Carnegie Observatories Astrophysics Series, Vol. 2:
Measuring and Modeling the Universe
ed. W. L. Freedman (Cambridge: Cambridge Univ. Press)

Theoretical Overview of Cosmic Microwave Background Anisotropy

EDWARD L. WRIGHT
Department of Astronomy, University of California, Los Angeles

Abstract

The theoretical basis for the prediction of anisotropies in the cosmic microwave background is very well developed. Very low amplitude density and temperature perturbations produce small gravitational effects, leading to an anisotropy that is a combination of temperature fluctuations at the surface of last scattering and gravitational redshifts both at last scattering and along the path to the observer. All of the primary anisotropy can be handled by linear perturbation theory, which allows a very accurate calculation of the predicted anisotropy from different models of the Universe.

1.1 Introduction

The first predictions of the anisotropy of the cosmic microwave background (CMB) were published shortly after the CMB was discovered by Penzias & Wilson (1965). Sachs & Wolfe (1967) calculated the anisotropies due to gravitational potential fluctuations produced by density perturbations (Figure 1.1). Because the density perturbations are given by the second derivative of the gravitational potential fluctuation in Poisson's equation, the Sachs-Wolfe effect dominates the temperature fluctuations at large scales or low spherical harmonic index ℓ. Sachs & Wolfe predicted $\Delta T/T \approx 10^{-2}$ at large scales. This prediction, which failed by a factor of 10^3, is based on correct physics but incorrect input assumptions: prior to the discovery of the CMB no one knew how uniform the Universe was on large scales.

Silk (1968) computed the density perturbations needed at the recombination epoch at $z \approx 10^3$ in order to produce galaxies, and predicted $\Delta T/T \approx 3 \times 10^{-4}$ on arcminute scales. Silk (1967) calculated the damping of waves that were partially optically thick during recombination. This process, known as "Silk damping," greatly reduces the CMB anisotropy for small angular scales.

Observations by Conklin (1969) and then Henry (1971) showed that there was a dipole anisotropy in the CMB corresponding to the motion of the Solar System with respect to the average velocity of the observable Universe. There is a discussion of the dipole observations and their interpretation in Peebles (1971) that is still valid today, except that what was then a "tentative" dipole is now known to better than 1% accuracy, after a string of improved measurements starting with Corey & Wilkinson (1976) and ending with the *COBE* DMR (Bennett et al. 1996).

Peebles & Yu (1970) calculated the baryonic oscillations resulting from interactions be-

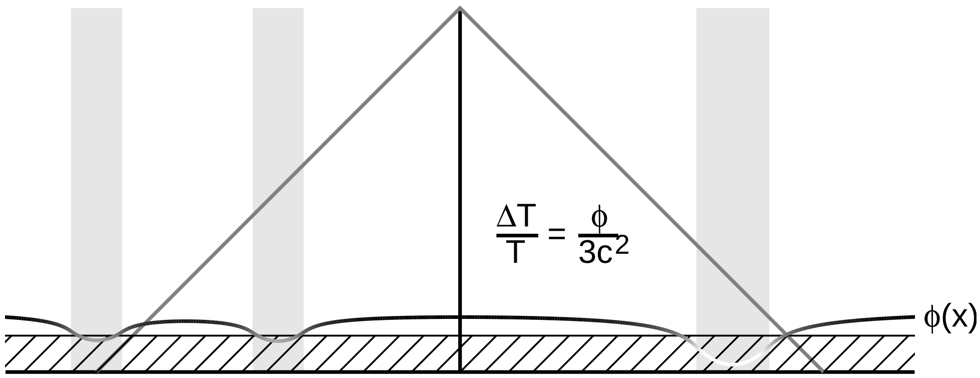

Fig. 1.1. Sachs & Wolfe (1967) predicted that density enhancements would be cold spots in the CMB, as shown in this conformal spacetime diagram.

tween photons and hydrogen in the early Universe, and also independently introduced the Harrison-Zel'dovich spectrum (Harrison 1970; Zel'dovich 1972). Pebbles & Yu predicted $\Delta T/T \approx 1.5\times10^{-4}$ on $1'$ scales and $\Delta T/T \approx 1.7\times10^{-3}$ on $7'$ scales.

Wilson & Silk (1981) further developed the theory of photon and matter interaction by scattering and gravity, and predicted $\Delta T/T = 100\ \mu$K for a single subtracted experiment with a 7° throw and with a 7° beam like *COBE*. Of course, when *COBE* was launched in 1989 it actually observed a much smaller anisotropy.

These early predictions of a large anisotropy were greatly modified by the addition of dark matter to the recipe for the cosmos. Observational upper limits on small-scale anisotropies had reached $\Delta T/T < 4\times10^{-5}$ on $1\farcm5$ scales (Uson & Wilkinson 1982), which was considerably less than the predictions from universes with just baryons and photons. Peebles (1982) computed the anisotropy expected in a universe "dominated by massive, weakly interacting particles" — in other words cold dark matter (CDM), although this paper predated the use of "cold dark matter."

Bond & Efstathiou (1987) calculated the correlation function of the CMB anisotropy, $C(\theta)$, and also the angular power spectrum, C_ℓ, in the CDM cosmology. This paper contains one of the first plots showing $\ell(\ell+1)C_\ell$ vs. ℓ, with peaks originally called the "Doppler" peaks but more properly called "acoustic peaks." This paper solved the Boltzmann equation describing the evolution of the photon distribution functions. Several authors developed these "Boltzmann codes," but the calculation of the angular power spectrum up to high ℓ was very slow. These codes described the conversion of inhomogeneity at the last-scattering surface into anisotropy on the observed sky by a set of differential equations evolving the coefficients of a Legendre polynomial expansion of the radiation intensity. Since the Universe is almost completely transparent after recombination, a ray-tracing approach is much more efficient. This great step in efficiency was implemented in the CMBFAST code by Seljak & Zaldarriaga (1996).

Hu & Dodelson (2002) give a recent review of CMB anisotropies, which includes a very good tutorial on the theory of $\Delta T/T$.

1.2 Results

A simple analysis of cosmological perturbations can be obtained from a consideration of the Newtonian approximation to a homogeneous and isotropic universe. Consider a test particle at radius R from an arbitrary center. Because the model is homogeneous the choice of center does not matter. The evolution of the velocity of the test particle is given by the energy equation

$$\frac{\upsilon^2}{2} = E_{tot} + \frac{GM}{R}. \tag{1.1}$$

If the total energy E_{tot} is positive, the Universe will expand forever since M, the mass (plus energy) enclosed within R, is positive, G is positive, and R is positive. In the absence of a cosmological constant or "dark energy," the expansion of the Universe will stop, leading to a recollapse if E_{tot} is negative. But this simple connection between E_{tot} and the fate of the Universe is broken in the presence of a vacuum energy density. The mass M is proportional to R^3 because the Universe is homogeneous and the Hubble velocity υ is given by $\upsilon = HR$. Thus $E_{tot} \propto R^2$.

We can find the total energy by plugging in the velocity $\upsilon_0 = H_0 R_0$ and the density ρ_0 in the Universe now. This gives

$$E_{tot} = \frac{(H_0 R_0 c)^2}{2} - \frac{4\pi G \rho_0 R_0^2}{3} = \frac{(H_0 R_0)^2}{2}\left(1 - \frac{\rho_0}{\rho_{crit}}\right), \tag{1.2}$$

with the critical density at time t_0 being $\rho_{crit} = 3H_0^2/(8\pi G)$. We define the ratio of density to critical density as $\Omega = \rho/\rho_{crit}$. This Ω includes all forms of matter and energy. Ω_m will be used to refer to the matter density.

