Annu. Rev. Astron. Astrophys. 2002. 40:171–216
doi: 10.1146/annurev.astro.40.060401.093926

COSMIC MICROWAVE BACKGROUND ANISOTROPIES

Wayne Hu[1,2,3] and Scott Dodelson[2,3]
[1]*Center for Cosmological Physics, University of Chicago, Chicago, Illinois 60637; email: whu@background.uchicago.edu*
[2]*NASA/Fermilab Astrophysics Center, P.O. Box 500, Batavia, Illinois 60510; email: dodelson@fnal.gov*
[3]*Department of Astronomy and Astrophysics, University of Chicago, Chicago, Illinois 60637*

Key Words background radiation, cosmology, theory, dark matter, early universe

■ **Abstract** Cosmic microwave background (CMB) temperature anisotropies have and will continue to revolutionize our understanding of cosmology. The recent discovery of the previously predicted acoustic peaks in the power spectrum has established a working cosmological model: a critical density universe consisting of mainly dark matter and dark energy, which formed its structure through gravitational instability from quantum fluctuations during an inflationary epoch. Future observations should test this model and measure its key cosmological parameters with unprecedented precision. The phenomenology and cosmological implications of the acoustic peaks are developed in detail. Beyond the peaks, the yet to be detected secondary anisotropies and polarization present opportunities to study the physics of inflation and the dark energy. The analysis techniques devised to extract cosmological information from voluminous CMB data sets are outlined, given their increasing importance in experimental cosmology as a whole.

INTRODUCTION

The field of cosmic microwave background (CMB) anisotropies has dramatically advanced over the past decade (see White et al. 1994), especially on its observational front. The observations have turned some of our boldest speculations about our Universe into a working cosmological model: namely, that the Universe is spatially flat, consists mainly of dark matter and dark energy, with the small amount of ordinary matter necessary to explain the light element abundances, and all the rich structure in it formed through gravitational instability from quantum mechanical fluctuations when the Universe was a fraction of a second old. Observations over the coming decade should pin down certain key cosmological parameters with unprecedented accuracy (Knox 1995, Jungman et al. 1996, Bond et al. 1997, Zaldarriaga et al. 1997, Eisenstein et al. 1999). These determinations will

have profound implications for astrophysics, as well as other disciplines. Particle physicists, for example, will be able to study neutrino masses, theories of inflation impossible to test at accelerators, and the mysterious dark energy or cosmological constant.

For the 28 years between the discovery of the CMB (Penzias & Wilson 1965) and the COBE DMR detection of 10^{-5} fluctuations in its temperature field across the sky (Smoot et al. 1992), observers searched for these anisotropies but found none except the dipole induced by our own motion (Smoot et al., 1977). They learned the hard way that the CMB is remarkably uniform. This uniformity is in stark contrast to the matter in the Universe, organized in very nonlinear structures like galaxies and clusters. The disparity between the smooth photon distribution and the clumpy matter distribution is due to radiation pressure. Matter inhomogeneities grow owing to gravitational instability, but pressure prevents the same process from occuring in the photons. Thus, even though both inhomogeneities in the matter in the Universe and anisotropies in the CMB apparently originated from the same source, these appear very different today.

Because the photon distribution is very uniform, perturbations are small, and linear response theory applies. This is perhaps the most important fact about CMB anisotropies. Because they are linear, predictions can be made as precisely as their sources are specified. If the sources of the anisotropies are also linear fluctuations, anisotropy formation falls in the domain of linear perturbation theory. There are then essentially no phenomenological parameters that need to be introduced to account for nonlinearities or gas dynamics or any other of a host of astrophysical processes that typically afflict cosmological observations.

CMB anisotropies in the working cosmological model, which we briefly review in "Observables," fall almost entirely under linear perturbation theory. The most important observables of the CMB are the power spectra of the temperature and polarization maps. Theory predicted, and now observations confirm, that the temperature power spectrum has a series of peaks and troughs. In "Acoustic Peaks," we discuss the origin of these acoustic peaks and their cosmological uses. Although they are the most prominent features in the spectrum and are the focus of the current generation of experiments, future observations will turn to even finer details, potentially revealing the physics at the two opposite ends of time. Some of these are discussed in "Beyond the Peaks." Finally, the past few years have witnessed important new advances, introduced in "Data Analysis," from a growing body of CMB data analysts on how best to extract the information contained in CMB data. Some of the fruits of this labor have already spread to other fields of astronomy.

OBSERVABLES

Standard Cosmological Paradigm

Whereas a review of the standard cosmological paradigm is not our intention (see Narkilar & Padmanabhan 2001 for a critical appraisal), we briefly introduce the observables necessary to parameterize it.

The expansion of the Universe is described by the scale factor $a(t)$, set to unity today, and by the current expansion rate, the Hubble constant $H_0 = 100\,h$ km sec^{-1} Mpc^{-1}, with $h \simeq 0.7$ (Freedman et al. 2001; compare Theureau et al. 1997, Sandage et al. 2000). The Universe is flat (no spatial curvature) if the total density is equal to the critical density, $\rho_c = 1.88\,h^2 \times 10^{-29}$ g cm^{-3}; it is open (negative curvature) if the density is less than this and closed (positive curvature) if greater. The mean densities of different components of the Universe control $a(t)$ and are typically expressed today in units of the critical density Ω_i, with an evolution with a specified by equations of state $w_i = p_i/\rho_i$, where p_i is the pressure of the ith component. Density fluctuations are determined by these parameters through the gravitational instability of an initial spectrum of fluctuations.

The working cosmological model contains photons, neutrinos, baryons, cold dark matter, and dark energy with densities proscribed within a relatively tight range. For the radiation, $\Omega_r = 4.17 \times 10^{-5}\,h^{-2}$ ($w_r = 1/3$). The photon contribution to the radiation is determined to high precision by the measured CMB temperature, $T = 2.728 \pm 0.004$ K (Fixsen et al. 1996). The neutrino contribution follows from the assumption of three neutrino species, a standard thermal history, and a negligible mass $m_\nu \ll 1$ eV. Massive neutrinos have an equation of state $w_\nu = 1/3 \to 0$ as the particles become nonrelativistic. For $m_\nu \sim 1$ eV this occurs at $a \sim 10^{-3}$ and can leave a small but potentially measurable effect on the CMB anisotropies (Ma & Bertschinger 1995, Dodelson et al. 1996).

For the ordinary matter or baryons, $\Omega_b \approx 0.02\,h^{-2}$ ($w_b \approx 0$), with statistical uncertainties at about the 10% level determined through studies of the light element abundances (for reviews, see Boesgaard & Steigman 1985, Schramm & Turner 1998, Tytler et al. 2000). This value is in strikingly good agreement with that implied by the CMB anisotropies themselves, as we shall see. There is very strong evidence that there is also substantial nonbaryonic dark matter. This dark matter must be close to cold ($w_m = 0$) for the gravitational instability paradigm to work (Peebles 1982), and when added to the baryons gives a total in nonrelativistic matter of $\Omega_m \simeq 1/3$. Because the Universe appears to be flat, the total $\Omega_{\rm tot}$ must equal one. Thus, there is a missing component to the inventory, dubbed *dark energy*, with $\Omega_\Lambda \simeq 2/3$. The cosmological constant ($w_\Lambda = -1$) is only one of several possible candidates, but we generally assume this form unless otherwise specified. Measurements of an accelerated expansion from distant supernovae (Riess et al. 1998, Perlmutter et al. 1999) provide entirely independent evidence for dark energy in this amount.

The initial spectrum of density perturbations is assumed to be a power law with a power law index or tilt of $n \approx 1$ corresponding to a scale-invariant spectrum. Likewise the initial spectrum of gravitational waves is assumed to be scale-invariant, with an amplitude parameterized by the energy scale of inflation E_i and compared with the initial density spectrum (see "Gravitational Waves"). Finally the formation of structure will eventually reionize the Universe at some redshift $6 \lesssim z_{\rm ri} \lesssim 20$.

Many of the features of the anisotropies will be produced even if these parameters fall outside the expected range or even if the standard paradigm is incorrect. Where appropriate, we try to point these out.

Cosmic Microwave Background Temperature Field

The basic observable of the CMB is its intensity as a function of frequency and direction on the sky $\hat{\mathbf{n}}$. Because the CMB spectrum is an extremely good blackbody (Fixsen et al. 1996) with a nearly constant temperature across the sky T, we generally describe this observable in terms of a temperature fluctuation $\Theta(\hat{\mathbf{n}}) = \Delta T/T$.

If these fluctuations are Gaussian, then the multipole moments of the temperature field

$$\Theta_{\ell m} = \int d\hat{\mathbf{n}} Y^*_{\ell m}(\hat{\mathbf{n}})\Theta(\hat{\mathbf{n}}) \tag{1}$$

are fully characterized by their power spectrum

$$\left\langle \Theta^*_{\ell m}\Theta_{\ell' m'}\right\rangle = \delta_{\ell\ell'}\delta_{mm'}C_\ell, \tag{2}$$

whose values as a function of ℓ are independent in a given realization. For this reason predictions and analyses are typically performed in harmonic space. On small sections of the sky where its curvature can be neglected, the spherical harmonic analysis becomes ordinary Fourier analysis in two dimensions. In this limit ℓ becomes the Fourier wavenumber. Because the angular wavelength $\theta = 2\pi/\ell$, large multipole moments corresponds to small angular scales with $\ell \sim 10^2$ representing degree scale separations. Likewise, because in this limit the variance of the field is $\int d^2\ell C_\ell/(2\pi)^2$, the power spectrum is usually displayed as

$$\Delta^2_T \equiv \frac{\ell(\ell+1)}{2\pi}C_\ell T^2, \tag{3}$$

the power per logarithmic interval in wavenumber for $\ell \gg 1$.

Figure 1 (*top*) shows observations of Δ_T along with the prediction of the working cosmological model, complete with the acoustic peaks mentioned in the "Introduction" and discussed extensively in "Acoustic Peaks," below. Whereas COBE first detected anisotropy on the largest scales (*inset*), observations in the past decade have pushed the frontier to smaller and smaller scales (*left to right* in the figure). The Microwave Anisotropy Probe (MAP) satellite, launched in June 2001, will go out to $\ell \sim 1000$, whereas the European satellite, Planck, scheduled for launch in 2007, will go a factor or two higher (see Figure 1, *bottom*).

The power spectra shown in Figure 1 all begin at $\ell = 2$ and exhibit large errors at low multipoles. This is because the predicted power spectrum is the average power in the multipole moment ℓ an observer would see in an ensemble of universes. However, a real observer is limited to one universe and one sky with its one set of $\Theta_{\ell m}$'s, $2\ell + 1$ numbers for each ℓ. This is particularly problematic for the monopole and dipole ($\ell = 0, 1$). If the monopole were larger in our vicinity than its average value, we would have no way of knowing it. Likewise for the dipole, we have no way of distinguishing a cosmological dipole from our own peculiar motion with respect to the CMB rest frame. Nonetheless, the monopole and

dipole—which we will often call simply Θ and v_γ—are of the utmost significance in the early Universe. It is precisely the spatial and temporal variation of these quantities, especially the monopole, which determine the pattern of anisotropies we observe today. A distant observer sees spatial variations in the local temperature or monopole, at a distance given by the lookback time, as a fine-scale angular anisotropy. Similarly, local dipoles appear as a Doppler-shifted temperature that is viewed analogously. In the jargon of the field this simple projection is referred to as the *freestreaming* of power from the monopole and dipole to higher multipole moments.

How accurately can the spectra ultimately be measured? As alluded to above, the fundamental limitation is set by "cosmic variance," the fact that there are only $2\ell + 1$ m-samples of the power in each multipole moment. This leads to an inevitable error of

$$\Delta C_\ell = \sqrt{\frac{2}{2\ell + 1}} C_\ell. \tag{4}$$

Allowing for further averaging over ℓ in bands of $\Delta\ell \approx \ell$, we see that the precision in the power spectrum determination scales as ℓ^{-1}, i.e., $\sim$1% at $\ell = 100$ and $\sim$0.1% at $\ell = 1000$. It is the combination of precision predictions and prospects for precision measurements that gives CMB anisotropies their unique stature.

There are two general caveats to these scalings. The first is that any source of noise, instrumental or astrophysical, increases the errors. If the noise is also Gaussian and has a known power spectrum, one simply replaces the power spectrum on the right-hand side of Equation 4 with the sum of the signal and noise power spectra (Knox 1995). This is why the errors for the Planck satellite increase near its resolution scale in Figure 1 (*bottom*). Because astrophysical foregrounds are typically non-Gaussian, it is usually also necessary to remove heavily contaminated regions, e.g., the galaxy. If the fraction of sky covered is $f_{\rm sky}$, then the errors increase by a factor of $f_{\rm sky}^{-1/2}$ and the resulting variance is usually dubbed "sample variance" (Scott et al. 1994). An $f_{\rm sky} = 0.65$ was chosen for the Planck satellite in Figure 1.

Cosmic Microwave Background Polarization Field

Whereas no polarization has yet been detected, general considerations of Thomson scattering suggest that up to 10% of the anisotropies at a given scale are polarized. Experimenters are currently hot on the trail, with upper limits approaching the expected level (Hedman et al. 2001, Keating et al. 2001). Thus, we expect polarization to be an extremely exciting field of study in the coming decade.

The polarization field can be analyzed in a way very similar to the temperature field, save for one complication. In addition to its strength, polarization also has an orientation, depending on relative strength of two linear polarization states. The classical literature has tended to describe polarization locally in terms of the Stokes

parameters Q and U[1], but recently cosmologists (Seljak 1997, Kamionkowski et al. 1997, Zaldarriaga & Seljak 1997) have found that the scalar E and pseudoscalar B, linear but nonlocal combinations of Q and U, provide a more useful description. Postponing the precise definition of E and B until "Polarization," below, we can, in complete analogy with Equation 1, decompose each of them in terms of multipole moments and then, following Equation 2, consider the power spectra,

$$\begin{aligned}\left\langle E^*_{\ell m} E_{\ell' m'}\right\rangle &= \delta_{\ell\ell'}\delta_{mm'} C_\ell^{EE},\\ \left\langle B^*_{\ell m} B_{\ell' m'}\right\rangle &= \delta_{\ell\ell'}\delta_{mm'} C_\ell^{BB},\\ \left\langle \Theta^*_{\ell m} E_{\ell' m'}\right\rangle &= \delta_{\ell\ell'}\delta_{mm'} C_\ell^{\Theta E}.\end{aligned} \tag{5}$$

Parity invariance demands that the cross correlation between the pseudoscalar B and the scalars Θ or E vanishes.

The polarization spectra shown in Figure 1 (*bottom*, plotted in μK following Equation 3) have several notable features. First, the amplitude of the *EE* spectrum of Equation 5 is indeed down from the temperature spectrum by a factor of 10. Second, the oscillatory structure of the *EE* spectrum is very similar to the temperature oscillations, although they are apparently out of phase but correlated with each other. Both of these features are a direct result of the simple physics of acoustic oscillations, as shown in "Acoustic Peaks" below. The final feature of the polarization spectra is the comparative smallness of the *BB* signal. Indeed, density perturbations do not produce B modes to first order. A detection of substantial B polarization, therefore, would be momentous. Whereas E polarization effectively doubles our cosmological information, supplementing that contained in C_ℓ, B detection would push us qualitatively forward into new areas of physics.

ACOUSTIC PEAKS

When the temperature of the Universe was $\sim$3000 K at a redshift $z_* \approx 10^3$, electrons and protons combined to form neutral hydrogen, an event usually known as recombination (Peebles 1968, Zel'dovich et al. 1969; see Seager et al. 2000 for recent refinements). Before this epoch, free electrons acted as glue between the photons and the baryons through Thomson and Coulomb scattering, so the cosmological plasma was a tightly coupled photon-baryon fluid (Peebles & Yu 1970). The spectrum depicted in Figure 1 can be explained almost completely by analyzing the behavior of this prerecombination fluid.

