will adopt this as the value of the redshift for both recombination and decoupling. Using Eqs. (4) and (58) yields a temperature at recombination of

$$T_{\mathrm{dec}} = T_0(1 + z_{\mathrm{dec}}) = 2970\ \mathrm{K}.$$

This is lower than our estimate of 3760 K because the **photons created by the formation of some atoms were then absorbed by other atoms, putting these atoms into excited states from which they were easier to ionize.** Thus a slightly cooler temperature was needed to complete the recombination process.

The conditions of recombination

It is important to remember that at times earlier than recombination, the radiation and matter shared a common temperature, whereas **after recombination, the temperatures of the radiation and matter must be distinguished.** It is the radiation temperature (the temperature of the CMB) that will be of interest after recombination.

The WMAP value for the time at which recombination and decoupling occurred is

$$[t_{\mathrm{dec}}]_{\mathrm{WMAP}} = 379^{+8}_{-7} \text{ kyr.} \tag{102}$$

Of course, **these events did not occur at a single instant of time**; the WMAP value of the decoupling time interval is

$$[\Delta t_{\mathrm{dec}}]_{\mathrm{WMAP}} = 118^{+3}_{-2} \text{ kyr.}$$

This corresponds to the surface of last scattering having a thickness (in redshift) of

$$[\Delta z_{\mathrm{dec}}]_{\mathrm{WMAP}} = 195 \pm 2.$$