From Equation 1.1 we can compute the time variation of Ω. Let

$$2E_{tot} = \upsilon^2 - \frac{2GM}{R} = H^2 R^2 - \frac{8\pi G \rho R^2}{3} = \text{const.} \tag{1.3}$$

If we divide this equation by $8\pi G \rho R^2/3$ we get

$$\frac{3H^2}{8\pi G \rho} - 1 = \frac{\text{const}'}{\rho R^2} = \Omega^{-1} - 1. \tag{1.4}$$

Thus $\Omega^{-1} - 1 \propto (\rho R^2)^{-1}$. When ρ declines with expansion at a rate faster than R^{-2} then the deviation of Ω from unity grows with expansion. This is the situation during the matter-dominated epoch with $\rho \propto R^{-3}$, so $\Omega^{-1} - 1 \propto R$. During the radiation-dominated epoch $\rho \propto R^{-4}$, so $\Omega^{-1} - 1 \propto R^2$. For Ω_0 to be within 0.9 and 1.1, Ω needed to be between 0.999 and 1.001 at the epoch of recombination, and within 10^{-15} of unity during nucleosynthesis. This fine-tuning problem is an aspect of the "flatness-oldness" problem in cosmology.

Inflation produces such a huge expansion that quantum fluctuations on the microscopic scale can grow to be larger than the observable Universe. These perturbations can be the seeds of structure formation and also will create the anisotropies seen by *COBE* for spherical harmonic indices $\ell \geq 2$. For perturbations that are larger than $\sim c_s t$ (or $\sim c_s/H$) we can ignore pressure gradients, since pressure gradients produce sound waves that are not able to cross the perturbation in a Hubble time. In the absence of pressure gradients, the density perturbation will evolve in the same way that a homogeneous universe does, and we can use the equation

$$\rho a^2 \left(\frac{1}{\Omega} - 1\right) = \text{const}, \tag{1.5}$$

the assumption that $\Omega \approx 1$ for early times, and $\Delta\rho \ll \rho$ as indicated by the smallness of the ΔT's seen by *COBE*, to derive

$$-\rho a^2 \left(\frac{1}{\Omega} - 1\right) \approx \rho_{crit} a^2 \Delta\Omega \approx \Delta\rho a^2 = \text{const}. \tag{1.6}$$

Hence,

$$\Delta\phi = \frac{G\Delta M}{R} = \frac{4\pi}{3} \frac{G\Delta\rho_0 (aL)^3}{aL} = \frac{1}{2} \frac{\Delta\rho_0}{\rho_{crit}} (H_0 L)^2, \tag{1.7}$$

where L is the comoving size of the perturbation. This is independent of the scale factor, so it does not change due to the expansion of the Universe.

During inflation (Guth 2003), the Universe is approximately in a steady state with constant H. Thus, the magnitude of $\Delta\phi$ for perturbations with physical scale c/H will be the same for all times during the inflationary epoch. But since this constant physical scale is aL, and the scale factor a changes by more than 30 orders of magnitude during inflation, this means that the magnitude of $\Delta\phi$ will be the same over 30 decades of comoving scale L. Thus, we get a strong prediction that $\Delta\phi$ will be the same on all observable scales from c/H_0 down to the scale that is no longer always larger than the sound speed horizon. This means that

$$\frac{\Delta\rho}{\rho} \propto L^{-2}, \tag{1.8}$$

so the Universe becomes extremely homogeneous on large scales even though it is quite inhomogeneous on small scales.

This behavior of $\Delta\phi$ being independent of scale is called *equal power on all scales.* It was originally predicted by Harrison (1970), Zel'dovich (1972), and Peebles & Yu (1970) based on a very simple argument: there is no scale length provided by the early Universe, and thus the perturbations should be *scale-free* — a power law. Therefore $\Delta\phi \propto L^m$. The gravitational potential divided by c^2 is a component of the metric, and if it gets comparable to unity then wild things happen. If $m < 0$ then $\Delta\phi$ gets large for small L, and many black holes would form. But we observe that this did not happen. Therefore $m \geq 0$. But if $m > 0$ then $\Delta\phi$ gets large on large scales, and the Universe would be grossly inhomogeneous. But we observe that this is not the case, so $m \leq 0$. Combining both results requires that $m = 0$, which is a *scale-invariant* perturbation power spectrum. This particular power-law power spectrum is called the Harrison-Zel'dovich spectrum. It was expected that the primordial perturbations should follow a Harrison-Zel'dovich spectrum because all other answers were wrong, but the inflationary scenario provides a good mechanism for producing a Harrison-Zel'dovich spectrum.

Sachs & Wolfe (1967) show that a gravitational potential perturbation produces an anisotropy of the CMB with magnitude

$$\frac{\Delta T}{T} = \frac{1}{3} \frac{\Delta\phi}{c^2}, \tag{1.9}$$

where $\Delta\phi$ is evaluated at the intersection of the line-of-sight and the surface of last scattering (or recombination at $z \approx 1100$). The $(1/3)$ factor arises because clocks run faster by a factor

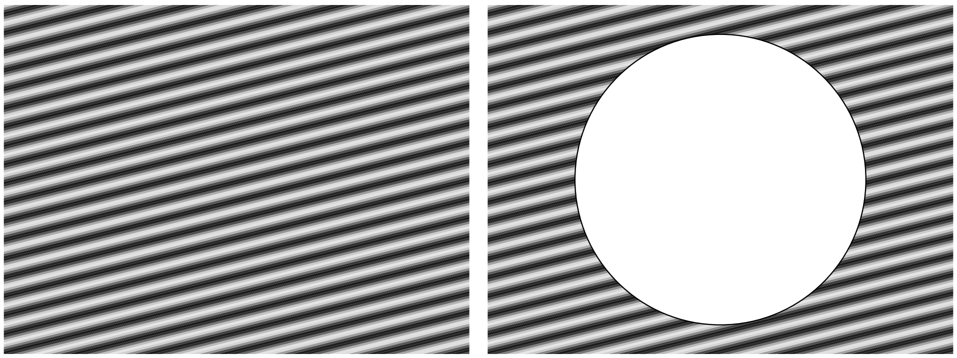

Fig. 1.2. *Left:* A plane wave on the last-scattering hypersurface. *Right:* The spherical intersection with our past light cone is shown.

$(1+\phi/c^2)$ in a gravitational potential, and we can consider the expansion of the Universe to be a clock. Since the scale factor is varying as $a \propto t^{2/3}$ at recombination, the faster expansion leads to a decreased temperature by $\Delta T/T = -(2/3)\Delta\phi/c^2$, which, when added to the normal gravitational redshift $\Delta T/T = \Delta\phi/c^2$ yields the (1/3) factor above. This is an illustration of the "gauge" problem in calculating perturbations in general relativity. The expected variation of the density contrast as the square of the scale factor for scales larger than the horizon in the radiation-dominated epoch is only obtained after allowance is made for the effect of the potential on the time. For a plane wave with wavenumber k we have $-k^2\Delta\phi = 4\pi G\Delta\rho$, or

$$\frac{\Delta\phi}{c^2} = -\frac{3}{2}(H/ck)^2\frac{\Delta\rho}{\rho_{crit}}, \tag{1.10}$$

so when $\rho \approx \rho_{crit}$ at recombination, the Sachs-Wolfe effect exceeds the physical temperature fluctuation $\Delta T/T = (1/4)\Delta\rho/\rho$ by a factor of $2(H/ck)^2$ if fluctuations are adiabatic (all component number densities varying by the same factor).

In addition to the physical temperature fluctuation and the gravitational potential fluctuation, there is a Doppler shift term. When the baryon fluid has a density contrast given by

$$\delta_b(x,t) = \frac{\Delta\rho_b}{\rho_b} = \delta_b \exp[ik(x-c_s t)], \tag{1.11}$$

where c_s is the sound speed, then

$$\frac{\partial\Delta\rho_b}{\partial t} = -ikc_s\delta_b\rho_b = -\rho_b\vec{\nabla}\cdot\vec{v} = ikv\rho_b. \tag{1.12}$$

As a result the velocity perturbation is given by $v = -c_s\delta_b$. But the sound speed is given by $c_s = \sqrt{\partial P/\partial\rho} = \sqrt{(4/3)\rho_\gamma c^2/(3\rho_b+4\rho_\gamma)} \approx c/\sqrt{3}$, since $\rho_\gamma > \rho_b$ at recombination ($z = 1100$). But the photon density is only slightly higher than the baryon density at recombination so the sound speed is about 20% smaller than $c/\sqrt{3}$. The Doppler shift term in the anisotropy is given by $\Delta T/T = v/c$, as expected. This results in $\Delta T/T$ slightly less than $\delta_b/\sqrt{3}$, which is nearly $\sqrt{3}$ larger than the physical temperature fluctuation given by $\Delta T/T = \delta_b/3$.