In "Basics," below, we start from the two basic equations of fluid mechanics and derive the salient characteristics of the anisotropy spectrum: the existence of

[1]There is also the possibility in general of circular polarization, described by the Stokes parameter V, but this is absent in cosmological settings.

peaks and troughs, the spacing between adjacent peaks, and the location of the peaks. These properties depend in decreasing order of importance on the initial conditions, the energy contents of the Universe before recombination, and those after recombination. Ironically, the observational milestones have been reached in almost the opposite order. Throughout the 1990s constraints on the location of the first peak steadily improved, culminating with precise determinations from the Toco (Miller et al. 1999), Boomerang (de Bernardis et al. 2000), and Maxima-1 (Hanany et al. 2000) experiments (see Figure 1, *top*). In the working cosmological model it shows up right where it should be if the present energy density of the Universe is equal to the critical density, i.e., if the Universe is flat. The skeptic should note that the working cosmological model assumes a particular form for the initial conditions and energy contents of the Universe before recombination, which we shall see have only recently been tested directly (with an as yet much lower level of statistical confidence) with the higher peaks.

In "Initial Conditions," below, we introduce the initial conditions that apparently are the source of all clumpiness in the Universe. In the context of ab initio models the term "initial conditions" refers to the physical mechanism that generates the primordial small perturbations. In the working cosmological model this mechanism is inflation and it sets the initial phase of the oscillations to be the same across all Fourier modes. Remarkably, from this fact alone comes the prediction that there will be peaks and troughs in the amplitude of the oscillations as a function of wavenumber. Additionally, the inflationary prediction of an approximately scale-invariant amplitude of the initial perturbations implies roughly scale-invariant oscillations in the power spectrum, and inflation generically predicts a flat Universe. These are all falsifiable predictions of the simplest inflationary models, and they have withstood the test against observations to date.

The energy contents of the Universe before recombination all leave their distinct signatures on the oscillations, as discussed in "Gravitational Forcing" and the two following sections. In particular, the cold dark matter and baryon signatures have now been seen in the data (Halverson et al. 2001, Netterfield et al. 2001, Lee et al. 2001). The coupling between electrons and photons is not perfect, especially as one approaches the epoch of recombination. As discussed in "Damping," below, this imperfect coupling leads to damping in the anisotropy spectrum: Very small scale inhomogeneities are smoothed out. The damping phenomenon has now been observed by the Cosmic Background Imager (CBI) experiment (Padin et al. 2001). Importantly, fluid imperfections also generate linear polarization as covered in "Polarization," below. Because the imperfection is minimal and appears only at small scales, the polarization generated is small and has not been detected to date.

After recombination the photons essentially travel freely to us today, so the problem of translating the acoustic inhomogeneities in the photon distribution at recombination to the anisotropy spectrum today is simply one of projection. This projection depends almost completely on one number, the angular diameter distance between us and the surface of last scattering. This number depends on the energy contents of the Universe after recombination through the expansion rate.

The hand-waving projection argument of "Basics" is formalized in "Integral Approach" in the process introducing the popular code used to compute anisotropies, CMBFAST. Finally, we discuss the sensitivity of the acoustic peaks to cosmological parameters in "Parameter Sensitivity."

Basics

For pedagogical purposes, let us begin with an idealization of a perfect photon-baryon fluid and neglect the dynamical effects of gravity and the baryons. Perturbations in this perfect fluid can be described by a simple continuity and a Euler equation that encapsulate the basic properties of acoustic oscillations.

The discussion of acoustic oscillations takes place exclusively in Fourier space. For example, we decompose the monopole of the temperature field into

$$\Theta_{\ell=0,m=0}(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} \Theta(\mathbf{k}) \tag{6}$$

and omit the subscript $_{00}$ on the Fourier amplitude. Because perturbations are very small, the evolution equations are linear, and different Fourier modes evolve independently. Therefore, instead of partial differential equations for a field $\Theta(\mathbf{x})$, we have ordinary differential equations for $\Theta(\mathbf{k})$. In fact, owing to rotational symmetry, all $\Theta(\mathbf{k})$ for a given k obey the same equations. Here and in the following sections we omit the wavenumber argument k when no confusion with physical space quantities will arise.

Temperature perturbations in Fourier space obey

$$\dot{\Theta} = -\frac{1}{3} k v_\gamma . \tag{7}$$

This equation for the photon temperature Θ, which does indeed look like the familiar continuity equation in Fourier space (derivatives ∇ become wavenumbers $i\mathbf{k}$), has a number of subtleties hidden in it, owing to the cosmological setting. First, the over-dot derivative here is with respect to conformal time $\eta \equiv \int dt/a(t)$. Because we are working in units in which the speed of light $c = 1$, η is also the maximum comoving distance a particle could have traveled since $t = 0$. It is often called the comoving horizon or more specifically the comoving particle horizon. The physical horizon is a times the comoving horizon.

Second, the photon fluid velocity, v_γ, has been written as a scalar instead of a vector. In the early universe only the velocity component parallel to the wavevector $\mathbf{k}$ is expected to be important because it alone has a source in gravity. Specifically, $\mathbf{v}_\gamma = -iv_\gamma \hat{\mathbf{k}}$. In terms of the moments introduced in "Observables," above, v_γ represents a dipole moment directed along $\mathbf{k}$. The factor of $1/3$ comes about because continuity conserves photon number, not temperature, and the number density $n_\gamma \propto T^3$. Finally, we reiterate that, for the time being, we are neglecting the effects of gravity.

The Euler equation for a fluid is an expression of momentum conservation. The momentum density of the photons is $(\rho_\gamma + p_\gamma)v_\gamma$, where the photon pressure

$p_\gamma = \rho_\gamma/3$. In the absence of gravity and viscous fluid imperfections, pressure gradients $\nabla p_\gamma = \nabla\rho_\gamma/3$ supply the only force. Because $\rho_\gamma \propto T^4$, this becomes $4k\Theta\bar{\rho}_\gamma/3$ in Fourier space. The Euler equation then becomes

$$\dot{v}_\gamma = k\Theta. \tag{8}$$

Differentiating the continuity equation and inserting the Euler equation yields the most basic form of the oscillator equation,

$$\ddot{\Theta} + c_s^2 k^2 \Theta = 0, \tag{9}$$

where $c_s \equiv \sqrt{\dot{p}/\dot{\rho}} = 1/\sqrt{3}$ is the sound speed in the (dynamically baryon-free) fluid. What this equation says is that pressure gradients act as a restoring force to any initial perturbation in the system, which thereafter oscillates at the speed of sound. Physically these temperature oscillations represent the heating and cooling of a fluid that is compressed and rarefied by a standing sound or acoustic wave. This behavior continues until recombination. Assuming negligible initial velocity perturbations, we have a temperature distribution at recombination of

$$\Theta(\eta_*) = \Theta(0)\cos(ks_*), \tag{10}$$

where $s = \int c_s\, d\eta \approx \eta/\sqrt{3}$ is the distance sound can travel by η, usually called the sound horizon. Asterisks denote evaluation at recombination z_*.

In the limit of scales that are large compared with the sound horizon $ks \ll 1$, the perturbation is frozen into its initial conditions. This is the gist of the statement that the large-scale anisotropies measured by COBE directly measure the initial conditions. On small scales the amplitude of the Fourier modes exhibit temporal oscillations, as shown in Figure 2 [with $\Psi = 0$, $\Psi_i = 3\Theta(0)$ for this idealization]. Modes that are caught at maxima or minima of their oscillation at recombination correspond to peaks in the power, i.e., the variance of $\Theta(k, \eta_*)$. Because sound takes half as long to travel half as far, modes corresponding to peaks follow a harmonic relationship $k_n = n\pi/s_*$, where n is an integer (see Figure 2*a*).

How does this spectrum of inhomogeneities at recombination appear to us today? Roughly speaking, a spatial inhomogeneity in the CMB temperature of wavelength λ appears as an angular anisotropy of scale $\theta \approx \lambda/D$, where $D(z)$ is the comoving angular diameter distance from the observer to redshift z. We address this issue more formally in "Integral Approach," below. In a flat universe $D_* = \eta_0 - \eta_* \approx \eta_0$, where $\eta_0 \equiv \eta\,(z = 0)$. In harmonic space the relationship implies a coherent series of acoustic peaks in the anisotropy spectrum, located at

$$\ell_n \approx n\ell_\mathrm{a}, \quad \ell_\mathrm{a} \equiv \pi D_*/s_*. \tag{11}$$

To get a feel for where these features should appear, note that in a flat matter-dominated universe $\eta \propto (1+z)^{-1/2}$ so that $\eta_*/\eta_0 \approx 1/30 \approx 2^\circ$. Equivalently, $\ell_1 \approx 200$. Notice that because we are measuring ratios of distances, the absolute distance scale drops out; we see in "Radiation Driving," below, that the Hubble

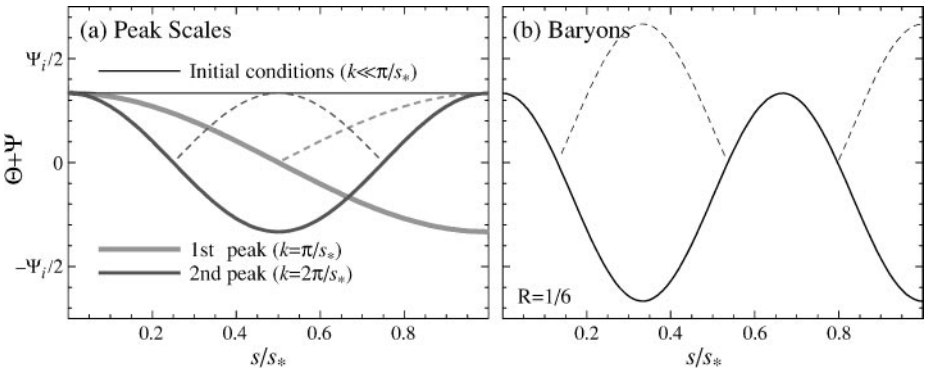

Figure 2 Idealized acoustic oscillations with time measured as the extent of the sound horizon relative to its value at recombination (s/s_*). (*a*) Peak scales: The wavemode that completes half an oscillation by recombination sets the physical scale of the first peak. Both minima and maxima correspond to peaks in power (*dashed lines*, absolute value), so higher peaks are integral multiples of this scale with equal height. Plotted here is the idealization of Equation 15 (constant potentials, no baryon loading). (*b*) Baryon loading: Baryon loading boosts the amplitudes of every other oscillation. Plotted here is the idealization of Equation 16 (constant potentials and baryon loading $R = 1/6$) for the third peak.

constant sneaks back into the problem because the Universe is not fully matter-dominated at recombination.

In a spatially curved universe the angular diameter distance no longer equals the coordinate distance, making the peak locations sensitive to the spatial curvature of the Universe (Doroshkevich et al. 1978, Kamionkowski et al. 1994). Consider first a closed universe with radius of curvature $R = H_0^{-1}|\Omega_{\rm tot} - 1|^{-1/2}$. Suppressing one spatial coordinate yields a two-sphere geometry with the observer situated at the pole (see Figure 3). Light travels on lines of longitude. A physical scale λ at fixed latitude given by the polar angle θ subtends an angle $\alpha = \lambda/R\sin\theta$. For $\alpha \ll 1$, a Euclidean analysis would infer a distance $D = R\sin\theta$, even though the coordinate distance along the arc is $d = \theta R$; thus

$$D = R\sin(d/R). \tag{12}$$

For open universes, simply replace sin with sinh. The result is that objects in an open (closed) universe are closer (further) than they appear, as if seen through a lens. In fact one way of viewing this effect is as the gravitational lensing caused by the background density (see "Gravitational Lensing," below). A given comoving scale at a fixed distance subtends a larger (smaller) angle in a closed (open) universe than a flat universe. This strong scaling with spatial curvature indicates that the observed first peak at $\ell_1 \approx 200$ constrains the geometry to be nearly spatially flat. We implicitly assume spatial flatness in the following sections unless otherwise stated.

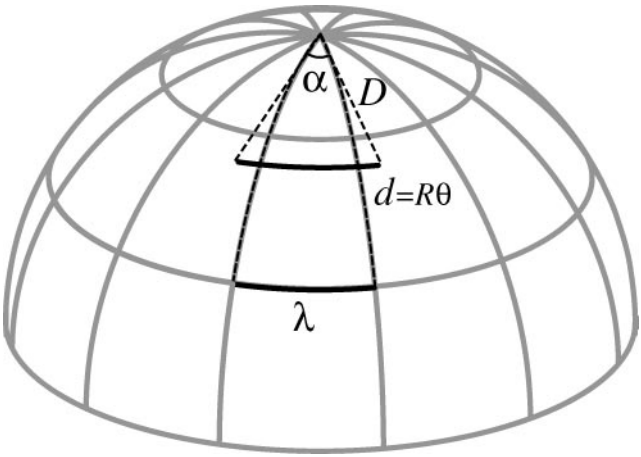

Figure 3 Angular diameter distance. In a closed universe objects are further than they appear to be from Euclidean (flat) expectations, corresponding to the difference between coordinate distance d and angular diameter distance D. Consequently, at a fixed coordinate distance, a given angle corresponds to a smaller spatial scale in a closed universe. Acoustic peaks therefore appear at larger angles or lower ℓ in a closed universe. The converse is true for an open universe.

Finally, in a flat dark energy-dominated universe the conformal age of the universe decreases approximately as $\eta_0 \to \eta_0\,(1 + \ln \Omega_m^{0.085})$. For reasonable Ω_m, this causes only a small shift of ℓ_1 to lower multipoles (see Figure 4, color insert) relative to the effect of curvature. Combined with the effect of the radiation near recombination, the peak locations provide a means to measure the physical age t_0 of a flat universe (Hu et al. 2001).

Initial Conditions

As suggested above, observations of the location of the first peak strongly point to a flat universe. This is encouraging news for adherents of inflation, a theory that initially predicted $\Omega_{\rm tot} = 1$ at a time when few astronomers would sign on to such a high value (see Liddle & Lyth 1993 for a review). However, the argument for inflation goes beyond the confirmation of flatness. In particular, the discussion of the previous subsection begs the question: Whence $\Theta(0)$, the initial conditions of the temperature fluctuations? The answer requires the inclusion of gravity and considerations of causality that point to inflation as the origin of structure in the Universe.

The calculations of the typical angular scale of the acoustic oscillations in the previous section are familiar in another context: the horizon problem. Because the sound speed is near the speed of light, the degree scale also marks the extent of a

causally connected region or particle horizon at recombination. For the picture in the previous section to hold, the perturbations must have been laid down while the scales in question were still far outside the particle horizon[2]. The recent observational verification of this basic peak structure presents a problem potentially more serious than the original horizon problem of approximate isotropy: The mechanism that smooths fluctuations in the Universe must also regenerate them with superhorizon-sized correlations at the 10^{-5} level. Inflation is an idea that solves both problems simultaneously.

The inflationary paradigm postulates that an early phase of near exponential expansion of the Universe was driven by a form of energy with negative pressure. In most models this energy is usually provided by the potential energy of a scalar field. The inflationary era brings the observable universe to a nearly smooth and spatially flat state. Nonetheless, quantum fluctuations in the scalar field are unavoidable and are carried to large physical scales by the expansion. Because an exponential expansion is self-similar in time, the fluctuations are scale-invariant, i.e., in each logarithmic interval in scale the contribution to the variance of the fluctuations is equal. Because the scalar field carries the energy density of the Universe during inflation, its fluctuations induce variations in the spatial curvature (Guth & Pi 1985, Hawking 1982, Bardeen et al. 1983). Instead of perfect flatness, inflation predicts that each scale will resemble a very slightly open or closed universe. This fluctuation in the geometry of the Universe is essentially frozen while the perturbation is outside the horizon (Bardeen 1980).

Formally, curvature fluctuations are perturbations to the space-space piece of the metric. In a Newtonian coordinate system, or gauge, where the metric is diagonal the spatial curvature fluctuation is called $\delta g_{ij} = 2a^2\Phi\delta_{ij}$ (see, e.g., Ma & Bertschinger 1995). The more familiar Newtonian potential is the time-time fluctuation $\delta g_{tt} = 2\Psi$ and is approximately $\Psi \approx -\Phi$. Approximate scale invariance then says that $\Delta^2_\Phi \equiv k^3 P_\Phi(k)/2\pi^2 \propto k^{n-1}$, where $P_\Phi(k)$ is the power spectrum of Φ and the tilt $n \approx 1$.