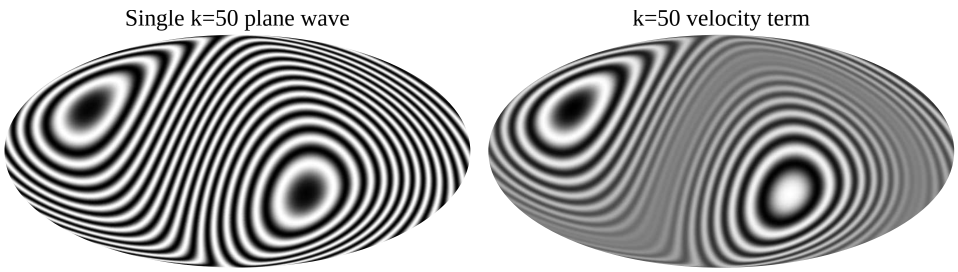

Fig. 1.3. *Left:* The scalar density and potential perturbation. *Right:* The vector velocity perturbation.

These plane-wave calculations need to be projected onto the sphere that is the intersection of our past light cone and the hypersurface corresponding to the time of recombination. Figure 1.2 shows a plane wave on these surfaces. The scalar density and potential perturbations produce a different pattern on the observed sky than the vector velocity perturbation. Figure 1.3 shows these patterns on the sky for a plane wave with $kR_{LS} = 50$, where R_{LS} is the radius of the last-scattering surface. The contribution of the velocity term is multiplied by $\cos\theta$, and since the RMS of this over the sphere is $\sqrt{1/3}$, the RMS contribution of the velocity term almost equals the RMS contribution from the density term since the speed of sound is almost $c/\sqrt{3}$.

The anisotropy is usually expanded in spherical harmonics:

$$\frac{\Delta T(\hat{n})}{T_0} = \sum_{\ell} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{n}). \tag{1.13}$$

Because the Universe is approximately isotropic the probability densities for all the different m's at a given ℓ are identical. Furthermore, the expected value of $\Delta T(\hat{n})$ is obviously zero, and thus the expected values of the $a_{\ell m}$'s is zero. But the variance of the $a_{\ell m}$'s is a measurable function of ℓ, defined as

$$C_\ell = \langle |a_{\ell m}|^2 \rangle. \tag{1.14}$$

Note that in this normalization C_ℓ and $a_{\ell m}$ are dimensionless. The harmonic index ℓ associated with an angular scale θ is given by $\ell \approx 180^\circ/\theta$, but the total number of spherical harmonics contributing to the anisotropy power at angular scale θ is given by $\Delta\ell \approx \ell$ times $2\ell+1$. Thus to have equal power on all scales one needs to have approximately $C_\ell \propto \ell^{-2}$. Given that the square of the angular momentum operator is actually $\ell(\ell+1)$, it is not surprising that the actual angular power spectrum of the CMB predicted by "equal power on all scales" is

$$C_\ell = \frac{4\pi\langle Q^2 \rangle}{5T_0^2} \frac{6}{\ell(\ell+1)}, \tag{1.15}$$

where $\langle Q^2 \rangle$ or Q^2_{rms-PS} is the expected variance of the $\ell = 2$ component of the sky, which must be divided by T_0^2 because the $a_{\ell m}$'s are defined to be dimensionless. The "4π" term arises because the mean of $|Y_{\ell m}|^2$ is $1/(4\pi)$, so the $|a_{\ell m}|^2$'s must be 4π times larger to compensate.

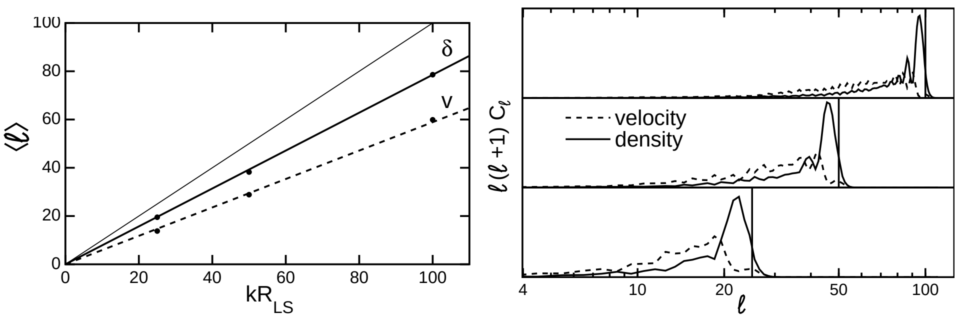

Fig. 1.4. *Left:* Mean ℓ is plotted vs. the wavenumber kR_{LS}. The light solid line shows $\ell = kR_{LS}$, while the solid line shows $\ell = (\pi/4)kR_{LS}$, and the dashed line shows $\ell = (3\pi/16)kR_{LS}$. *Right:* The angular power spectra for single k skies is plotted for $kR_{LS} = 25$, 50, and 100 (from bottom to top).

Finally, the quadrupole has 5 components, while C_ℓ is the variance of a single component, giving the "5" in the denominator. The *COBE* DMR experiment determined $\sqrt{\langle Q^2 \rangle} = 18\,\mu$K, and that the C_ℓ's from $\ell = 2$ to $\ell = 20$ were consistent with Equation 1.15.

The other common way of describing the anisotropy is in terms of

$$\Delta T_\ell^2 = \frac{T_0^2 \ell(\ell+1) C_\ell}{2\pi}. \tag{1.16}$$

Note these definitions give $\Delta T_2^2 = 2.4\langle Q^2 \rangle$. Therefore, the *COBE* normalized Harrison-Zel'dovich spectrum has $\Delta T_\ell^2 = 2.4 \times 18^2 = 778\,\mu\text{K}^2$ for $\ell \le 20$.

It is important to realize that the relationship between the wavenumber k and the spherical harmonic index ℓ is not a simple $\ell = kR_{LS}$. Figure 1.3 shows that while $\ell = kR_{LS}$ at the "equator" the poles have lower ℓ. In fact, if $\mu = \cos\theta$, where θ is the angle between the wave vector and the line-of-sight, then the "local ℓ" is given by $kR_{LS}\sqrt{1-\mu^2}$. The average of this over the sphere is $\langle \ell \rangle = (\pi/4)kR_{LS}$. For the velocity term the power goes to zero when $\mu = 0$ on the equator, so the average ℓ is smaller, $\langle \ell \rangle = (3\pi/16)kR_{LS}$, and the distribution of power over ℓ lacks the sharp cusp at $\ell = kR_{LS}$. As a result the velocity term, while contributing about 60% as much to the RMS anisotropy as the density term, does not contribute this much to the peak structure in the angular power spectrum. Thus the old nomenclature of "Doppler" peaks was not appropriate, and the new usage of "acoustic" peaks is more correct. Figure 1.4 shows the angular power spectrum from single k skies for both the density and velocity terms for several values of k, and a graph of the variance-weighted mean ℓ vs. kR_{LS}. These curves were computed numerically but have the expected forms given by the spherical Bessel function j_ℓ for the density term and j'_ℓ for the velocity term.

Seljak (1994) considered a simple model in which the photons and baryons are locked together before recombination, and completely noninteracting after recombination. Thus the opacity went from infinity to zero instantaneously. Prior to recombination there were two fluids, the photon-baryon fluid and the CDM fluid, which interacted only gravitationally. The baryon-photon fluid has a sound speed of about $c/\sqrt{3}$ while the dark matter fluid has a sound speed of zero. Figure 1.5 shows a conformal spacetime diagrams with a traveling wave in the baryon-photon fluid and the stationary wave in the CDM. The CDM dominates

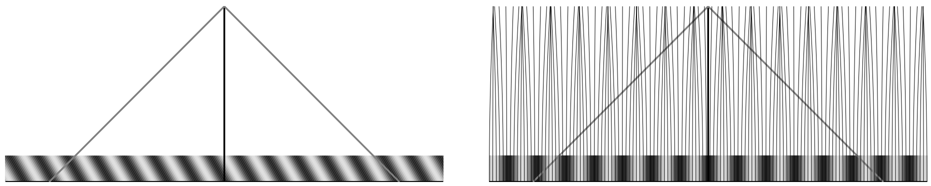

Fig. 1.5. On the left a conformal spacetime diagram showing a traveling wave in the baryon-photon fluid. On the right, the stationary CDM wave and the world lines of matter falling into the potential wells. For this wavenumber the density contrast in the baryon-photon fluid has undergone one-half cycle of its oscillation and is thus in phase with the Sachs-Wolfe effect from the CDM. This condition defines the first acoustic peak.

the potential, so the large-scale structure (LSS) forms in the potential wells defined by the CDM.

In Seljak's simple two-fluid model, there are five variables to follow: the density contrast in the CDM and baryons, δ_c and δ_b, the velocities of these fluids υ_c and υ_b, and the potential ϕ. The photon density contrast is $(4/3)\delta_b$. In Figure 1.6 the density contrasts are plotted vs. the scale factor for several values of the wavenumber. To make this plot the density contrasts were adjusted for the effect of the potential on the time, with

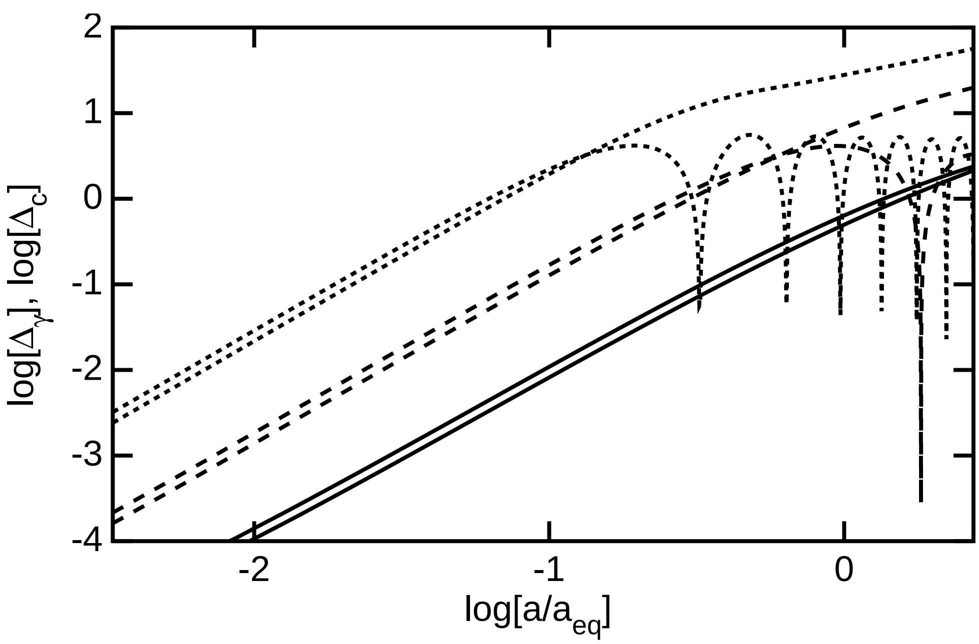

Fig. 1.6. Density contrasts in the CDM and the photons for wavenumbers κ = 5, 20, and 80 (see Fig. 1.7) as a function of the scale factor relative to the scale factor and when matter and radiation densities were equal. The photon density contrast starts out slightly larger than the CDM density contrast but oscillates.

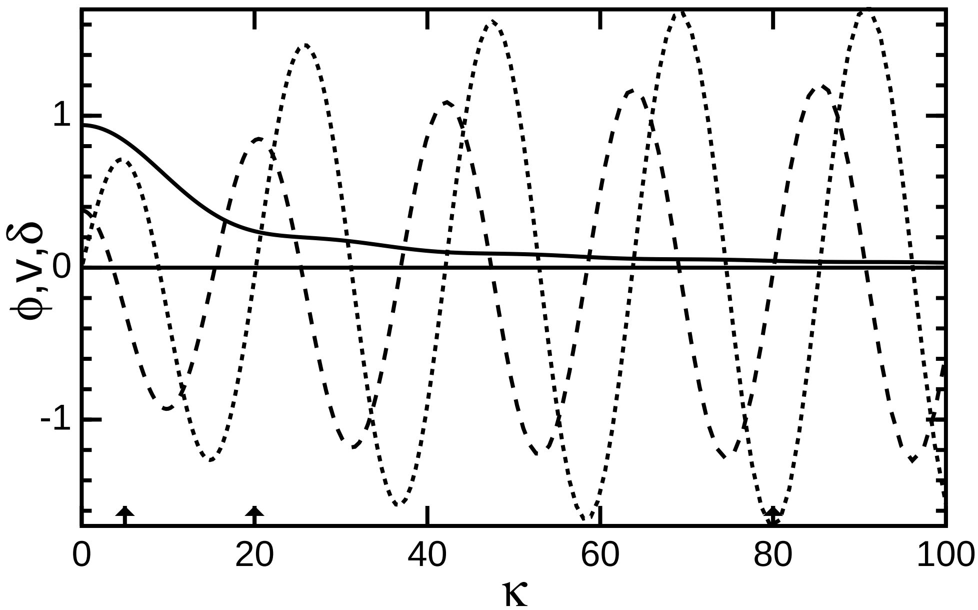

Fig. 1.7. Density contrasts at recombination as a function of wavenumber κ. The arrows on the x-axis indicate the values of κ for δ vs. a, as plotted in Figure 1.6. The solid curve shows the potential (the initial potential is always $\phi = 1$), the long dashed curve curve shows the combined potential plus density effect on the CMB temperature, while the short dashed curve shows the velocity of the baryon-photon fluid.

$$\Delta_c = \delta_c + 3H \int (\phi/c^2) dt \tag{1.17}$$

and

$$\Delta_\gamma = \delta_\gamma + 4H \int (\phi/c^2) dt. \tag{1.18}$$

Remembering that ϕ is negative when δ is positive, the two terms on the right-hand side of the above equations cancel almost entirely at early times, leaving a small residual growing like a^2 prior to a_{eq}, the scale factor when the matter density and the radiation density were equal. Thus these adjusted density contrasts evolve like $\Omega^{-1} - 1$ in homogeneous universes.

Figure 1.7 shows the potential that survives to recombination and produces LSS, the potential plus density effect on the CMB temperature, and the velocity of the baryons as function of wavenumber. Close scrutiny of the potential curve in the plot shows the baryonic wiggles in the LSS that may be detectable in the large redshift surveys by the 2dF and SDSS groups.

A careful examination of the angular power spectrum allows several cosmological parameters to be derived. The baryon to photon ratio and the dark matter to baryon density ratio can both be derived from the amplitudes of the first two acoustic peaks. Since the photon density is known precisely, the peak amplitudes determine the baryon density $\omega_b = \Omega_b h^2$ and the cold dark matter density $\omega_c = \Omega_{CDM} h^2$. The matter density is given by $\Omega_m = \Omega_b + \Omega_{CDM}$. The amplitude $\langle Q^2 \rangle$ and spectral index n of the primordial density perturbations are also

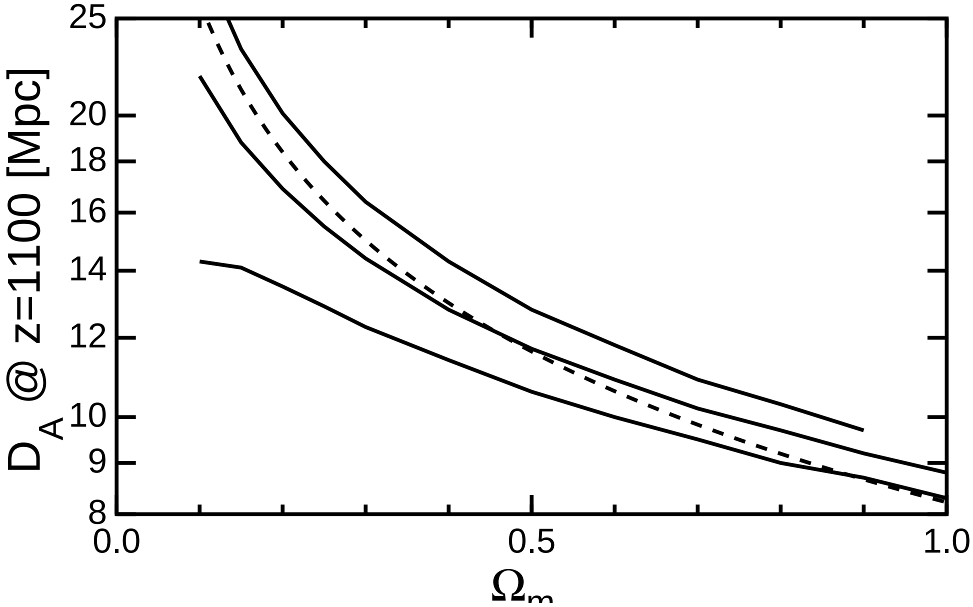

Fig. 1.8. The angular size distance vs. Ω_m for $H_0 = 60$ km s^{-1} Mpc^{-1} and three different values of Ω_{tot}(0.9, 1, and 1.1, from top to bottom). The dashed curve shows $\Omega_m^{-1/2}$.