Now let us relate the inflationary prediction of scale-invariant curvature fluctuations to the initial temperature fluctuations. Newtonian intuition based on the Poisson equation $k^2\Phi = 4\pi G a^2 \delta\rho$ tells us that on large scales (small k) density and hence temperature fluctuations should be negligible compared with Newtonian potential. General relativity says otherwise because the Newtonian potential is also a time-time fluctuation in the metric. It corresponds to a temporal shift of $\delta t/t = \Psi$. The CMB temperature varies as the inverse of the scale factor, which in turn depends on time as $a \propto t^{2/[3(1+p/\rho)]}$. Therefore, the fractional change in the CMB temperature

$$\Theta = -\frac{\delta a}{a} = -\frac{2}{3}\left(1 + \frac{p}{\rho}\right)^{-1} \frac{\delta t}{t}. \tag{13}$$

[2]Recall that the comoving scale k does not vary with time. At very early times, then, the wavelength k^{-1} is much larger than the horizon η.

Thus, a temporal shift produces a temperature perturbation of $-\Psi/2$ in the radiation-dominated era (when $p=\rho/3$) and $-2\Psi/3$ in the matter-dominated epoch ($p=0$) (Sachs & Wolfe 1967, Peacock 1991, White & Hu 1997). The initial temperature perturbation is therefore inextricably linked with the initial gravitational potential perturbation. Inflation predicts scale-invariant initial fluctuations in both the CMB temperature and the spatial curvature in the Newtonian gauge.

Alternate models that seek to obey causality can generate curvature fluctuations only inside the particle horizon. Because the perturbations are then not generated at the same epoch independent of scale, there is no longer a unique relationship between the phase of the oscillators. That is, the argument of the cosine in Equation 10 becomes $ks_* + \phi(\mathbf{k})$, where ϕ is a phase that can in principle be different for different wavevectors, even those with the same magnitude k. This leads to temporal incoherence in the oscillations and hence a washing out of the acoustic peaks (Albrecht et al. 1996), most notably in cosmological defect models (Allen et al. 1997, Seljak et al. 1997). Complete incoherence is not a strict requirement of causality because there are other ways to synch up the oscillations. For example, many isocurvature models, in which the initial spatial curvature is unperturbed, are coherent because their oscillations begin with the generation of curvature fluctuations at horizon crossing (Hu & White 1996). Still they typically have $\phi \neq 0$ (compare Turok 1996). Independent of the angular diameter distance D_*, the ratio of the peak locations gives the phase: $\ell_1 : \ell_2 : \ell_3 \sim 1 : 2 : 3$ for $\phi = 0$. Likewise independent of a constant phase, the spacing of the peaks $\ell_n - \ell_{n-1} = \ell_A$ gives a measure of the angular diameter distance (Hu & White 1996). The observations, which indicate coherent oscillations with $\phi = 0$, therefore, have provided a nontrivial test of the inflationary paradigm and supplied a substantially more stringent version of the horizon problem for contenders to solve.

Gravitational Forcing

We saw above that fluctuations in a scalar field during inflation get turned into temperature fluctuations via the intermediary of gravity. Gravity affects Θ in more ways than this. The Newtonian potential and spatial curvature alter the acoustic oscillations by providing a gravitational force on the oscillator. The Euler equation (8) gains a term on the right-hand side owing to the gradient of the potential $k\Psi$. The main effect of gravity then is to make the oscillations a competition between pressure gradients $k\Theta$ and potential gradients $k\Psi$ with an equilibrium when $\Theta + \Psi = 0$.

Gravity also changes the continuity equation. Because the Newtonian curvature is essentially a perturbation to the scale factor, changes in its value also generate temperature perturbations by analogy to the cosmological redshift $\delta\Theta = -\delta\Phi$, and so the continuity equation (7) gains a contribution of $-\dot{\Phi}$ on the right-hand side.

These two effects bring the oscillator equation (9) to

$$\ddot{\Theta} + c_s^2 k^2 \Theta = -\frac{k^2}{3}\Psi - \ddot{\Phi}. \tag{14}$$

In a flat universe and in the absence of pressure Φ and Ψ are constant. Also, in the absence of baryons, $c_s^2 = 1/3$, so the new oscillator equation is identical to Equation 9 with Θ replaced by $\Theta + \Psi$. The matter-dominated epoch is then

$$[\Theta + \Psi](\eta) = [\Theta + \Psi](\eta_{\rm md})\cos(ks) = \frac{1}{3}\Psi(\eta_{\rm md})\cos(ks), \quad (15)$$

where $\eta_{\rm md}$ represents the start of the matter-dominated epoch (see Figure 2*a*). We have used the matter-dominated "initial conditions" for Θ given in the previous section, assuming large scales, $ks_{\rm md} \ll 1$.

The results from the idealization of "Basics," above, carry through with a few exceptions. Even without an initial temperature fluctuation to displace the oscillator, acoustic oscillations would arise by the infall and compression of the fluid into gravitational potential wells. Because it is the effective temperature $\Theta + \Psi$ that oscillates, they occur even if $\Theta(0) = 0$. The quantity $\Theta + \Psi$ can be thought of as an effective temperature in another way: After recombination, photons must climb out of the potential well to the observer and thus suffer a gravitational redshift of $\Delta T/T = \Psi$. The effective temperature fluctuation is therefore also the observed temperature fluctuation. We now see that the large scale limit of Equation 15 recovers the famous Sachs-Wolfe result that the observed temperature perturbation is $\Psi/3$ and overdense regions correspond to cold spots on the sky (Sachs & Wolfe 1967). When $\Psi < 0$, although Θ is positive, the effective temperature $\Theta + \Psi$ is negative. The plasma begins effectively rarefied in gravitational potential wells. As gravity compresses the fluid and pressure resists, rarefaction becomes compression and rarefaction again. The first peak corresponds to the mode that is caught in its first compression by recombination. The second peak at roughly half the wavelength corresponds to the mode that went through a full cycle of compression and rarefaction by recombination. We will use this language of the compression and rarefaction phase inside initially overdense regions, but one should bear in mind that there are an equal number of initially underdense regions with the opposite phase.

Baryon Loading

So far we have been neglecting the baryons in the dynamics of the acoustic oscillations. To see whether this is a reasonable approximation consider the photon-baryon momentum density ratio $R = (p_b + \rho_b)/(p_\gamma + \rho_\gamma) \approx 30\Omega_b h^2 (z/10^3)^{-1}$. For typical values of the baryon density this number is of order unity at recombination, and so we expect baryonic effects to begin appearing in the oscillations just as they are frozen in.

Baryons are conceptually easy to include in the evolution equations because their momentum density provides extra inertia in the joint Euler equation for pressure and potential gradients to overcome. Because inertial and gravitational mass are equal, all terms in the Euler equation except the pressure gradient are

multiplied by $1+R$, leading to the revised oscillator equation (Hu & Sugiyama, 1995)

$$c_s^2 \frac{d}{d\eta}\left(c_s^{-2}\dot{\Theta}\right) + c_s^2 k^2 \Theta = -\frac{k^2}{3}\Psi - c_s^2 \frac{d}{d\eta}\left(c_s^{-2}\dot{\Phi}\right), \tag{16}$$

where we have used the fact that the sound speed is reduced by the baryons to $c_s = 1/\sqrt{3(1+R)}$.

To get a feel for the implications of the baryons take the limit of constant R, Φ, and Ψ. Then $d^2(R\Psi)/d\eta^2 (= 0)$ may be added to the left-hand side to again put the oscillator equation in the form of Equation 9 with $\Theta \to \Theta + (1+R)\Psi$. The solution then becomes

$$[\Theta + (1+R)\Psi](\eta) = [\Theta + (1+R)\Psi](\eta_{\rm md})\cos(ks). \tag{17}$$

Aside from the lowering of the sound speed, which decreases the sound horizon, baryons have two distinguishing effects: They enhance the amplitude of the oscillations and shift the equilibrium point to $\Theta = -(1+R)\Psi$ (see Figure 2*b*). These two effects are intimately related and are easy to understand because the equations are exactly those of a mass $m = 1+R$ on a spring in a constant gravitational field. For the same initial conditions, increasing the mass causes the oscillator to fall further in the gravitational field, leading to larger oscillations and a shifted zero point.

The shifting of the zero point of the oscillator has significant phenomenological consequences. Because it is still the effective temperature $\Theta + \Psi$ that is the observed temperature, the zero-point shift breaks the symmetry of the oscillations. The baryons enhance only the compressional phase, i.e., every other peak. For the working cosmological model these are the first, third, fifth . . .. Physically, the extra gravity provided by the baryons enhances compression into potential wells.

These qualitative results remain true in the presence of a time-variable R. An additional effect arises owing to the adiabatic damping of an oscillator with a time-variable mass. Because the energy/frequency of an oscillator is an adiabatic invariant, the amplitude must decay as $(1+R)^{-1/4}$. This can also be understood by expanding the time derivatives in Equation 16 and identifying the $\dot{R}\dot{\Theta}$ term as the remnant of the familiar expansion drag on baryon velocities.

Radiation Driving

We have hitherto also been neglecting the energy density of the radiation in comparison to the matter. The matter-to-radiation energy-density ratio scales as $\rho_m/\rho_r \approx 24\Omega_m h^2 (z/10^3)^{-1}$ and so is also of order unity at recombination for reasonable parameters. Moreover, fluctuations corresponding to the higher peaks entered the sound horizon at an earlier time, during radiation domination.

Including the radiation changes the expansion rate of the Universe and hence the physical scale of the sound horizon at recombination. It introduces yet another potential ambiguity in the interpretation of the location of the peaks. Fortunately,

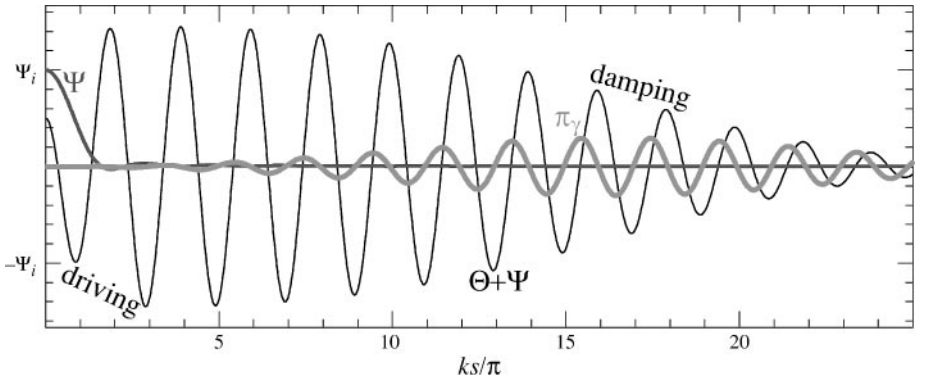

Figure 5 Radiation driving and diffusion damping. The decay of the potential Ψ drives the oscillator in the radiation-dominated epoch. Diffusion generates viscosity π_γ, i.e., a quadrupole moment in the temperature, which damps oscillations and generates polarization. Plotted here is the numerical solution to Equation 18 and Equation 19 for a mode with wavelength much smaller than the sound horizon at decoupling, $ks_* \gg 1$.

the matter-radiation ratio has another effect in the power spectrum by which it can be distinguished. Radiation drives the acoustic oscillations by making the gravitational force evolve with time (Hu & Sugiyama 1995). Matter does not.

The exact evolution of the potentials is determined by the relativistic Poisson equation. However, qualitatively, we know that the background density is decreasing with time, so unless the density fluctuations in the dominant component grow unimpeded by pressure, potentials will decay. In particular, in the radiation-dominated era once pressure begins to fight gravity at the first compressional maxima of the wave, the Newtonian gravitational potential and spatial curvature must decay (see Figure 5).

This decay actually drives the oscillations: It is timed to leave the fluid maximally compressed with no gravitational potential to fight as it turns around. The net effect is doubled because the redshifting from the spatial metric fluctuation Φ also goes away at the same time. When the Universe becomes matter dominated, the gravitational potential is no longer determined by photon-baryon density perturbations but by the pressureless cold dark matter. Therefore, the amplitudes of the acoustic peaks increase as the cold dark matter-to-radiation ratio decreases (Seljak 1994, Hu & Sugiyama 1995). Density perturbations in any form of radiation will stop growing around horizon crossing and lead to this effect. The net result is that across the horizon scale at matter-radiation equality ($k_{\rm eq} \equiv (4-2\sqrt{2})/\eta_{\rm eq}$) the acoustic amplitude increases by a factor of 4–5 (Hu & Sugiyama 1996). By eliminating gravitational potentials, photon-baryon acoustic oscillations eliminate the alternating peak heights from baryon loading. The observed high third peak (Halverson et al. 2001) is a good indication that cold dark matter both exists and dominates the energy density at recombination.

Damping

The photon-baryon fluid has slight imperfections corresponding to shear viscosity and heat conduction in the fluid (Silk 1968, Weinberg 1971). These imperfections damp acoustic oscillations. To consider these effects, we now present the equations of motion of the system in their full form, including separate continuity and Euler equations for the baryons. Formally the continuity and Euler equations follow from the covariant conservation of the joint stress-energy tensor of the photon-baryon fluid. Because photon and baryon numbers are separately conserved, the continuity equations are unchanged,

$$\dot{\Theta} = -\frac{k}{3}v_\gamma - \dot{\Phi}, \quad \dot{\delta}_b = -kv_b - 3\dot{\Phi}, \tag{18}$$

where δ_b and v_b are the density perturbation and fluid velocity of the baryons. The Euler equations contain qualitatively new terms:

$$\dot{v}_\gamma = k(\Theta + \Psi) - \frac{k}{6}\pi_\gamma - \dot{\tau}(v_\gamma - v_b),$$
$$\dot{v}_b = -\frac{\dot{a}}{a}v_b + k\Psi + \dot{\tau}(v_\gamma - v_b)/R. \tag{19}$$

For the baryons the first term on the right accounts for cosmological expansion, which makes momenta decay as a^{-1}. The third term on the right accounts for momentum exchange in the Thomson scattering between photons and electrons (protons are very tightly coupled to electrons via Coulomb scattering), with $\dot{\tau} \equiv n_e \sigma_T a$, the differential Thomson optical depth, and is compensated by its opposite in the photon Euler equation. These terms are the origin of heat conduction imperfections. If the medium is optically thick across a wavelength, $\dot{\tau}/k \gg 1$ and the photons and baryons cannot slip past each other. As it becomes optically thin, slippage dissipates the fluctuations.

In the photon Euler equation there is an extra force on the right-hand side owing to anisotropic stress gradients or radiation viscosity in the fluid, π_γ. The anisotropic stress is directly proportional to the quadrupole moment of the photon temperature distribution. A quadrupole moment is established by gradients in v_γ as photons from, for example, neighboring temperature crests meet at a trough (see Figure 6, "*damping and polarization*"). However, it is destroyed by scattering. Thus, $\pi_\gamma = 2(kv_\gamma/\dot{\tau})A_\mathrm{v}$, where the order unity constant can be derived from the Boltzmann equation as $A_\mathrm{v} = 16/15$ (Kaiser 1983). Its evolution is shown in Figure 5. With the continuity Equation 7, $kv_\gamma \approx -3\dot{\Theta}$, so viscosity takes the form of a damping term. The heat conduction term can be shown to have a similar effect by expanding the Euler equations in $k/\dot{\tau}$. The final oscillator equation including both terms becomes

$$c_s^2 \frac{d}{d\eta}\left(c_s^{-2}\dot{\Theta}\right) + \frac{k^2 c_s^2}{\dot{\tau}}[A_\mathrm{v} + A_\mathrm{h}]\dot{\Theta} + c_s^2 k^2 \Theta = -\frac{k^2}{3}\Psi - c_s^2 \frac{d}{d\eta}\left(c_s^{-2}\dot{\Phi}\right), \tag{20}$$

where the heat conduction coefficient $A_h = R^2/(1+R)$. Thus, we expect the inhomogeneities to be damped by an exponential factor of order $e^{-k^2\eta/\dot{\tau}}$ (see Figure 5). The damping scale k_d is thus of order $\sqrt{\dot{\tau}/\eta}$, corresponding to the geometric mean of the horizon and the mean free path. Damping can be thought of as the result of the random walk in the baryons that takes photons from hot regions into cold and vice versa (Silk 1968). Detailed numerical integration of the equations of motion are required to track the rapid growth of the mean free path and damping length through recombination itself. These calculations show that the damping scale is of order $k_d s_* \approx 10$, leading to a substantial suppression of the oscillations beyond the third peak.

How does this suppression depend on the cosmological parameters? As the matter density $\Omega_m h^2$ increases, the horizon η_* decreases because the expansion rate goes up. Because the diffusion length is proportional to $\sqrt{\eta_*}$, it too decreases as the matter density goes up but not as much as the angular diameter distance D_*, which is also inversely proportional to the expansion rate. Thus, more matter translates into more damping at a fixed multipole moment; conversely, it corresponds to slightly less damping at a fixed peak number. The dependence on baryons is controlled by the mean free path, which is in turn controlled by the free electron density: The increase in electron density caused by an increase in the baryons is partially offset by a decrease in the ionization fraction caused by recombination. The net result under the Saha approximation is that the damping length scales approximately as $(\Omega_b h^2)^{-1/4}$. Accurate fitting formulae for this scale in terms of cosmological parameters can be found in (Hu & White 1997c).

Polarization

The dissipation of the acoustic oscillations leaves a signature in the polarization of CMB in its wake (see e.g., Hu & White 1997a and references therein for a more complete treatment). Much like reflection off of a surface, Thomson scattering induces a linear polarization in the scattered radiation. Consider incoming radiation in the $-\mathbf{x}$ direction scattered at right angles into the $\mathbf{z}$ direction (see Figure 7, *top left*). Heuristically, incoming radiation shakes an electron in the direction of its electric field vector or polarization $\hat{\epsilon}'$, causing it to radiate with an outgoing polarization parallel to that direction. However, because the outgoing polarization $\hat{\epsilon}$ must be orthogonal to the outgoing direction, incoming radiation that is polarized parallel to the outgoing direction cannot scatter, leaving only one polarization state, more generally, the Thomson differential cross section $d\sigma_T/d\Omega \propto |\hat{\epsilon}' \cdot \hat{\epsilon}|^2$.

Unlike the reflection of sunlight off a surface, the incoming radiation comes from all angles. If it were completely isotropic in intensity, radiation coming along the $\hat{\mathbf{y}}$ would provide the polarization state that is missing from that coming along $\hat{\mathbf{x}}$, leaving the net outgoing radiation unpolarized. Only a quadrupole temperature anisotropy in the radiation generates a net linear polarization from Thomson scattering. As we have seen, a quadrupole can only be generated causally by the

motion of photons and only if the Universe is optically thin to Thomson scattering across this scale (i.e., it is inversely proportional to $\dot{\tau}$). Polarization generation suffers from a Catch-22: The scattering that generates polarization also suppresses its quadrupole source.

The fact that the polarization strength is of order the quadrupole in amplitude explains the shape and height of the polarization spectra in Figure 1*b*. The monopole and dipole Θ and v_γ are of the same order of magnitude at recombination, but their oscillations are $\pi/2$ out of phase as follows from Equation 9 and Equation 10. Because the quadrupole is of order $kv_\gamma/\dot{\tau}$ (see Figure 5), the polarization spectrum should be smaller than the temperature spectrum by a factor of order $k/\dot{\tau}$ at recombination. As in the case of the damping, the precise value requires numerical work (Bond & Efstathiou 1987) because $\dot{\tau}$ changes so rapidly near recombination. Calculations show a steady rise in the polarized fraction with increasing l or k to a maximum of about 10% before damping destroys the oscillations and hence the dipole source. Because v_γ is out of phase with the monopole, the polarization peaks should also be out of phase with the temperature peaks. Indeed, Figure 1*b* shows that this is the case. Furthermore, the phase relation also tells us that the polarization is correlated with the temperature perturbations. Because the correlation power $C_\ell^{\Theta E}$ is the product of the two, it exhibits oscillations at twice the acoustic frequency.

Until now, we have focused on the polarization strength without regard to its orientation. The orientation, like a two-dimensional vector, is described by two components, E and B. The E and B decomposition is simplest to visualize in the small scale limit, where spherical harmonic analysis coincides with Fourier analysis (Seljak 1997). The wavevector **k** picks out a preferred direction against which the polarization direction is measured (see Figure 7, *top right*). Because the linear polarization is a "headless vector" that remains unchanged upon a 180° rotation, the two numbers E and B that define it represent polarization aligned or orthogonal with the wavevector (positive and negative E) and crossed at $\pm 45°$ (positive and negative B).

In linear theory scalar perturbations such as the gravitational potential and temperature perturbations have only one intrinsic direction associated with them, that provided by **k**, and the orientation of the polarization inevitably takes it cue from that one direction, thereby producing an E-mode. The generalization to an all-sky characterization of the polarization changes none of these qualitative features. The E-mode and the B-mode are formally distinguished by the orientation of the Hessian of the Stokes parameters that define the direction of the polarization itself. This geometric distinction is preserved under summation of all Fourier modes as well as the generalization of Fourier analysis to spherical harmonic analysis.

The acoustic peaks in the polarization appear exclusively in the EE power spectrum of Equation 5. This distinction is very useful, as it allows a clean separation of this effect from those occuring beyond the scope of the linear perturbation theory of scalar fluctuations: in particular, gravitational waves). Moreover, in the working cosmological model, the polarization peaks and correlation are precise predictions of the temperature peaks, as they depend on the same physics. As

such, their detection would represent a sharp test of the implicit assumptions of the working model, especially its initial conditions and ionization history.

Integral Approach

The discussion in the previous sections suffices for a qualitative understanding of the acoustic peaks in the power spectra of the temperature and polarization anisotropies. To refine this treatment we must consider more carefully the sources of anisotropies and their projection into multipole moments.

Because the description of the acoustic oscillations takes place in Fourier space, the projection of inhomogeneities at recombination onto anisotropies today has an added level of complexity. An observer today sees the acoustic oscillations in effective temperature as they appeared on a spherical shell at $\mathbf{x} = D_* \hat{n}$ at recombination, where $\hat{n}$ is the direction vector and $D_* = \eta_0 - \eta_*$ is the distance light can travel between recombination and the present (see Figure 6). Having solved for the Fourier amplitude $[\Theta + \Psi](k, \eta_*)$, we can expand the exponential in Equation 6 in terms of spherical harmonics, so the observed anisotropy today is

$$\Theta(\hat{\mathbf{n}}, \eta_0) = 4\pi \sum_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}) \left[(-i)^\ell \int \frac{d^3k}{(2\pi)^3} \, a_\ell(k) Y^*_{\ell m}(\hat{\mathbf{k}}) \right], \tag{21}$$

where the projected source $a_\ell(k) = [\Theta + \Psi](\mathbf{k}, \eta_*) j_\ell(kD_*)$. Because the spherical harmonics are orthogonal, Equation 1 implies that $\Theta_{\ell m}$ today is given by the integral in square brackets today. A given plane wave actually produces a range of anisotropies in angular scale, as is obvious from Figure 6. The one-to-one mapping between wavenumber and multipole moment described in "Basics," above, is only approximately true and comes from the fact that the spherical Bessel function $j_\ell(kD_*)$ is strongly peaked at $kD_* \approx \ell$. Notice that this peak corresponds to contributions in the direction orthogonal to the wavevector where the correspondence between ℓ and k is one-to-one (see Figure 6).

Projection is less straightforward for other sources of anisotropy. We have hitherto neglected the fact that the acoustic motion of the photon-baryon fluid also produces a Doppler shift in the radiation that appears to the observer as a temperature anisotropy as well. In fact, we argued above that $v_b \approx v_\gamma$ is of comparable magnitude but out of phase with the effective temperature. If the Doppler effect projected in the same way as the effective temperature, it would wash out the acoustic peaks. However, the Doppler effect has a directional dependence as well because it is only the line-of-sight velocity that produces the effect. Formally, it is a dipole source of temperature anisotropies and hence has an $\ell = 1$ structure. The coupling of the dipole and plane-wave angular momenta implies that the projection of the Doppler effect involves a combination of $j_{\ell \pm 1}$ that may be rewritten as $j'_\ell(x) \equiv d j_\ell(x)/dx$. The structure of j'_ℓ lacks a strong peak at $x = \ell$. Physically this corresponds to the fact that the velocity is irrotational and hence has no component in the direction orthogonal to the wavevector (see Figure 6). Correspondingly, the Doppler effect cannot produce strong peak structures (Hu & Sugiyama 1995). The

observed peaks must be acoustic peaks in the effective temperature, not Doppler peaks.

There is one more subtlety involved when passing from acoustic oscillations to anisotropies. Recall from "Radiation Driving," above, that radiation leads to decay of the gravitational potentials. Residual radiation after decoupling therefore implies that the effective temperature is not precisely $[\Theta + \Psi](\eta_*)$. The photons actually have slightly shallower potentials to climb out of and lose the perturbative analogue of the cosmological redshift, so the $[\Theta + \Psi](\eta_*)$ overestimates the difference between the true photon temperature and the observed temperature. This effect of course is already in the continuity equation for the monopole Equation 18, so the source in Equation 21 gets generalized to

$$a_\ell(k) = [\Theta + \Psi](\eta_*) j_l(kD_*) + v_b(k, \eta_*) j'_\ell(kD_*) + \int_{\eta_*}^{\eta_0} d\eta (\dot{\Psi} - \dot{\Phi}) j_l(kD). \quad (22)$$

The last term vanishes for constant gravitational potentials but is nonzero if residual radiation driving exists, as it will in low $\Omega_m h^2$ models. Note that residual radiation driving is particularly important because it adds in phase with the monopole: The potentials vary in time only near recombination, so the Bessel function can be set to $j_l(kD_*)$ and removed from the η integral. This complication has the effect of decreasing the multipole value of the first peak ℓ_1 as the matter-radiation ratio at recombination decreases (Hu & Sugiyama 1995). Finally, we mention that time varying potentials can also play a role at very late times owing to nonlinearities or, a cosmological constant. Those contributions, to be discussed more in "Integrated Sachs-Wolfe Effect," are sometimes referred to as late integrated Sachs-Wolfe effects and do not add coherently with $[\Theta + \Psi](\eta_*)$.

Putting these expressions together and squaring, we obtain the power spectrum of the acoustic oscillations

$$C_\ell = \frac{2}{\pi} \int \frac{dk}{k} k^3 a_\ell^2(k). \quad (23)$$

This formulation of the anisotropies in terms of projections of sources with specific local angular structure can be completed to include all types of sources of temperature and polarization anisotropies at any given epoch in time, linear or nonlinear: the monopole, dipole, and quadrupole sources arising from density perturbations, vorticity, and gravitational waves (Hu & White 1997b). In a curved geometry one replaces the spherical Bessel functions with ultraspherical Bessel functions (Abbott & Schaefer 1986, Hu et al. 1998). Precision in the predictions of the observables is then limited only by the precision in the prediction of the sources. This formulation is ideal for cases in which the sources are governed by nonlinear physics, even though the CMB responds linearly, as we see in "Beyond the Peaks," below.

Perhaps more importantly, the widely used CMBFAST code (Seljak & Zaldarriaga 1996) exploits these properties to calculate the anisotropies in linear perturbation efficiently. It numerically solves for the smoothly varying sources

on a sparse grid in wavenumber, interpolating in the integrals for a handful of ℓ's in the smoothly varying C_ℓ. It has largely replaced the original ground breaking codes (Wilson & Silk 1981, Bond & Efstathiou 1984, Vittorio & Silk 1984) based on tracking the rapid temporal oscillations of the multipole moments that simply reflect structure in the spherical Bessel functions themselves.

Parameter Sensitivity

The phenomenology of the acoustic peaks in the temperature and polarization is essentially described by four observables and the initial conditions (Hu et al. 1997). These are the angular extents of the sound horizon $\ell_a \equiv \pi D_*/s_*$, the particle horizon at matter radiation equality $\ell_{eq} \equiv k_{eq} D_*$, and the damping scale $\ell_d \equiv k_d D_*$, as well as the value of the baryon-photon momentum density ratio R_*. ℓ_a sets the spacing between of the peaks; ℓ_{eq} and ℓ_d compete to determine their amplitude through radiation driving and diffusion damping. R_* sets the baryon loading and, along with the potential well depths set by ℓ_{eq}, fixes the modulation of the even and odd peak heights. The initial conditions set the phase, or equivalently the location of the first peak in units of ℓ_a, and an overall tilt n in the power spectrum.

In the model of Figure 1, these numbers are $\ell_a = 301$ ($\ell_1 = 0.73\ell_a$), $\ell_{eq} = 149$, $\ell_d = 1332$, $R_* = 0.57$, and $n = 1$, and in this family of models the parameter sensitivity is approximately (Hu et al. 2001)

$$\frac{\Delta\ell_a}{\ell_a} \approx -0.24\frac{\Delta\Omega_m h^2}{\Omega_m h^2} + 0.07\frac{\Delta\Omega_b h^2}{\Omega_b h^2} - 0.17\frac{\Delta\Omega_\Lambda}{\Omega_\Lambda} - 1.1\frac{\Delta\Omega_{tot}}{\Omega_{tot}},$$

$$\frac{\Delta\ell_{eq}}{\ell_{eq}} \approx 0.5\frac{\Delta\Omega_m h^2}{\Omega_m h^2} - 0.17\frac{\Delta\Omega_\Lambda}{\Omega_\Lambda} - 1.1\frac{\Delta\Omega_{tot}}{\Omega_{tot}},$$

$$\frac{\Delta\ell_d}{\ell_d} \approx -0.21\frac{\Delta\Omega_m h^2}{\Omega_m h^2} + 0.20\frac{\Delta\Omega_b h^2}{\Omega_b h^2} - 0.17\frac{\Delta\Omega_\Lambda}{\Omega_\Lambda} - 1.1\frac{\Delta\Omega_{tot}}{\Omega_{tot}}, \quad (24)$$

and $\Delta R_*/R_* \approx 1.0\Delta\Omega_b h^2/\Omega_b h^2$. Current observations indicate that $\ell_a = 304 \pm 4$, $\ell_{eq} = 168 \pm 15$, $\ell_d = 1392 \pm 18$, $R_* = 0.60 \pm 0.06$, and $n = 0.96 \pm 0.04$ (Knox et al. 2001; see also Wang et al. 2001, Pryke et al. 2001, de Bernardis et al. 2002) if gravitational wave contributions are subdominant and the reionization redshift is low, as assumed in the working cosmological model (see "Standard Cosmological Paradigm," above).

The acoustic peaks therefore contain three rulers for the angular diameter distance test for curvature, i.e., deviations from $\Omega_{tot} = 1$. However, contrary to popular belief, any one of these alone is not a standard ruler whose absolute scale is known even in the working cosmological model. This is reflected in the sensitivity of these scales to other cosmological parameters. For example, the dependence of ℓ_a on $\Omega_m h^2$, and hence on the Hubble constant, is quite strong. However, in combination with a measurement of the matter-radiation ratio from ℓ_{eq}, this degeneracy is broken.

The weaker degeneracy of ℓ_a on the baryons can likewise be broken from a measurement of the baryon-photon ratio R_*. The damping scale ℓ_d provides an additional consistency check on the implicit assumptions of the working model, e.g., recombination and the energy contents of the Universe during this epoch. What makes the peaks so valuable for this test is that the rulers are standardizable and contain a built-in consistency check.

There remains a weak but perfect degeneracy between Ω_{tot} and Ω_Λ because they both appear only in D_*. This is called the angular diameter distance degeneracy in the literature and can readily be generalized to dark energy components beyond the cosmological constant assumed here. Because the effect of Ω_Λ is intrinsically so small, it creates a correspondingly small ambiguity in Ω_{tot} for reasonable values of Ω_Λ. The down side is that dark energy can never be isolated through the peaks alone because it only takes a small amount of curvature to mimic its effects. The evidence for dark energy through the CMB comes about by allowing for external information. The most important is the nearly overwhelming direct evidence for $\Omega_m < 1$ from local structures in the Universe. The second is the measurements of a relatively high Hubble constant $h \approx 0.7$; combined with a relatively low $\Omega_m h^2$ that is preferred in the CMB data, it implies $\Omega_m < 1$, but at low significance currently.