easily observed. Finally the angular scale of the peaks depends on the ratio of the angular size distance at recombination to the distance sound can travel before recombination. Since the speed of sound is close to $c/\sqrt{3}$, this sound travel distance is primarily affected by the age of the Universe at $z = 1100$. The age of the Universe goes like $t \propto \rho^{-1/2} \propto \Omega_m^{-1/2} h^{-1}$. The angular size distance is proportional to h^{-1} as well, so the Hubble constant cancels out. The angular size distance is almost proportional to $\Omega_m^{-1/2}$, but this relation is not quite exact. Figure 1.8 compares the angular size distance to $\Omega_m^{-1/2}$. One sees that a peak position that corresponds to $\Omega_{tot} = 0.95$ if $\Omega_m = 0.2$ can also be fit by $\Omega_{tot} = 1.1$ if $\Omega_m = 1$. Thus, to first order the peak position is a good measure of Ω_{tot}.

The CMBFAST code by Seljak & Zaldarriaga (1996) provides the ability to quickly compute the angular power spectrum C_ℓ. Typically CMBFAST runs in about 1 minute for a given set of cosmological parameters. However, two different groups have developed even faster methods to evaluate C_ℓ. Kaplinghat, Knox, & Skordis (2002) have published the Davis Anisotropy Shortcut (DASh), with code available for download. This program interpolates among precomputed C_ℓ's. Kosowsky, Milosavljević, & Jiminez (2002) discuss combinations of the parameters that produce simple changes in the power spectrum, and also allow accurate and fast interpolation between C_ℓ's. These shortcuts allow the computation of a C_ℓ from model parameters in about 1 second. This allows the rapid computation of the likelihood of a given data set D for a set of model parameters M, $L(D|M)$. When computing the likelihood for high signal-to-noise ratio observations of a small area of the sky, biases due to the non-Gaussian shape of the likelihood are common. This can be avoided using the offset log-normal form for the likelihood $L(C_\ell)$ advocated by Bond, Jaffe, & Knox (2000).

Table 1.1. *Beam Size and Calibration Corrections*

Experiment	θ_B ($'$)	$100(\Delta\theta)/\theta$	$100(\Delta dT)/dT$
COBE	420.0	–0.3	...
ARCHEOPS	15.0	–0.2	2.7
BOOMERanG	12.9	10.5	–7.6
MAXIMA	10.0	–0.9	0.2
DASI	5.0	–0.4	0.7
VSA	3.0	–0.5	–1.7
CBI	1.5	–1.1	–0.2

The likelihood is a probability distribution over the data, so $\int L(D|M)dD = 1$ for any M. It is not a probability distribution over the models, so one should never attempt to evaluate $\int L dM$. For example, one could consider the likelihood as a function of the model parameters H_0 in km s^{-1} Mpc^{-1} and Ω_m for flat ΛCDM models, or one could use the parameters t_0 in seconds and Ω_m. For any (H_0, Ω_m) there is a corresponding (t_0, Ω_m) that makes exactly the same predictions, and therefore gives the same likelihood. But the integral of the likelihood over $dt_0 d\Omega_m$ will be much larger than the integral of the likelihood over $dH_0 d\Omega_m$ just because of the Jacobian of the transformation between the different parameter sets.

Wright (1994) gave the example of determining the primordial power spectrum power-law index n, $P(k) = A(k/k_0)^n$. Marginalizing over the amplitude by integrating the likelihood over A gives very different results for different values of k_0. Thus, it is very unfortunate that Hu & Dodelson (2002) still accept integration over the likelihood.

Instead of integrating over the likelihood one needs to define the *a posteriori* probability of the models $p_f(M)$ based on an *a priori* distribution $p_i(M)$ and Bayes' theorem:

$$p_f(M) \propto p_i(M)L(D|M). \tag{1.19}$$

It is allowable to integrate p_f over the space of models because the prior will transform when changing variables so as to keep the integral invariant.

In the modeling reported here, the *a priori* distribution is chosen to be uniform in ω_b, ω_c, n, Ω_V, Ω_{tot}, and z_{ri}. In doing the fits, the model C_ℓ's are adjusted by a factor of $\exp[a + b\ell(\ell+1)]$ before comparison with the data. Here a is a calibration adjustment, and b is a beam size correction that assumes a Gaussian beam. For *COBE*, a is the overall amplitude scaling parameter instead of a calibration correction. Marginalization over the calibration and beam size corrections for each experiment, and the overall spectral amplitude, is done by maximizing the likelihood, not by integrating the likelihood. Table 1.1 gives these beam and calibration corrections for each experiments. All of these corrections are less than the quoted uncertainties for these experiments. BOOMERanG stands out in the table for having honestly reported its uncertainties: $\pm 11\%$ for the beam size and $\pm 10\%$ for the gain. The likelihood is given by

$$-2\ln L = \chi^2 \quad = \quad \sum_j \Big\{ f(a_j/\sigma[a_j]) + f(b_j/\sigma[b_j])$$

Table 1.2. *Cosmic Parameters from pre-WMAP CMB Data only*

Parameter	Mean	σ	Units
ω_b	0.0206	0.0020	
ω_c	0.1343	0.0221	
Ω_V	0.3947	0.2083	
Ω_k	−0.0506	0.0560	
z_{ri}	7.58	3.97	
n	0.9409	0.0379	
H_0	51.78	12.26	km s^{-1} Mpc^{-1}
t_0	15.34	1.60	Gyr
Γ	0.2600	0.0498	

$$+ \sum_i f([Z^o_{ij} - Z^c_{ij}]/\sigma[Z_{ij}])\Big\}, \tag{1.20}$$

where j indexes over experiments, i indexes over points within each experiment, $Z = \ln(C_\ell + N_\ell)$ in the offset log normal approach of Bond et al. (2000), and N_ℓ is the noise bias. Since for *COBE* a is the overall normalization, $\sigma(a)$ is set to infinity for this term to eliminate it from the likelihood. The function $f(x)$ is x^2 for small $|x|$ but switches to $4(|x|-1)$ when $|x| > 2$. This downweighs outliers in the data. Most of the experiments have double tabulated their data. I have used both the even and odd points in my fits, but I have multiplied the σ's by $\sqrt{2}$ to compensate. Thus, I expect to get χ^2 per degree of freedom close to 0.5 but should have the correct sensitivity to cosmic parameters.