The upshot is that precise measurements of the acoustic peaks yield precise determinations of four fundamental parameters of the working cosmological model: $\Omega_m h^2$, $\Omega_m h^2$, D_*, and n. More generally, the first three can be replaced by ℓ_a, ℓ_{eq}, ℓ_d, and R_* to extend these results to models in which the underlying assumptions of the working model are violated.

BEYOND THE PEAKS

Once the acoustic peaks in the temperature and polarization power spectra have been scaled, the days of splendid isolation of CMB theory, analysis, and experiment will have ended. Beyond and beneath the peaks lies a wealth of information about the evolution of structure in the Universe and its origin in the early universe. As CMB photons traverse the large-scale structure of the Universe on their journey from the recombination epoch, they pick up secondary temperature and polarization anisotropies. These depend on the intervening dark matter, dark energy, baryonic gas density, and temperature distributions, and even the existence of primordial gravity waves, so the potential payoff of their detection is enormous. The price for this extended reach is the loss of the ability both to make precise predictions, owing to uncertain and/or nonlinear physics, and to make precise measurements, owing to the cosmic variance of the primary anisotropies and the relatively greater importance of galactic and extragalactic foregrounds.

We begin in the following section with a discussion of the matter power spectrum to set the framework for the discussion of secondary anisotropies. Secondaries can be divided into two classes: those caused by gravitational effects and those induced by scattering off of electrons. The former are treated in “Gravitational

Secondaries" and the latter in "Scattering Secondaries." Secondary anisotropies are often non-Gaussian, so they show up not only in the power spectra of "Observables," above, but in higher point functions as well. We briefly discuss non-Gaussian statistics in "Non-Gaussianity," below. All of these topics are subjects of current research to which this review can only serve as an introduction.

Matter Power Spectrum

The same balance between pressure and gravity that is responsible for acoustic oscillations determines the power spectrum of fluctuations in nonrelativistic matter. This relationship is often obscured by focusing on the density fluctuations in the pressureless cold dark matter itself, so we instead consider the matter power spectrum from the perspective of the Newtonian potential.

PHYSICAL DESCRIPTION After recombination, without the pressure of the photons, the baryons simply fall into the Newtonian potential wells with the cold dark matter, an event usually referred to as the end of the Compton drag epoch. We claimed in "Radiation Driving," above, that above the horizon at matter-radiation equality the potentials are nearly constant. This follows from the following dynamics: Where pressure gradients are negligible, infall into some initial potential causes a potential flow of $v_{\rm tot} \sim (k\eta)\Psi_i$ (see Equation 19) and causes density enhancements by continuity of $\delta_{\rm tot} \sim -(k\eta)v_{\rm tot} \sim -(k\eta)^2\Psi_i$. The Poisson equation says that the potential at this later time $\Psi \sim -(k\eta)^{-2}\delta_{\rm tot} \sim \Psi_i$, so this rate of growth is exactly right to keep the potential constant. Formally, this Newtonian argument only applies in general relativity for a particular choice of coordinates (Bardeen 1980), but the rule of thumb is that if what is driving the expansion (including spatial curvature) can also cluster unimpeded by pressure, the gravitational potential will remain constant.

Because the potential is constant in the matter-dominated epoch, the large-scale observations of COBE set the overall amplitude of the potential power spectrum today. Translated into density, this is the well-known COBE normalization. It is usually expressed in terms of δ_H, the matter density perturbation at the Hubble scale today. Because the observed temperature fluctuation is approximately $\Psi/3$ (Sachs & Wolfe 1967),

$$\frac{\Delta_T^2}{T^2} \approx \frac{1}{9}\Delta_\Psi^2 \approx \frac{1}{4}\delta_H^2, \tag{25}$$

where the second equality follows from the Poisson equation in a fully matter-dominated universe with $\Omega_m = 1$. The observed COBE fluctuation of $\Delta_T \approx 28\ \mu$K (Smoot et al. 1992) implies $\delta_H \approx 2 \times 10^{-5}$. For corrections for $\Omega_m < 1$, where the potential decays because the dominant driver of the expansion cannot cluster, see Bunn & White (1997).

On scales below the horizon at matter-radiation equality, we have seen in "Radiation Driving," above, that pressure gradients from the acoustic oscillations themselves impede the clustering of the dominant component, i.e., the photons, and lead

to decay in the potential. Dark matter density perturbations remain but grow only logarithmically from their value at horizon crossing, which (just as for large scales) is approximately the initial potential, $\delta_m \approx -\Psi_i$. The potential for modes that have entered the horizon already will therefore be suppressed by $\Psi \propto -\delta_m/k^2 \sim \Psi_i/k^2$ at matter domination (neglecting the logarithmic growth), again according to the Poisson equation. The ratio of Ψ at late times to its initial value is called the *transfer function*. On large scales, then, the transfer function is close to one, and it falls off as k^{-2} on small scales. If the baryons fraction ρ_b/ρ_m is substantial, baryons alter the transfer function in two ways. First, their inability to cluster below the sound horizon causes further decay in the potential between matter-radiation equality and the end of the Compton drag epoch. Second, the acoustic oscillations in the baryonic velocity field kinematically cause acoustic wiggles in the transfer function (Hu & Sugiyama 1996). These wiggles in the matter power spectrum are related to the acoustic peaks in the CMB spectrum, like twins separated at birth, and are actively being pursued by the largest galaxy surveys (Percival et al. 2001). For fitting formulas for the transfer function that includes these effects see Eisenstein & Hu (1998).

COSMOLOGICAL IMPLICATIONS The combination of the COBE normalization, the matter transfer function, and the near scale-invariant initial spectrum of fluctuations tells us that by the present, fluctuations in the cold dark matter or baryon density fields will have gone nonlinear for all scales $k \gtrsim 10^{-1}h\,\mathrm{Mpc}^{-1}$. It is a great triumph of the standard cosmological paradigm that there is just enough growth between $z_* \approx 10^3$ and $z = 0$ to explain structures in the Universe across a wide range of scales.

In particular, because this nonlinear scale also corresponds to galaxy clusters, measurements of their abundance yield a robust measure of the power near this scale for a given matter density Ω_m. The agreement between the COBE normalization and the cluster abundance at low $\Omega_m \sim 0.3$–0.4 was pointed out immediately following the COBE result (e.g., White et al. 1993, Bartlett & Silk 1993) and is one of the strongest pieces of evidence for the parameters in the working cosmological model (Ostriker & Steinhardt 1995, Krauss & Turner 1995).

More generally, the comparison between large-scale structure and the CMB is important in that it breaks degeneracies between effects owing to deviations from power law initial conditions and the dynamics of the matter and energy contents of the Universe. Any dynamical effect that reduces the amplitude of the matter power spectrum corresponds to a decay in the Newtonian potential that boosts the level of anisotropy (see "Radiation Driving," above, and "Integrated Sachs-Wolfe Effect," below). Massive neutrinos are a good example of physics that drives the matter power spectrum down and the CMB spectrum up.

The combination is even more fruitful in the relationship between the acoustic peaks and the baryon wiggles in the matter power spectrum. Our knowledge of the physical distance between adjacent wiggles provides our best calibrated standard ruler for cosmology (Eisenstein et al. 1998). For example, at very low z the radial

distance out to a galaxy is cz/H_0. The unit of distance is therefore h^{-1} Mpc, and a knowledge of the true physical distance corresponds to a determination of h. At higher redshifts the radial distance depends sensitively on the background cosmology (especially the dark energy), so a future measurement of baryonic wiggles at, for instance, $z \sim 1$ would be a powerful test of dark energy models. To a lesser extent, the shape of the transfer function, which mainly depends on the matter-radiation scale in h Mpc^{-1}, i.e., $\Omega_m h$, is another standard ruler (see, e.g., Tegmark et al. 2001 for a recent assessment), more heralded than the wiggles but less robust owing to degeneracy with other cosmological parameters.

For scales corresponding to $k \gtrsim 10^{-1} h\,\mathrm{Mpc}^{-1}$, density fluctuations are nonlinear by the present. Numerical N-body simulations show that the dark matter is bound up in a hierarchy of virialized structures or halos (see Bertschinger 1998 for a review). The statistical properties of the dark matter and the dark matter halos have been extensively studied in the working cosmological model. Less certain are the properties of the baryonic gas. We see that both enter into the consideration of secondary CMB anisotropies.

Gravitational Secondaries

Gravitational secondaries arise from two sources: the differential redshift from time-variable metric perturbations (Sachs & Wolfe 1967) and gravitational lensing. There are many examples of the former, one of which we have already encountered in "Integral Approach" in the context of potential decay in the radiation-dominated era. Such gravitational potential effects are usually called the integrated Sachs-Wolfe (ISW) effect in linear perturbation theory (see next section), the Rees-Sciama ("Rees-Sciama and Moving Halo Effects," below) effect in the nonlinear regime, and the gravitational wave effect for tensor perturbations ("Gravitation Waves," below). Gravitational waves and lensing also produce B-modes in the polarization (see "Polarization," below), by which they may be distinguished from acoustic polarization.

INTEGRATED SACHS-WOLFE EFFECT As we have seen in the previous section, the potential on a given scale decays whenever the expansion is dominated by a component whose effective density is smooth on that scale. This occurs at late times in an $\Omega_m < 1$ model at the end of matter domination and the onset dark energy (or spatial curvature) domination. If the potential decays between the time a photon falls into a potential well and when it climbs out, it gets a boost in temperature of $\delta\Psi$ owing to the differential gravitational redshift and $-\delta\Phi \approx \delta\Psi$ owing to an accompanying contraction of the wavelength (see "Gravitational Forcing," above).

Potential decay owing to residual radiation was introduced in "Integral Approach," above, but that owing to dark energy or curvature at late times induces much different changes in the anisotropy spectrum. What makes the dark energy or curvature contributions different from those owing to radiation is the longer length of time over which the potentials decay, on order of the Hubble time today. Residual radiation produces its effect quickly, so the distance over which photons

feel the effect is much smaller than the wavelength of the potential fluctuation. Recall that this meant that $j_l(kD)$ in the integral in Equation 23 could be set to $j_l(kD_*)$ and removed from the integral. The final effect then is proportional to $j_l(kD_*)$ and adds in phase with the monopole.

The ISW projection, indeed the projection of all secondaries, is much different (see Figure 6). Because the duration of the potential change is much longer, photons typically travel through many peaks and troughs of the perturbation. This cancellation implies that many modes have virtually no impact on the photon temperature. The only modes that do have an impact are those with wavevectors perpendicular to the line of sight, so that along the line of sight the photon does not pass through crests and troughs. What fraction of the modes contributes to the effect then? For a given wavenumber k and line of sight instead of the full spherical shell at radius $4\pi k^2 dk$, only the ring $2\pi k dk$ with $\mathbf{k} \perp \mathbf{n}$ participates. Thus, the anisotropy induced is suppressed by a factor of k (or $\ell = kD$ in angular space). Mathematically, this arises in the line-of-sight integral of Equation 23 from the integral over the oscillatory Bessel function $\int dx\, j_\ell(x) \approx (\pi/2\ell)^{1/2}$ (see also Figure 6).

The ISW effect thus generically shows up only at the lowest ℓ's in the power spectrum (Kofman & Starobinskii 1985). This spectrum is shown in Figure 7 (*bottom left*). Secondary anisotropy predictions in this figure are for a model with $\Omega_{\rm tot} = 1, \Omega_\Lambda = 2/3, \Omega_b h^2 = 0.02, \Omega_m h^2 = 0.16, n = 1$, and $z_{\rm ri} = 7$ and an inflationary energy scale $E_i \ll 10^{16}$ GeV. The ISW effect is especially important in that it is extremely sensitive to the dark energy: its amount, equation of state, and clustering properties (Coble et al. 1997, Caldwell et al. 1998, Hu 1998). Unfortunately, being confined to the low multipoles, the ISW effect suffers severely from the cosmic variance in Equation 4 in its detectability. Perhaps more promising is its correlation with other tracers of the gravitational potential [e.g., X-ray background (Boughn et al. 1998) and gravitational lensing (see below)].

This type of cancellation behavior and corresponding suppression of small-scale fluctuations is a common feature of secondary temperature and polarization anisotropies from large-scale structure and is quantified by the Limber equation (Limber 1954) and its CMB generalization (Hu & White, 1996; Hu, 2000a). It is the central reason why secondary anisotropies tend to be smaller than the primary ones from $z_* \approx 10^3$ despite the intervening growth of structure.

REES-SCIAMA AND MOVING HALO EFFECTS The ISW effect is linear in the perturbations. Cancellation of the ISW effect on small scales leaves second-order and nonlinear analogues in its wake (Rees & Sciama 1968). From a single isolated structure, the potential along the line of sight can change not only from evolution in the density profile but more importantly from its bulk motion across the line of sight. In the context of clusters of galaxies, this is called the moving cluster effect (Birkinshaw & Gull 1983). More generally, the bulk motion of dark matter halos of all masses contributes to this effect (Tuluie & Laguna 1995, Seljak 1996b), and their clustering gives rise to a low level of anisotropies on a range of scales but is never the leading source of secondary anisotropies on any scale (see Figure 7, *bottom left*).

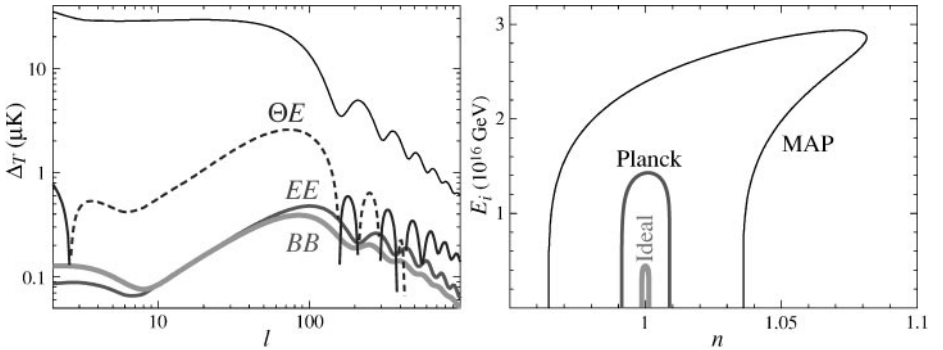

Figure 8 Gravitational waves and the energy scale of inflation E_i. (*Left*) Temperature and polarization spectra from an initial scale-invariant gravitational wave spectrum with power $\propto E_i^4 = (4 \times 10^{16}\ \text{GeV})^4$. (*Right*) 95% confidence upper limits statistically achievable on E_i and the scalar tilt n by the MAP and Planck satellites as well as an ideal experiment out to $\ell = 3000$ in the presence of gravitational lensing B-modes.

GRAVITATIONAL WAVES A time-variable tensor metric perturbation similarly leaves an imprint in the temperature anisotropy (Sachs & Wolfe 1967). A tensor metric perturbation can be viewed as a standing gravitational wave and produces a quadrupolar distortion in the spatial metric. If its amplitude changes, it leaves a quadrupolar distortion in the CMB temperature distribution (Polnarev 1985). Inflation predicts a nearly scale-invariant spectrum of gravitational waves. Their amplitude depends strongly on the energy scale of inflation,[3] (power $\propto E_i^4$) (Rubakov et al. 1982, Fabbri & Pollock 1983), and its relationship to the curvature fluctuations discriminates between particular models for inflation. Detection of gravitational waves in the CMB therefore provides our best hope to study the particle physics of inflation.

Gravitational waves, like scalar fields, obey the Klein-Gordon equation in a flat universe, and their amplitudes begin oscillating and decaying once the perturbation crosses the horizon. Although this process occurs even before recombination, rapid Thomson scattering destroys any quadrupole anisotropy that develops (see "Damping," above). This fact dicates the general structure of the contributions to the power spectrum (see Figure 8, *left panel*): They are enhanced at $\ell = 2$ or the present quadrupole and sharply suppressed at a multipole larger than that of the first peak (Abbott & Wise 1984, Starobinskii 1985, Crittenden et al. 1993). As is the case for the ISW effect, confinement to the low multipoles means the isolation of gravitational waves is severely limited by cosmic variance.

The signature of gravitational waves in the polarization is more distinct. Because gravitational waves cause a quadrupole temperature anisotropy at the end of recombination, they also generate a polarization. The quadrupole generated by

[3] $E_i^4 \equiv V(\phi)$, the potential energy density associated with the scalar field(s) driving inflation.

a gravitational wave has its main angular variation transverse to the wavevector itself (Hu & White 1997a). The resulting polarization has components directed both along or orthogonal to the wavevector and at 45° angles to it. Gravitational waves therefore generate a nearly equal amount of E and B mode polarization when viewed at a distance that is much greater than a wavelength of the fluctuation (Kamionkowski et al. 1997, Zaldarriaga & Seljak 1997). The B-component presents a promising means of measuring the gravitational waves from inflation and hence the energy scale of inflation (see Figure 8, *right panel*). Models of inflation correspond to points in the n, E_i plane (Dodelson et al. 1997). Therefore, the anticipated constraints discriminate among different models of inflation, probing fundamental physics at scales well beyond those accessible in accelerators.

GRAVITATIONAL LENSING The gravitational potentials of large-scale structure also lens the CMB photons. Because lensing conserves surface brightness, it only affects anisotropies and hence is second order in perturbation theory (Blanchard & Schneider 1987). The photons are deflected according to the angular gradient of the potential projected along the line of sight with a weighting of $2(D_* - D)/(D_* D)$. Again the cancellation of parallel modes implies that it is mainly the large-scale potentials that are responsible for deflections. Specifically, the angular gradient of the projected potential peaks at a multipole $\ell \sim 60$, corresponding to scales of a $k \sim$ few 10^{-2} Mpc^{-1} (Hu 2000b). The deflections are therefore coherent below the degree scale. The coherence of the deflection should not be confused with its rms value, which in the model of Figure 1 is a few arcminutes.

This large coherence and small amplitude ensures that linear theory in the potential is sufficient to describe the main effects of lensing. Because lensing is a one-to-one mapping of the source and image planes, it simply distorts the images formed from the acoustic oscillations in accord with the deflection angle. This warping naturally also distorts the mapping of physical scales in the acoustic peaks to angular scales (see "Integral Approach," above) and hence smooths features in the temperature and polarization (Seljak 1996a). The smoothing scale is the coherence scale of the deflection angle $\Delta\ell \approx 60$ and is sufficiently wide to alter the acoustic peaks with $\Delta\ell \sim 300$. The contributions, shown in Figure 7 (*bottom left*), are therefore negative (*dashed*) on scales corresponding to the peaks.

For the polarization, the remapping not only smooths the acoustic power spectrum but actually generates B-mode polarization (see Figure 1 and Zaldarriaga & Seljak 1998). Remapping by the lenses preserves the orientation of the polarization but warps its spatial distribution in a Gaussian random fashion and hence does not preserve the symmetry of the original E-mode. The B-modes from lensing set a detection threshold for gravitational waves for a finite patch of sky (Hu 2002).

Gravitational lensing also generates a small amount of power in the anisotropies on its own, but this is only noticable beyond the damping tail, where diffusion has destroyed the primary anisotropies (see Figure 7, *bottom left*). On these small scales, the anisotropy of the CMB is approximately a pure gradient on the sky, and the inhomogeneous distribution of lenses introduces ripples in the gradient

on the scale of the lenses (Seljak & Zaldarriaga 2000). In fact, the moving halo effect of "Rees-Sciama and Moving Halo Effects," above, can be described as the gravitational lensing of the dipole anisotropy owing to the peculiar motion of the halo (Birkinshaw & Gull, 1983).

Because the lensed CMB distribution is not linear in the fluctuations, it is not completely described by changes in the power spectrum. Much of the recent work in the literature has been devoted to utilizing the non-Gaussianity to isolate lensing effects (Bernardeau 1997, 1998; Zaldarriaga & Seljak 1999; Zaldarriaga 2000) and their cross-correlation with the ISW effect (Goldberg & Spergel 1999, Seljak & Zaldarriaga 1999). In particular, there is a quadratic combination of the anisotropy data that optimally reconstructs the projected dark matter potentials for use in this cross-correlation (Hu 2001b). The cross-correlation is especially important in that in a flat universe it is a direct indication of dark energy and can be used to study the properties of the dark energy beyond a simple equation of state (Hu 2002).

Scattering Secondaries

From the observations both of the lack of a Gunn-Peterson trough (Gunn & Peterson 1965) in quasar spectra and its preliminary detection (Becker et al. 2001), we know that hydrogen was reionized at $z_{\rm ri} \gtrsim 6$. This is thought to occur through the ionizing radiation of the first generation of massive stars (see, e.g., Loeb & Barkana 2001 for a review). The consequent recoupling of CMB photons to the baryons causes a few percent of them to be rescattered. Linearly, rescattering induces three changes to the photon distribution: suppression of primordial anisotropy, generation of large angle polarization, and a large angle Doppler effect. The latter two are suppressed on small scales by the cancellation highlighted in "Integrated Sachs-Wolfe Effect", above. Nonlinear effects can counter this suppression; these are the subject of active research and are outlined in "Modulated Doppler Effect," below.

PEAK SUPPRESSION Like scattering before recombination, scattering at late times suppresses anisotropies in the distributions that have already formed. Reionization therefore suppresses the amplitude of the acoustic peaks by the fraction of photons rescattered, approximately the optical depth $\sim\tau_{\rm ri}$ (see Figure 7, *bottom right*, *dotted line* and negative, *dashed line*; contributions corresponding to $|\delta\Delta_T^2|^{1/2}$ between the $z_{\rm ri}=7$ and $z_{\rm ri}=0$ models). Unlike the plasma before recombination, the medium is optically thin, so the mean free path and diffusion length of the photons is of order the horizon itself. New acoustic oscillations cannot form. On scales approaching the horizon at reionization, inhomogeneities have yet to be converted into anisotropies (see "Integral Approach," above) so large angle fluctuations are not suppressed. Whereas these effects are relatively large compared with the expected precision of future experiments, they mimic a change in the overall normalization of fluctuations except at the lowest, cosmic variance limited, multipoles.

LARGE-ANGLE POLARIZATION The rescattered radiation becomes polarized because, as discussed in "Integral Approach," temperature inhomogeneities become anisotropies by projection, passing through quadrupole anisotropies when the perturbations are on the horizon scale at any given time. The result is a bump in the power spectrum of the *E*-polarization on angular scales corresponding to the horizon at reionization (see Figure 1). Because of the low optical depth of reionization and the finite range of scales that contribute to the quadrupole, the polarization contributions are on the order of tenths of μK on scales of $\ell \sim$ few. In a perfect, foreground free world, this is not beyond the reach of experiments and can be used to isolate the reionization epoch (Hogan et al. 1982, Zaldarriaga et al. 1997). As in the ISW effect, cancellation of contributions along the line of sight guarantees a sharp suppression of contributions at higher multipoles in linear theory. Spatial modulation of the optical depth owing to density and ionization (see "Modulated Doppler Effects," below) does produce higher order polarization but at an entirely negligible level in most models (Hu 2000a).

DOPPLER EFFECT Naively, velocity fields of order $v \sim 10^{-3}$ ($c = 1$; see, e.g., Strauss & Willick 1995 for a review) and optical depths of a few percent would imply a Doppler effect that rivals the acoustic peaks themselves. That this is not the case is the joint consequence of the cancellation described in "Integrated Sachs-Wolfe Effect," above, and the fact that the acoustic peaks are not "Doppler peaks" (see "Integral Approach," above). Because the Doppler effect comes from the peculiar velocity along the line of sight, it retains no contributions from linear modes with wavevectors perpendicular to the line of sight. However, as we have seen, these are the only modes that survive cancellation (see Figure 6 and Kaiser 1984). Consequently, the Doppler effect from reionization is strongly suppressed and is entirely negligible below $\ell \sim 10^2$ unless the optical depth in the reionization epoch approaches unity (see Figure 7, *bottom right*).

MODULATED DOPPLER EFFECTS The Doppler effect can survive cancellation if the optical depth has modulations in a direction orthogonal to the bulk velocity. This modulation can be the result of either density or ionization fluctuations in the gas. Examples of the former include the effect in clusters and linear as well as nonlinear large-scale structures.

Cluster modulation The strongly nonlinear modulation provided by the presence of a galaxy cluster and its associated gas leads to the kinetic Sunyaev-Zel'dovich effect. Cluster optical depths on order of 10^{-2} and peculiar velocities of 10^{-3} imply signals in the 10^{-5} regime in individual arcminute-scale clusters, which are of course rare objects. Although this signal is reasonably large, it is generally dwarfed by the thermal Sunyaev-Zel'dovich effect (see below) and has yet to be detected with high significance (see Carlstrom et al. 2001 and references therein). The kinetic Sunyaev-Zel'dovich effect has negligible impact on the power spectrum of anisotropies, owing to the rarity of clusters, and can be included as part of the general density modulation effect below.

Linear modulation At the opposite extreme, linear density fluctuations modulate the optical depth and give rise to a Doppler effect, as pointed out by Ostriker & Vishniac (1986) and calculated by Vishniac (1987) (see also Efstathiou & Bond 1987). The result is a signal at the μK level peaking at $\ell \sim$ few $\times 10^3$ that increases roughly logarithmically with the reionization redshift (see Figure 7, *bottom right*).

General density modulation Both the cluster and linear modulations are limiting cases of the more general effect of density modulation by the large-scale structure of the Universe. For the low reionization redshifts currently expected ($z_{\rm ri} \approx 6$–7), most of the effect comes neither from clusters nor the linear regime but from intermediate-scale dark matter halos. An upper limit to the total effect can be obtained by assuming the gas density traces the dark matter density (Hu 2000a) and implies signals on the order of $\Delta_T \sim$ few μK at $\ell > 10^3$ (see Figure 7, *bottom right*). Based on simulations, this assumption should hold in the outer profiles of halos (Pearce et al. 2001, Lewis et al. 2000), but gas pressure will tend to smooth out the distribution in the cores of halos and reduce small scale contributions. In the absence of substantial cooling and star formation, these net effects can be modeled under the assumption of hydrostatic equilibrium (Komatsu & Seljak 2001) in the halos and included in a halo approach to the gas distribution (Cooray 2001).

Ionization modulation Finally, optical depth modulation can also come from variations in the ionization fraction (Aghanim et al. 1996, Gruzinov & Hu 1998, Knox et al. 1998). Predictions for this effect are the most uncertain, as it involves both the formation of the first ionizing objects and the subsequent radiative transfer of the ionizing radiation (Bruscoli et al. 2000, Benson et al. 2001). It is, however, unlikely to dominate the density-modulated effect except perhaps at very high multipoles of $\ell \sim 10^4$ (crudely estimated, following Gruzinov & Hu 1998) (Figure 7, *bottom right*).

SUNYAEV-ZEL'DOVICH EFFECT Internal motion of the gas in dark matter halos also gives rise to Doppler shifts in the CMB photons. As in the linear Doppler effect, shifts that are first order in the velocity are canceled as photons scatter off of electrons moving in different directions. At second order in the velocity, there is a residual effect. For clusters of galaxies where the temperature of the gas can reach $T_e \sim 10$ keV, the thermal motions are a substantial fraction of the speed of light $v_{\rm rms} = (3\ T_e/m_e)^{1/2} \sim 0.2$. The second-order effect represents a net transfer of energy between the hot electron gas and the cooler CMB and leaves a spectral distortion in the CMB where photons on the Rayleigh-Jeans side are transferred to the Wien tail. This effect is called the thermal Sunyaev-Zel'dovich (SZ) effect (Sunyaev & Zel'dovich 1972). Because the net effect is of order $\tau_{\rm cluster}\, T_e/m_e \propto n_e T_e$, it is a probe of the gas pressure. Like all CMB effects, once imprinted, distortions relative to the redshifting background temperature remain unaffected by cosmological

dimming, so one might hope to find clusters at high redshift using the SZ effect. However, the main effect comes from the most massive clusters because of the strong temperature weighting, and these have formed only recently in the standard cosmological model.

Great strides have recently been made in observing the SZ effect in individual clusters, following pioneering attempts that spanned two decades (Birkinshaw 1999). The theoretical basis has remained largely unchanged save for small relativistic corrections as T_e/m_e approches unity. Both developments are comprehensively reviewed by Carlstrom et al. (2001). Here we instead consider the implications of the SZ effect as a source of secondary anisotropies.

The SZ effect from clusters provides the most substantial contribution to temperature anisotropies beyond the damping tail. On scales much larger than an arcminute, in which clusters are unresolved, contributions to the power spectrum appear as uncorrelated shot noise ($C_\ell = \text{const.}$ or $\Delta_T \propto \ell$). The additional contribution owing to the spatial correlation of clusters turns out to be almost negligible in comparison because of the rarity of clusters (Komatsu & Kitayama 1999). Below this scale, contributions turn over as the clusters become resolved. Though there has been much recent progress in simulations (Refregier et al. 2000, Seljak et al. 2001, Springel et al. 2001), dynamic range still presents a serious limitation.

Much recent work has been devoted to semi-analytic modeling following the technique of Cole & Kaiser (1988), in which the SZ correlations are described in terms of the pressure profiles of clusters, their abundance, and their spatial correlations [now commonly referred to as an application of the "halo model" (see Komatsu & Kitayama 1999, Atrio-Barandela & Mücket 1999, Cooray 2001, Komatsu & Seljak 2001)]. We show the predictions of a simplified version in Figure 7, (*bottom right*, Rayleigh-Jeans limit), where the pressure profile is approximated by the dark matter halo profile and the virial temperature of the halo. Although this treatment is comparatively crude, the inaccuracies that result are dwarfed by "missing physics" in both the simulations and more sophisticated modeling, e.g., the nongravitational sources and sinks of energy that change the temperature and density profile of the cluster, often modeled as a uniform "preheating" of the intercluster medium (Holder & Carlstrom 2001).

Although the SZ effect is expected to dominate the power spectrum of secondary anisotropies, it does not necessarily make the other secondaries unmeasurable or contaminate the acoustic peaks. Its distinct frequency signature can be used to isolate it from other secondaries (see, e.g., Cooray et al. 2000). Additionally, it mainly comes from massive clusters that are intrinsically rare. Hence, contributions to the power spectrum are non-Gaussian and concentrated in rare, spatially localized regions. Removal of regions identified as clusters through X rays and optical surveys or ultimately through high-resolution CMB maps themselves can greatly reduce contributions at large angular scales at which they are unresolved (Persi et al. 1995, Komatsu & Kitayama 1999).

Non-Gaussianity

As we have seen, most of the secondary anisotropies are not linear in nature and hence produce non-Gaussian signatures. Non-Gaussianity in the lensing and SZ signals is important for their isolation. The same is true for contaminants such as galactic and extragalactic foregrounds. Finally, the lack of an initial non-Gaussianity in the fluctuations is a testable prediction of the simplest inflationary models (Guth & Pi 1985, Bardeen et al. 1983). Consequently, non-Gaussianity in the CMB is a very active field of research. The primary challenge in studies of non-Gaussianity is in choosing the statistic that quantifies it. Non-Gaussianity tells us what the distribution is not, not what it is. The secondary challenge is to optimize the statistic against the Gaussian "noise" of the primary anisotropies and instrumental or astrophysical systematics.

Early theoretical work on the bispectrum, the harmonic analogue of the three-point function, addressed its detectability in the presence of the cosmic variance of the Gaussian fluctuations (Luo 1994) and showed that the inflationary contribution is not expected to be detectable in most models (Allen et al. 1987, Falk et al. 1993, Gangui et al. 1994). The bispectrum is defined by a triplet of multipoles, or configuration, that defines a triangle in harmonic space. The large cosmic variance in an individual configuration is largely offset by the great number of possible triplets. Interest was spurred by reports of significant signals in specific bispectrum configurations in the COBE maps (Ferreira et al. 1998) that turned out to be caused by systematic errors (Banday et al. 2000). Recent investigations have focused on the signatures of secondary anisotropies (Goldberg & Spergel 1999, Cooray & Hu 2000). These turned out to be detectable with experiments that have both high resolution and angular dynamic range but require the measurement of a wide range of configurations of the bispectrum. Data analysis challenges for measuring the full bispectrum largely remain to be addressed (see Heavens 1998, Spergel & Goldberg 1999, Phillips & Kogut 2001).