The scientific results such as the mean values and the covariance matrix of the parameters can be determined by integrations over parameter space weighted by p_f. Table 1.2 shows the mean and standard deviation of the parameters determined by integrating over the *a posteriori* probability distribution of the models. The evaluation of integrals over multi-dimensional spaces can require a large number of function evaluations when the dimensionality of the model space is large, so a Monte Carlo approach can be used. To achieve an accuracy of $\mathcal{O}(\epsilon)$ in a Monte Carlo integration requires $\mathcal{O}(\epsilon^{-2})$ function evaluations, while achieving the same accuracy with a gridding approach requires $\mathcal{O}(\epsilon^{-n/2})$ evaluations when second-order methods are applied on each axis. The Monte Carlo approach is more efficient for more than four dimensions. When the CMB data get better, the likelihood gets more and more sharply peaked as a function of the parameters, so a Gaussian approximation to $L(M)$ becomes more accurate, and concerns about banana-shaped confidence intervals and long tails in the likelihood are reduced. The Monte Carlo Markov Chain (MCMC) approach using the Metropolis-Hastings algorithm to generate models drawn from p_f is a relatively fast way to evaluate these integrals (Lewis & Bridle 2002). In the MCMC, a "trial" set of parameters is sampled from the proposal density $p_t(P';P)$, where P is the current location in parameter space, and P' is the new location. Then the trial location is accepted with a probability given by

$$\lambda = \frac{p_f(P')}{p_f(P)} \frac{p_t(P;P')}{p_t(P';P)}. \tag{1.21}$$

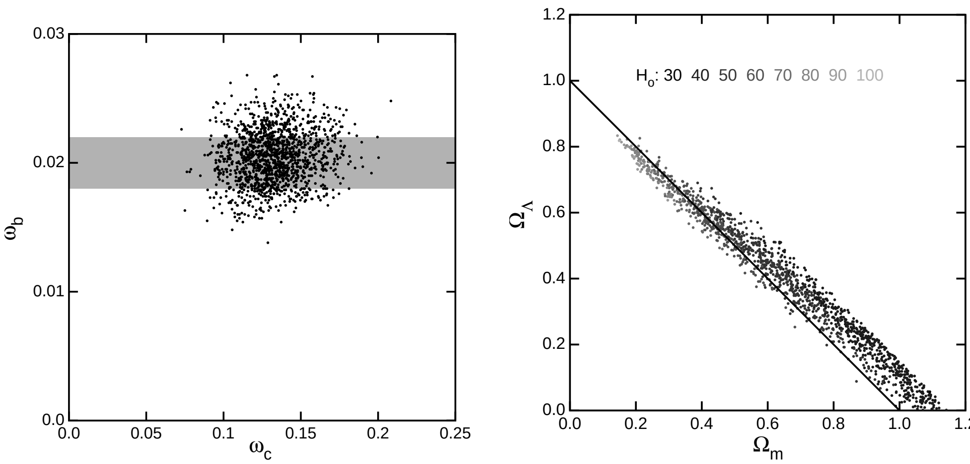

Fig. 1.9. *Left:* Clouds of models drawn from the *a posteriori* distribution based on the CMB data set as of 19 November 2002. The gray band shows the Big Bang nucleosynthesis determination of ω_b ($\pm 2\sigma$) from Burles, Nollett, & Turner (2001). *Right:* the same set of models in the Ω_m, Ω_Λ plane.

When a trial is accepted the Markov chain one sets $P = P'$. This algorithm guarantees that the accepted points in parameter space are sampled from the *a posteriori* probability distribution.

The most common choice for the proposal density is one that depends only on the parameter change $P' - P$. If the proposal density is a symmetric function then the ratio $p_t(P;P')/p_t P';P)$ $= 1$ and λ is then just the ratio of *a posteriori* probabilities. But the most efficient choice for the proposal density is $p_f(P)$ which is not a function of the parameter change, because this choice makes $\lambda = 1$ and all trials are accepted. However, if one knew how to sample models from p_f, why waste time calculating the likelihoods?

Just plotting the cloud of points from MCMC gives a useful indication of the allowable parameter ranges that are consistent with the data. I have done some MCMC calculations using the DASh (Kaplinghat et al. 2002) to find the C_ℓ's. I found DASh to be user unfriendly and too likely to terminate instead of reporting an error for out-of-bounds parameter sets, but it was fast. Figure 1.9 shows the range of baryon and CDM densities consistent with the CMB data set from *COBE* (Bennett et al. 1996), ARCHEOPS (Amblard 2003), BOOMERanG (Netterfield et al. 2002), MAXIMA (Lee et al. 2001), DASI (Halverson et al. 2002), VSA (Scott et al. 2003), and CBI (Pearson et al. 2003), and the range of matter and vacuum densities consistent with these data. The Hubble constant is strongly correlated with position on this diagram. Figure 1.10 shows the distribution of t_0 for models consistent with this pre-*WMAP* CMB data set. The relative uncertainty in t_0 is much smaller than the relative uncertainty in H_0 because the low-H_0 models have low vacuum energy density (Ω_V), and thus low values of the product $H_0 t_0$. The CMB data are giving a reasonable value for t_0 without using information on the distances or ages of objects, which is an interesting confirmation of the Big Bang model.

Peacock & Dodds (1994) define a shape parameter for the observed LSS power spectrum, $\Gamma = \Omega_m h \exp(-2\Omega_b)$. There are other slightly different definitions of Γ in use, but this will be used consistently here. Peacock & Dodds determine $\Gamma = 0.255 \pm 0.017 + 0.32(n^{-1} - 1)$. The CMB data specify n, so the slope correction in the last term is only 0.020 ± 0.013. Hence, the LSS power spectrum wants $\Gamma = 0.275 \pm 0.02$. The models based only on the pre-*WMAP* CMB data give the distribution in Γ shown in Figure 1.11, which is clearly consistent with the LSS data.

Two examples of flat ($\Omega_{tot} = 1$) models with equal power on all scales ($n = 1$), plotted on the pre-*WMAP* data set, are shown in Figure 1.12. Both these models are acceptable fits, but the ΛCDM model is somewhat favored based on the positions of the peaks. The rise in C_ℓ at low ℓ for the ΛCDM model is caused by the late integrated Sachs-Wolfe effect, which is due to the changing potential that occurs for $z < 1$ in this model. The potential changes because the density contrast stops growing when Λ dominates while the Universe continues to expand at an accelerating rate. The potential change during a photon's passage through a structure produces a temperature change given by $\Delta T/T = 2\Delta\phi/c^2$ (Fig. 1.13). The factor of 2 is the same factor of 2 that enters into the gravitational deflection of starlight by the Sun. The effect should be correlated with LSS that we can see at $z \approx 0.6$. Boughn & Crittenden (2003) have looked for this correlation using *COBE* maps compared to radio source count maps from the NVSS, and Boughn, Crittenden, & Koehrsen (2002) have looked at the correlation of *COBE* and the X-ray background. As of now the correlation has not been seen, which is an area of concern for ΛCDM, since the (non)correlation implies $\Omega_\Lambda = 0 \pm 0.33$ with roughly Gaussian errors. This correlation should arise primarily from redshifts near $z = 0.6$, as shown in Figure

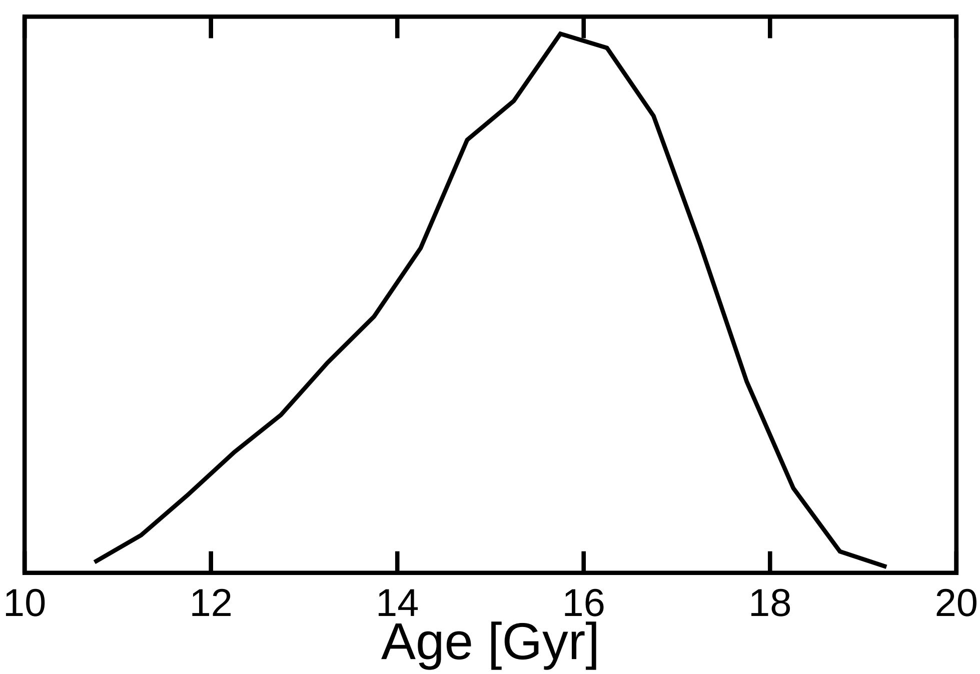

Fig. 1.10. Distribution of the age of the Universe based only on the pre-*WMAP* CMB data.