The trispectrum, the harmonic analogue of the four-point function, also has advantages for the study of secondary anisotropies. Its great number of configurations are specified by a quintuplet of multipoles that correspond to the sides and diagonal of a quadrilateral in harmonic space (Hu 2001a). The trispectrum is important in that it quantifies the covariance of the power spectrum across multipoles that are often very strong in nonlinear effects, e.g., the SZ effect (Cooray 2001). It is also intimately related to the power spectra of quadratic combinations of the temperature field and has been applied to study gravitational lensing effects (Bernardeau 1997, Zaldarriaga 2000, Hu 2001a).

The bispectrum and trispectrum quantify non-Gaussianity in harmonic space and have clear applications for secondary anisotropies. Tests for non-Gaussianity localized in angular space include the Minkowski functionals (including the genus) (Winitzki & Kosowsky 1997), the statistics of temperature extrema (Kogut et al. 1996), and wavelet coefficients (Aghanim & Forni 1999). These may be more useful for examining foreground contamination and trace amounts of topological defects.

DATA ANALYSIS

The very large CMB data sets that have begun arriving require new, innovative tools of analysis. The fundamental tool for analyzing CMB data—the likelihood function—has been used since the early days of anisotropy searches (Readhead et al. 1989, Bond et al. 1991, Dodelson & Jubas 1993). Brute-force likelihood analyses (Tegmark & Bunn 1995) were performed even on the relatively large COBE data set, with 6000 pixels in its map. Present data sets are a factor of 10 larger, and this factor will soon increase by yet another factor of 100. The brute-force approach, the time for which scales as the number of pixels cubed, no longer suffices.

In response, analysts have devised a host of techniques that move beyond the early brute-force approach. The simplicity of CMB physics—owing to linearity—is mirrored in analysis by the apparent Gaussianity of both the signal and many sources of noise. In the Gaussian limit, "optimal statistics" are easy to identify. These compress the data so that all of the information is retained, but the subsequent analysis—because of the compression—becomes tractable.

The Gaussianity of the CMB is not shared by other cosmological systems because gravitational nonlinearities turn an initially Gaussian distribution into a non-Gaussian one. Nontheless, many of the techniques devised to study the CMB have been proposed for studying the 3D galaxy distribution (Tegmark et al. 1998), the 2D galaxy distribution (Efstathiou & Moody 2001, Huterer et al. 2001), the Lyman alpha forest (Hui et al. 2001), and the shear field from weak lensing (Hu & White 2001), among others. Indeed, these techniques are now indispensible, powerful tools for all cosmologists, and we would be remiss not to at least outline them in a disussion of the CMB, the context in which many of them were developed.

Figure 9 summarizes the path from the data analyis starting point, a timestream of data points, to the end, the determination of cosmological parameters. Preceding this starting point comes the calibration and the removal of systematic errors from the raw data, but being experiment specific, we do not attempt to cover such

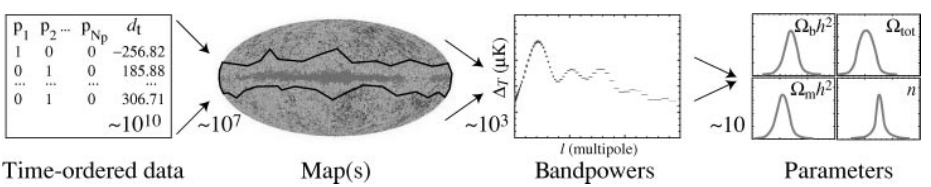

Figure 9 Data pipeline and radical compression. Maps are constructed for each frequency channel from the data timestreams, combined, and cleaned of foreground contamination by spatial (represented here by excising the galaxy) and frequency information. Bandpowers are extracted from the maps and cosmological parameters from the bandpowers. Each step involves a substantial reduction in the number of parameters needed to describe the data, from potentially $10^{10} \rightarrow 10$ for the Planck satellite.

issues here.[4] Each step radically compresses the data by reducing the number of parameters used to describe it. Although this data pipeline and our discussion below focus on temperature anisotropies, similar steps have been elucidated for polarization (Bunn 2001, Tegmark & de Oliveira-Costa 2001, Lewis et al. 2001).

Mapmaking

An experiment can be characterized by the data d_t taken at many different times; a "pointing matrix," P_{ti}, relating the data timestream to the underlying signal at pixelized positions indexed by i; and a noise matrix $C_{d,tt'}$ characterizing the covariance of the noise in the timestream. A model for the data then is $d_t = P_{ti}\Theta_i + n_t$ (with implicit sum over the repeating index i); it is the sum of signal plus noise. Here n_t is drawn from a distribution (often Gaussian) with mean zero and covariance $\langle n_t n_{t'}\rangle = C_{d,tt'}$. In its simplest form the pointing matrix $\mathbf{P}$ contains rows—which correspond to a particular time—with all zeroes in it except for one column with a one (see Figure 9). That column corresponds to the particular pixel observed at the time of interest. Typically, a pixel will be scanned many times during an experiment, so a given column will have many ones in it, corresponding to the many times the pixel has been observed.

Given this model, a well-posed question is, What is the optimal estimator for the signal Θ_i? In effect, what is the best way to construct a map? The answer stems from the likelihood function $\mathcal{L}$, defined as the probability of getting the data, given the theory $\mathcal{L} \equiv P[\text{data}|\text{theory}]$. In this case the theory is the set of parameters Θ_i,

$$\mathcal{L}_\Theta(d_t) = \frac{1}{(2\pi)^{N_t/2}\sqrt{\det \mathbf{C}_d}} \exp\left[-\frac{1}{2}\,(d_t - P_{ti}\Theta_i)\, C^{-1}_{d,tt'}(d_{t'} - P_{t'j}\Theta_j)\right]. \quad (26)$$

That is, the noise, the difference between the data and the modulated signal, is assumed to be Gaussian with covariance $\mathbf{C}_d$.

Two important theorems are useful in the construction of a map and more generally in each step of the data pipeline (Tegmark et al. 1997). The first is Bayes' theorem. In this context, it says that $P[\Theta_i|d_t]$, the probability that the temperatures are equal to Θ_i given the data, is proportional to the likelihood function times a prior $P(\Theta_i)$. Thus, with a uniform prior,

$$P[\Theta_i|d_t] \propto P[d_t|\Theta_i] \equiv \mathcal{L}_\Theta(d_t), \quad (27)$$

with the normalization constant determined by requiring the integral of the probability over all Θ_i to be equal to one. The probability on the left is the one of interest. The most likely values of Θ_i therefore are those that maximize the likelihood function. Because the log of the likelihood function in question, Equation 26, is quadratic in the parameters Θ_i, it is straightforward to find this maximum point. Differentiating the argument of the exponential with respect to Θ_i and setting to

[4]Aside from COBE, experiments to date have had a sizable calibration error (∼5–10%) that must be factored into the interpretation of Figure 1.

zero leads immediately to the estimator

$$\hat{\Theta}_i = C_{N,ij} P_{jt} C^{-1}_{d,tt'} d_{t'}, \tag{28}$$

where $\mathbf{C}_N \equiv (\mathbf{P}^{\rm tr} \mathbf{C}_d^{-1} \mathbf{P})^{-1}$. As the notation suggests, the mean of the estimator is equal to the actual Θ_i, and the variance is equal to $\mathbf{C}_N$.

The second theorem states that this maximum likelihood estimator is also the minimum variance estimator. The Cramer-Rao inequality says no estimator can measure the Θ_i with errors smaller than the diagonal elements of $\mathbf{F}^{-1}$, where the Fisher matrix is defined as

$$F_{\Theta,ij} \equiv \left\langle -\frac{\partial^2 \ln \mathcal{L}_\Theta}{\partial \Theta_i \partial \Theta_j} \right\rangle. \tag{29}$$

Inspection of Equation 26 shows that, in this case, the Fisher matrix is precisely equal to $\mathbf{C}_N^{-1}$. Therefore, the Cramer-Rao theorem implies that the estimator of Equation 28 is optimal: It has the smallest possible variance (Tegmark 1997a). No information is lost if the map is used in subsequent analysis instead of the timestream data, but huge factors of compression have been gained. For example, in the recent Boomerang experiment (Netterfield et al. 2001), the timestream contained 2×10^8 numbers, and the map had only 57,000 pixels. The map resulted in compression by a factor of 3500.

There are numerous complications that must be dealt with in realistic applications of Equation 28. Perhaps the most difficult is estimation of $\mathbf{C}_d$, the timestream noise covariance. This typically must be done from the data itself (Ferreira & Jaffe 2000, Stompor et al. 2001). Even if $\mathbf{C}_d$ were known perfectly, evaluation of the map involves inverting $\mathbf{C}_d$, a process that scales as the number of raw data points cubed. For both of these problems, the assumed "stationarity" of $C_{d,tt'}$ (it depends only on $t - t'$) is of considerable utility. Iterative techniques to approximate matrix inversion can also assist in this process (Wright et al. 1996). Another issue that has received much attention is the choice of pixelization. The CMB community has converged on the Healpix pixelization scheme (http://www.eso.org/science/healpix/), now freely available.

Perhaps the most dangerous complication arises from astrophysical foregrounds, both within and from outside the Galaxy, the main ones being synchrotron, bremmsstrahlung, dust, and point source emission. All of the main foregrounds have spectral shapes different than the blackbody shape of the CMB. Modern experiments typically observe at several different frequencies, so a well-posed question is, How can we best extract the CMB signal from the different frequency channels (Bouchet & Gispert 1999)? The blackbody shape of the CMB relates the signal in all the channels, leaving one free parameter. Similarly, if the foreground shapes are known, each foreground comes with just one free parameter per pixel. A likelihood function for the data can again be written down and the best estimator for the CMB amplitude determined analytically. In the absence of foregrounds, one would extract the CMB signal by weighting the frequency channels according to inverse noise. When foregrounds are present, the optimal combination of different

frequency maps is a more clever weighting that subtracts out the foreground contribution (Dodelson 1997). One can do better if the pixel-to-pixel correlations of the foregrounds can also be modeled from power spectra (Tegmark & Efstathiou 1996) or templates derived from external data.

This picture is complicated somewhat because the foreground shapes are not precisely known, varying across the sky, e.g., from a spatially varying dust temperature. This too can be modeled in the covariance and addressed in the likelihood analysis (Tegmark 1998, White 1998). The resulting cleaned CMB map is obviously noisier than if foregrounds were not around, but the multiple channels keep the degradation managable. For example, the errors on some cosmological parameters coming from Planck may degrade by almost a factor of 10 as compared with the no-foreground case. However, many errors will not degrade at all, and even the degraded parameters will still be determined with unprecedented precision (Knox 1999, Prunet et al. 2000, Tegmark et al. 2000).

Many foregrounds tend to be highly non-Gaussian, in particular well-localized regions of the map. These pixels can be removed from the map as was done for the region around the galactic disk for COBE. This technique can also be highly effective against point sources. Indeed, even if there is only one frequency channel, external foreground templates set the form of the additional contributions to $\mathbf{C}_N$, which, when properly included, immunize the remaining operations in the data pipeline to such contaminants (Bond et al. 1998). The same technique can be used with templates of residual systematics or constraints imposed on the data, from, e.g., the removal of a dipole.

Bandpower Estimation

Figure 9 indicates that the next step in the compression process is extracting bandpowers from the map. What is a bandpower and how can it be extracted from the map? To answer these questions, we must construct a new likelihood function, one in which the estimated Θ_i are the data. No theory predicts an individual Θ_i, but all predict the distribution from which the individual temperatures are drawn. For example, if the theory predicts Gaussian fluctuations, then Θ_i is distributed as a Gaussian with mean zero and covariance equal to the sum of the noise covariance matrix C_N and the covariance due to the finite sample of the cosmic signal C_S. Inverting Equation 1 and using Equation 2 for the ensemble average leads to

$$C_{S,ij} \equiv \langle \Theta_i \Theta_j \rangle = \sum_\ell \Delta^2_{T,\ell} W_{\ell,ij}, \tag{30}$$

where $\Delta^2_{T,\ell}$ depends on the theoretical parameters through C_ℓ (see Equation 3). Here W_ℓ, the window function, is proportional to the Legendre polynomial $P_\ell(\hat{n}_i \cdot \hat{n}_j)$ and a beam and pixel smearing factor b^2_ℓ. For example, a Gaussian beam of width σ dictates that the observed map is actually a smoothed picture of true signal, insensitive to structure on scales smaller than σ. If the pixel scale is much smaller than the beam scale, $b^2_\ell \propto e^{-\ell(\ell+1)\sigma^2}$. Techniques for handling

asymmetric beams have also recently been developed (Wu et al. 2001, Wandelt & Gorski 2001, Souradeep & Ratra 2001, Fosalba et al. 2001). Using bandpowers corresponds to assuming that $\Delta^2_{T,\ell}$ is constant over a finite range, or band, of ℓ, equal to B_a for $\ell_a - \delta\ell_a/2 < \ell < \ell_a + \delta\ell_a/2$. Figure 1 gives a sense of the width and number of bands N_b probed by existing experiments.

For Gaussian theories, then, the likelihood function is

$$\mathcal{L}_B(\Theta_i) = \frac{1}{(2\pi)^{N_p/2}\sqrt{\det \mathbf{C}_\Theta}} \exp\left(-\frac{1}{2}\Theta_i C^{-1}_{\Theta,ij}\Theta_j\right), \tag{31}$$

where $\mathbf{C}_\Theta = \mathbf{C}_S + \mathbf{C}_N$ and N_p is the number of pixels in the map. As before, $\mathcal{L}_B$ is Gaussian in the anisotropies Θ_i, but in this case Θ_i are not the parameters to be determined; the theoretical parameters are the B_a, upon which the covariance matrix depends. Therefore, the likelihood function is not Gaussian in the parameters, and there is no simple, analytic way to find the point in parameter space (which is multidimensional depending on the number of bands being fit) at which $\mathcal{L}_B$ is a maximum. An alternative is to evaluate $\mathcal{L}_B$ numerically at many points in a grid in parameter space. The maximum of the $\mathcal{L}_B$ on this grid then determines the best fit values of the parameters. Confidence levels on, e.g., B_1 can be determined by finding the region within which $\int_a^b dB_1[\Pi_{i=2}^{N_b}\int dB_i]\mathcal{L}_B = 0.95$, e.g., for 95% limits.

This possibility is no longer viable owing to the sheer volume of data. Consider the Boomerang experiment with $N_p = 57{,}000$. A single evaluation of $\mathcal{L}_B$ involves computation of the inverse and determinant of the $N_p \times N_p$ matrix $\mathbf{C}_\Theta$, both of which scale as N_p^3. Whereas this single evaluation might be possible with a powerful computer, a single evaluation does not suffice. The parameter space consists of $N_b = 19$ bandpowers equally spaced from $l_a = 100$ up to $l_a = 1000$. A blindly placed grid on this space would require at least 10 evaluations in each dimension, so the time required to adequately evaluate the bandpowers would scale as $10^{19}N_p^3$. No computer can do this. The situation is rapidly getting worse (better) because Planck will have of order 10^7 pixels and be sensitive to of order 10^3 bands.