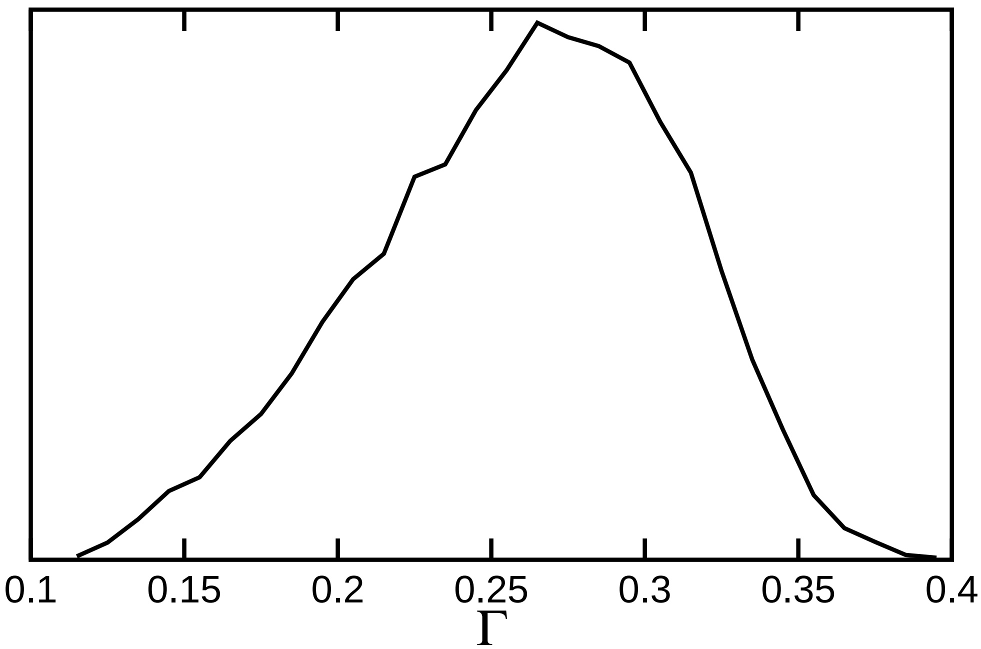

Fig. 1.11. Distribution of the LSS power spectrum shape parameter $\Gamma = \Omega_m h \exp(-2\Omega_b)$ from the pre-*WMAP* CMB data.

1.14. The coming availability of LSS maps based on deep all-sky infrared surveys (Maller 2003) should allow a better search for this correlation.

In addition to the late integrated Sachs-Wolfe effect from Λ, reionization should also enhance C_ℓ at low ℓ, as would an admixture of tensor waves. Since Λ, τ_{ri} and T/S all increase C_ℓ at low ℓ, and this increase is not seen, one has an upper limit on a weighted sum of all these parameters. If Λ is finally detected by the correlation between improved CMB and LSS maps, or if a substantial τ_{ri}, such as the $\tau = 0.1$ predicted by Cen (2003), is detected by the correlation between the E-mode polarization and the anisotropy (Zaldarriaga 2003), then one gets a greatly strengthened limit on tensor waves.

1.3 Discussion

The observed anisotropy of the CMB has an angular power spectrum that is in excellent agreement with the predictions of the ΛCDM model. But the CMB angular power spectrum is also consistent with an Einstein-de Sitter model having $\Omega_m = 1$ and a low value of $H_0 \approx 40$ km s^{-1} Mpc^{-1}. The observed lack of the expected correlation between the CMB and LSS due to the late integrated Sachs-Wolfe effect in ΛCDM slightly favors the $\Omega_m = 1$ “super Sandage” CDM model (sSCDM), which, like ΛCDM, is also consistent with the shape of the matter power spectrum $P(k)$ and the baryon fraction in clusters of galaxies. But sSCDM disagrees with the actual measurements of H_0 and with the supernova data for an accelerating Universe. Thus, ΛCDM is the overall best fit, but further efforts to confirm the CMB-LSS correlation should be encouraged.

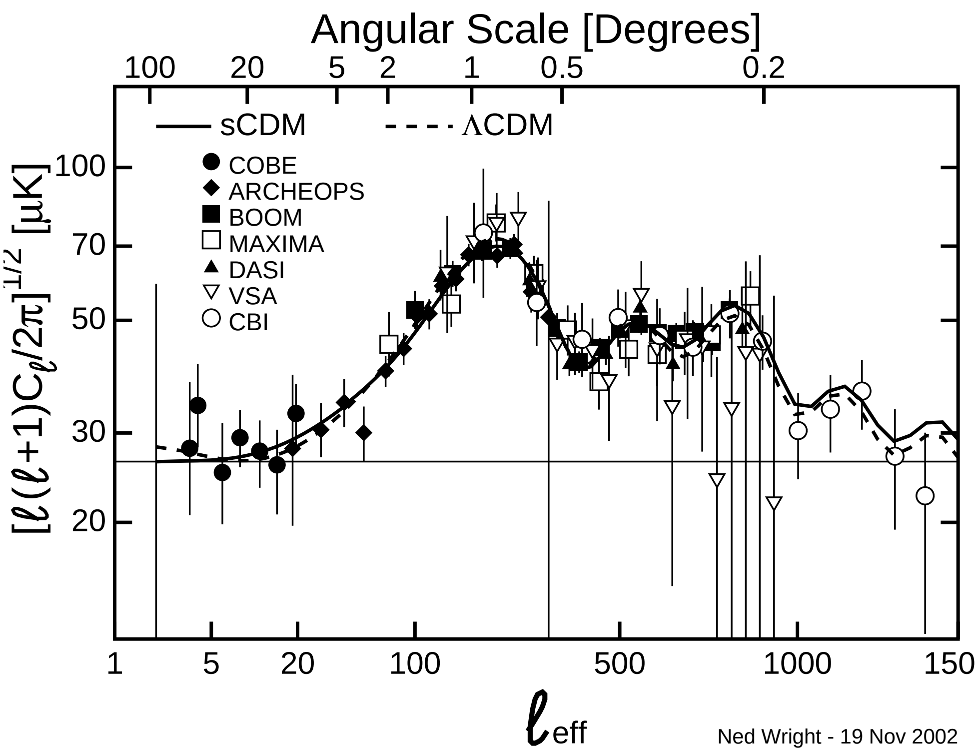

Fig. 1.12. Two flat $n = 1$ models. One shows ΛCDM with $\Omega_\Lambda = 2/3$. The best fit gives $\omega_b = 0.022$ and $\omega_c = 0.132$, implying $H_0 = 68$ km s^{-1} Mpc^{-1}. The other fit shows $\Omega_\Lambda = 0$ with $\omega_b = 0.021$ and $\omega_c = 0.196$, implying $H_0 = 47$ km s^{-1} Mpc^{-1}.

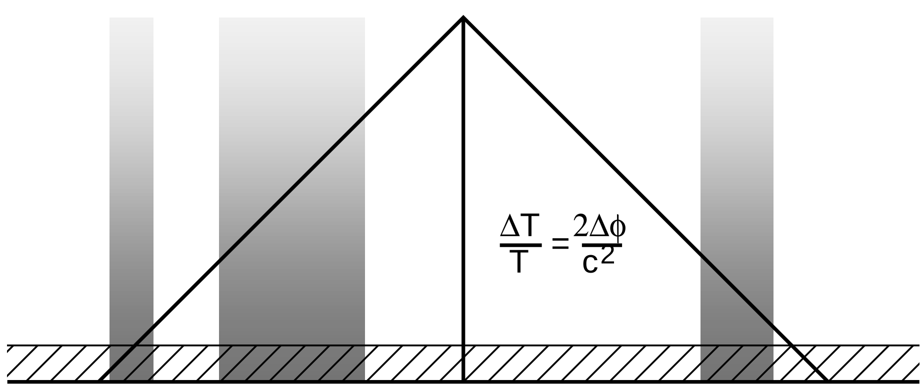

Fig. 1.13. Fading potentials cause large-scale anisotropy correlated with LSS due to the late integrated Sachs-Wolfe effect.

1.4 Conclusions

The pre-*WMAP* CMB angular power spectrum assembled from multiple experiments is very well fit by a six-parameter model. Of these six parameters, the vacuum energy density Ω_V and the redshift of reionization z_{ri} are still poorly determined from CMB data alone. However, the well-determined parameters either match independent determinations or the expectations from inflation:

- The baryon density ω_b is determined to 10% and agrees with the value from Big Bang nucleosynthesis.
- The age of the Universe is determined to 11% and agrees with determinations from white dwarf cooling (Rich 2003), main sequence turnoffs, and radioactive decay.
- The predicted shape of the LSS power spectrum $P(k)$ agrees with the observed shape.
- The curvature Ω_k is determined to 4% and agrees with the expected value from inflation.
- The spectral index n is determined to 4% and agrees with the expected value from inflation.

The angular power spectrum of the CMB can be computed using well-understood physics and linear perturbation theory. The current data set agrees with the predictions of inflation happening less than 1 picosecond after the Big Bang, the observations of light isotope abundances from the first three minutes after the Big Bang, and the observations of LSS in the current Universe. The inflationary scenario and the hot Big Bang model appear to be solidly based on confirmed quantitative predictions.

The greatly improved CMB data expected from *WMAP*, and later *Planck*, should dramatically improve our knowledge of the Universe.

References

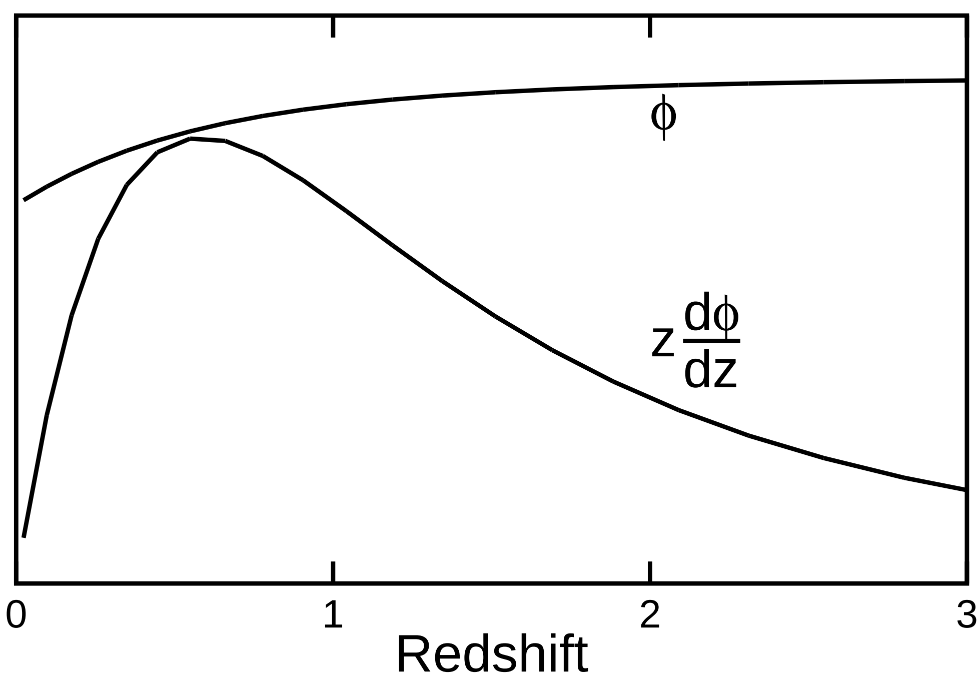

Fig. 1.14. Change of potential vs. redshift in a ΛCDM model. Note that the most significant changes occur near $z = 0.6$.

Amblard, A. 2003, Carnegie Observatories Astrophysics Series, Vol. 2: Measuring and Modeling the Universe, ed. W. L. Freedman (Pasadena: Carnegie Observatories, http://www.ociw.edu/ociw/symposia/series/symposium2/proceedings.html)
Bennett, C. L., et al. 1996, ApJL, 464, L1
Bond, J. R., & Efstathiou, G. 1987, MNRAS, 226, 655
Bond, J. R., Jaffe, A. H., & Knox, L. 2000, ApJ, 533, 19
Boughn, S. P., & Crittenden, R. G. 2003, Phys. Rev. Lett., submitted (astro-ph/0111281)
Boughn, S. P., Crittenden, R. G., & Koehrsen, G. P. 2002, ApJ, 580, 672
Burles, S., Nollett, K. M., & Turner, M. S. 2001, Phys. Rev. D., 63, 063512
Cen, R. 2003, ApJ, in press (astro-ph/02010473)
Conklin, E. K. 1969, Nature, 222, 971
Corey, B. E., & Wilkinson, D. T. 1976, BAAS, 8, 351
Guth, A. H. 2003, Carnegie Observatories Astrophysics Series, Vol. 2: Measuring and Modeling the Universe, ed. W. L. Freedman (Cambridge: Cambridge Univ. Press)
Halverson, N. W., et al. 2002, ApJ, 568, 38
Harrison, E. R. 1970, Phys Rev D, 1, 2726
Henry, P. S. 1971, Nature, 231, 516
Hu, W., & Dodelson, S. 2002, ARA&A, 40, 171
Kaplinghat, M., Knox, L., & Skordis, C. 2002, ApJ, 575, 665
Kosowsky, A., Milosavljević, M., & Jiminez, R. 2002, Phys. Rev. D, 66, 063007
Lee, A. T., et al. 2001, ApJ, 561, L1
Lewis, A., & Bridle, S. 2002, Phys. Rev. D, 66, 103511
Maller, A. 2003, Carnegie Observatories Astrophysics Series, Vol. 2: Measuring and Modeling the Universe, ed. W. L. Freedman (Pasadena: Carnegie Observatories, http://www.ociw.edu/ociw/symposia/series/symposium2/proceedings.html)
Netterfield, C. B., et al. 2002, ApJ, 571, 604
Peacock, J. A., & Dodds, S. J. 1994, MNRAS, 267, 1020
Pearson, T. J., et al. 2003, ApJ, in press (astro-ph/0205388)
Peebles, P. J. E. 1971, Physical Cosmology (Princeton: Princeton Univ. Press)
——. 1982, ApJ, 263, L1
Peebles, P. J. E., & Yu, J. T. 1970, ApJ, 162, 815
Penzias, A. A., & Wilson, R. W. 1965, ApJ, 142, 419
Rich, M. 2003, Carnegie Observatories Astrophysics Series, Vol. 2: Measuring and Modeling the Universe, ed. W. L. Freedman (Pasadena: Carnegie Observatories, http://www.ociw.edu/ociw/symposia/series/symposium2/proceedings.html)
Sachs, R. K., & Wolfe, A. M. 1967, ApJ, 147, 73
Scott, P. F., et al. 2003, MNRAS, in press (astro-ph/0205380)
Seljak, U. 1994, ApJ, 435, L87
Seljak, U., & Zaldarriaga, M. 1996, ApJ, 469, 437
Silk, J. 1967, Nature, 215, 1155
——. 1968, ApJ, 151, 459
Uson, J. M., & Wilkinson, D. T. 1982, Phys. Rev. Lett., 49, 1463
Wilson, M. L., & Silk, J. 1981, ApJ, 243, 14
Wright. E. L. 1994, in The Proceedings of the 1994 CWRU Workshop on CMB Anisotropies Two Years After Cobe: Observations, Theory and the Future, ed. L. M. Krauss (World Scientific: Singapore), 21
Zaldarriaga, M. 2003, Carnegie Observatories Astrophysics Series, Vol. 2: Measuring and Modeling the Universe, ed. W. L. Freedman (Cambridge: Cambridge Univ. Press)
Zel'dovich, Ya. B. 1972, MNRAS, 160, 1P