It is clear that a "smart" sampling of the likelihood in parameter space is necessary. The numerical problem, searching for the local maximum of a function, is well posed, and a number of search algorithms might be used. $\mathcal{L}_B$ tends to be sufficiently structureless that these techniques suffice. Bond et al. (1998) proposed the Newton-Raphson method, which has become widely used. One expands the derivative of the log of the likelihood function—which vanishes at the true maximum of $\mathcal{L}_B$—around a trial point in parameter space, $B_a^{(0)}$. Keeping terms second order in $B_a - B_a^{(0)}$ leads to

$$\hat{B}_a = \hat{B}_a^{(0)} + \hat{F}^{-1}_{B,ab}\frac{\partial \ln \mathcal{L}_B}{\partial B_b}, \tag{32}$$

where the curvature matrix $\hat{F}_{B,ab}$ is the second derivative of $-\ln \mathcal{L}_B$ with respect to B_a and B_b. Note the subtle distinction between the curvature matrix and the Fisher matrix in Equation 29, $\mathbf{F} = \langle\hat{\mathbf{F}}\rangle$. In general, the curvature matrix depends

on the data, on the Θ_i. In practice, though, analysts typically use the inverse of the Fisher matrix in Equation 32. In that case, the estimator becomes

$$\hat{B}_a = \hat{B}_a^{(0)} + \frac{1}{2} F^{-1}_{B,ab} \left(\Theta_i C^{-1}_{\Theta,ij} \frac{\partial C_{\Theta,jk}}{\partial B_b} C^{-1}_{\Theta,ki} \Theta_i - C^{-1}_{\Theta,ij} \frac{\partial C_{\Theta,ji}}{\partial B_b} \right), \quad (33)$$

quadratic in the data Θ_i. The Fisher matrix is equal to

$$F_{B,ab} = \frac{1}{2} C^{-1}_{\Theta,ij} \frac{\partial C_{\Theta,jk}}{\partial B_a} C^{-1}_{\Theta,kl} \frac{\partial C_{\Theta,li}}{\partial B_b}. \quad (34)$$

In the spirit of the Newton-Raphson method, Equation 33 is used iteratively but often converges after just a handful of iterations. The usual approximation is then to take the covariance between the bands as the inverse of the Fisher matrix evaluated at the convergent point $\mathbf{C}_B = \mathbf{F}_B^{-1}$. Indeed, Tegmark (1997b) derived the identical estimator by considering all unbiased quadratic estimators and identifying this one as the one with the smallest variance.

Although the estimator in Equation 33 represents a $\sim 10^{Nb}$ improvement over brute force coverage of the parameter space—converging in just several iterations—it still requires operations that scale as N_p^3. One means of speeding up the calculations is to transform the data from the pixel basis to the so-called signal-to-noise basis, based on an initial guess as to the signal and throwing out those modes that have low signal to noise (Bond 1995, Bunn & Sugiyama 1995). The drawback is that this procedure still requires at least one N_p^3 operation and potentially many, as the guess at the signal improves by iteration. Methods to truly avoid this prohibitive N_p^3 scaling (Oh et al. 1999, Wandelt & Hansen 2001) have been devised for experiments with particular scan strategies, but the general problem remains. A potentially promising approach involves extracting the real space correlation functions as an intermediate step between the map and the bandpowers (Szapudi et al. 2001). Another involves consistently analyzing coarsely pixelized maps with finely pixelized submaps (Dore et al. 2001).

Cosmological Parameter Estimation

The huge advantage of bandpowers is that they represent the natural meeting ground of theory and experiment. The above two sections outline some of the steps involved in extracting them from the observations. Once they are extracted, any theory can be compared with the observations without knowledge of experimental details. The simplest way to estimate the cosmological parameters $\{c_i\}$ is to approximate the likelihood as

$$\mathcal{L}_c(\hat{B}_a) \approx \frac{1}{(2\pi)^{N_c/2}\sqrt{\det \mathbf{C}_B}} \exp\left[-\frac{1}{2} (\hat{B}_a - B_a) C^{-1}_{B,ab} (\hat{B}_b - B_b) \right] \quad (35)$$

and evaluate it at many points in parameter space (the bandpowers depend on the cosmological parameters). Because the number of cosmological parameters in the working model is $N_c \sim 10$, this represents a final radical compression of

information in the original timestream, which recall has up to $N_t \sim 10^{10}$ data points.

In the approximation that the bandpower covariance $\mathbf{C}_B$ is independent of the parameters c, maximizing the likelihood is the same as minimizing χ^2 the argument of the exponential. This has been done by dozens of groups over the past few years, especially since the release of CMBFAST (Seljak & Zaldarriaga, 1996), which allows fast computation of theoretical spectra. Even after all the compression summarized in Figure 9, these analyses are still computationally cumbersome owing to the large numbers of parameters varied. Various methods of speeding up spectra computation have been proposed (Tegmark & Zaldarriaga 2000), based on the understanding of the physics of peaks outlined in "Acoustic Peaks," above, and Monte Carlo explorations of the likelihood function (Christensen et al. 2001).

Again the inverse Fisher matrix gives a quick and dirty estimate of the errors. Here the analogue of Equation 29 for the cosmological parameters becomes

$$F_{c,ij} = \frac{\partial B_a}{\partial c_i} C^{-1}_{B,ab} \frac{\partial B_b}{\partial c_j}. \tag{36}$$

In fact, this estimate has been widely used to forecast the optimal errors on cosmological parameters, given a proposed experiment and a band covariance matrix C_B that includes diagonal sample and instrumental noise variance. The reader should be aware that no experiment to date has even come close to achieving the precision implied by such a forecast!

As we enter the age of precision cosmology, a number of caveats will become increasingly important. No theoretical spectra are truly flat in a given band, so the question of how to weight a theoretical spectrum to obtain B_a can be important. In principle, one must convolve the theoretical spectra with window functions (Knox 1999) distinct from those in Equation 30 to produce B_a. Among recent experiments, Degree Angular Scale Interferometer (DASI) (Pryke et al. 2001), among others, have provided these functions. Another complication arises because the true likelihood function for B_a is not Gaussian, i.e., not of the form in Equation 35. The true distribution is skewed: The cosmic variance of Equation 4 leads to larger errors for an upward fluctuation than for a downward fluctuation. The true distribution is closer to log-normal (Bond et al. 2000), and several groups have already accounted for this in their parameter extractions.

DISCUSSION

Measurements of the acoustic peaks in the CMB temperature spectrum have already shown that the Universe is nearly spatially flat and began with a nearly scale-invariant spectrum of curvature fluctuations, consistent with the simplest of inflationary models. In a remarkable confirmation of a prediction of big bang nucleosynthesis, the CMB measurements have now verified that baryons account for about 4% of the critical density. Further, they suggest that the matter

density is about 10 times higher than this, implying the existence of nonbaryonic dark matter and dark energy.

Future measurements of the morphology of the peaks in the temperature and polarization should determine the baryonic and dark matter content of the Universe with exquisite precision. Beyond the peaks, gravitational wave imprint on the polarization, the gravitational lensing of the CMB, and gravitational and scattering secondary anisotropies hold the promise of elucidating the physics of inflation and the impact of dark energy on structure formation.

The once and future success of the CMB anisotropy enterprise rests on three equally important pillars: advances in experimental technique, precision in theory, and development of data analysis techniques. The remarkable progress in the field over the past decade owes much to the efforts of researchers in all three disciplines. That much more effort will be required to fulfill the bright promise of CMB suggests that the field will remain active and productive for years to come.

ACKNOWLEDGMENTS

W.H. thanks Fermilab, where this review was written, for their hospitality. W.H. was supported by NASA NAG5-10840 and the DOE OJI program. S.D. was supported by the DOE, by NASA grant NAG 5-10842 at Fermilab, and by NSF Grant PHY-0079251 at Chicago.

The *Annual Review of Astronomy and Astrophysics* is online at http://astro.annualreviews.org

LITERATURE CITED

Abbott LF, Schaefer RK. 1986. *Ap. J.* 308:546–62

Abbott LF, Wise MB. 1984. *Nucl. Phys.* B244: 541–48

Aghanim N, Desert FX, Puget JL, Gispert R. 1996. *Astron. Astrophys.* 311:1–11

Aghanim N, Forni O. 1999. *Astron. Astrophys.* 347:409–18

Albrecht A, Coulson D, Ferreira P, Magueijo J. 1996. *Phys. Rev. Lett.* 76:1413–16

Allen B, Caldwell RR, Dodelson S, Knox L, Shellard EPS, et al. 1997. *Phys. Rev. Lett.* 79:2624–27

Allen TJ, Grinstein B, Wise MB. 1987. *Phys. Lett.* B197:66–70

Atrio-Barandela F, Mücket JP. 1999. *Ap. J.* 515:465–70

Banday AJ, Zaroubi S, Górski KM. 2000. *Ap. J.* 533:575–87

Bardeen JM. 1980. *Phys. Rev. D* 22:1882–905

Bardeen JM, Steinhardt PJ, Turner MS. 1983. *Phys. Rev. D* 28:679–93

Bartlett JG, Silk J. 1993. *Ap. J. Lett.* 407:L45–48

Becker R, Fan X, White R, Strauss M, Narayanan V, et al. 2001. *Astron. J.* 122:2850–57

Benson AJ, Nusser A, Sugiyama N, Lacey CG. 2001. *MNRAS* 320:153–76

Bernardeau F. 1997. *Astron. Astrophys.* 324:15–26

Bernardeau F. 1998. *Astron. Astrophys.* 338: 767–76

Bertschinger E. 1998. *Annu. Rev. Astron. Astrophys.* 36:599–654

Birkinshaw M. 1999. *Phys. Rep.* 310:97–195

Birkinshaw M, Gull SF. 1983. *Nature* 302:315–17

Blanchard A, Schneider J. 1987. *Astron. Astrophys.* 184:1–2
Boesgaard AM, Steigman G. 1985. *Annu. Rev. Astron. Astrophys.* 23:319–78
Bond JR, Efstathiou G, Lubin P, Meinhold P. 1991. *Phys. Rev. Lett.* 66:2179–82
Bond JR. 1995. *Phys. Rev. Lett.* 74:4369–72
Bond JR, Efstathiou G. 1984. *Ap. J. Lett.* 285: L45–48
Bond JR, Efstathiou G. 1987. *MNRAS* 226:655–87
Bond JR, Efstathiou G, Tegmark M. 1997. *MNRAS* 291:L33–41
Bond JR, Jaffe AH, Knox L. 1998. *Phys. Rev. D* 57:2117–37
Bond JR, Jaffe AH, Knox L. 2000. *Ap. J.* 533: 19–37
Bouchet F, Gispert R. 1999. *New Astron.* 4:443–79
Boughn SP, Crittenden RG, Turok NG. 1998. *New Astron.* 3:275–91
Bruscoli M, Ferrara A, Fabbri R, Ciardi B. 2000. *MNRAS* 318:1068–72
Bunn EF. 2001. *Phys. Rev. D* 65:043003
Bunn EF, Sugiyama N. 1995. *Ap. J.* 446:49–53
Bunn EF, White M. 1997. *Ap. J.* 480:6–21
Caldwell RR, Dave R, Steinhardt PJ. 1998. *Phys. Rev. Lett.* 80:1582–85
Carlstrom J, Joy M, Grego L, Holder G, Holzapfel W, et al. 2001. *IAP Conf.* In press (astro-ph/0103480)
Christensen N, Meyer R, Knox L, Luey B. 2001. *Class. Quantum Gravity* 18:2677
Coble K, Dodelson S, Frieman JA. 1997. *Phys. Rev. D* 55:1851–59
Cole S, Kaiser N. 1988. *MNRAS* 233:637–48
Cooray A. 2001. *Phys. Rev. D* 64:063514
Cooray A, Hu W. 2000. *Ap. J.* 534:533–50
Cooray A, Hu W, Tegmark M. 2000. *Ap. J.* 540:1–13
Crittenden R, Bond JR, Davis RL, Efstathiou G, Steinhardt PJ. 1993. *Phys. Rev. Lett.* 71:324–27
de Bernardis P, Ade P, Bock J, Bond J, Borrill J, et al. 2002. *Ap. J.* 564:559–66
de Bernardis P, Ade PAR, Bock JJ, Bond JR, Borrill J, et al. 2000. *Nature* 404:955–59
Dodelson S. 1997. *Ap. J.* 482:577–87
Dodelson S, Gates E, Stebbins A. 1996. *Ap. J.* 467:10–18
Dodelson S, Jubas JM. 1993. *Phys. Rev. Lett.* 70:2224–27
Dodelson S, Kinney WH, Kolb EW. 1997. *Phys. Rev. D* 56:3207–15
Dore O, Knox L, Peel A. 2001. *Phys. Rev. D* 64:083001
Doroshkevich AG, Zel'Dovich YB, Sunyaev RA. 1978. *Sov. Astron.* 22:523–28
Efstathiou G, Bond JR. 1987. *MNRAS* 227: 33P–38
Efstathiou G, Moody SJ. 2001. *MNRAS* 325: 1603–15
Eisenstein DJ, Hu W, Tegmark M. 1998. *Ap. J. Lett.* 504:L57–60
Eisenstein DJ, Hu W. 1998. *Ap. J.* 496:605–14
Eisenstein DJ, Hu W, Tegmark M. 1999. *Ap. J.* 518:2–23
Fabbri R, Pollock MD. 1983. *Phys. Lett.* B125: 445–48
Falk T, Rangarajan R, Srednicki M. 1993. *Ap. J. Lett.* 403:L1–3
Ferreira PG, Jaffe AH. 2000. *MNRAS* 312:89–102
Ferreira PG, Magueijo J, Gorski KM. 1998. *Ap. J. Lett.* 503:L1–4
Fixsen DJ, Cheng ES, Gales JM, Mather JC, Shafer RA, et al. 1996. *Ap. J.* 473:576–87
Fosalba P, Dore O, Bouchet FR. 2001. *Phys. Rev. D* 65:063003
Freedman WL, Madore BF, Gibson BK, Ferrarese L, Kelson DD, et al. 2001. *Ap. J.* 553:47–72
Gangui A, Lucchin F, Matarrese S, Molkenach S. 1994. *Ap. J.* 430:447–57
Goldberg DM, Spergel DN. 1999. *Phys. Rev. D* 59:103002
Gruzinov A, Hu W. 1998. *Ap. J.* 508:435–39
Gunn JE, Peterson BA. 1965. *Ap. J.* 142:1633–36
Guth AH, Pi SY. 1985. *Phys. Rev. D* 32:1899–920
Halverson NW, Leitch EM, Pryke C, Kovac J, Carlstrom JE, et al. 2001. *Ap. J.* In press (astro-ph/0104489)
Hanany S, Ade P, Balbi A, Bock J, Borrill J, et al. 2000. *Ap. J. Lett.* 545:L5–9

Hawking SW. 1982. *Phys. Lett.* B115:295–97
Heavens AF. 1998. *MNRAS* 299:805–8
Hedman MM, Barkats D, Gundersen JO, Staggs ST, Winstein B. 2001. *Ap. J. Lett.* 548:L111–14
Hogan CJ, Kaiser N, Rees MJ. 1982. *Philos. Trans. R. Soc. London Ser.* 307:97–109
Holder G, Carlstrom J. 2001. *Ap. J.* 558:515–59
Hu W. 1998. *Ap. J.* 506:485–94
Hu W. 2000a. *Ap. J.* 529:12–25
Hu W. 2000b. *Phys. Rev. D* 62:043007
Hu W. 2001a. *Phys. Rev. D* 64:083005
Hu W. 2001b. *Ap. J. Lett.* 557:L79–83
Hu W. 2002. *Phys. Rev. D* 65:023003
Hu W, Fukugita M, Zaldarriaga M, Tegmark M. 2001. *Ap. J.* 549:669–80
Hu W, Seljak U, White M, Zaldarriaga M. 1998. *Phys. Rev. D* 57:3290–301
Hu W, Sugiyama N. 1995. *Ap. J.* 444:489–506
Hu W, Sugiyama N. 1996. *Ap. J.* 471:542–70
Hu W, Sugiyama N, Silk J. 1997. *Nature* 386: 37–43
Hu W, White M. 1996. *Ap. J.* 471:30–51
Hu W, White M. 1997a. *New Astron.* 2:323–44
Hu W, White M. 1997b. *Phys. Rev. D* 56:596–615
Hu W, White M. 1997c. *Ap. J.* 479:568–79
Hu W, White M. 2001. *Ap. J.* 554:67–73
Hui L, Burles S, Seljak U, Rutledge RE, Magnier E, et al. 2001. *Ap. J.* 552:15–35
Huterer D, Knox L, Nichol RC. 2001. *Ap. J.* 555:547–57
Jungman G, Kamionkowski M, Kosowsky A, Spergel DN. 1996. *Phys. Rev. D* 54:1332–44
Kaiser N. 1983. *MNRAS* 202:1169–80
Kaiser N. 1984. *Ap. J.* 282:374–81
Kamionkowski M, Kosowsky A, Stebbins A. 1997. *Phys. Rev. D* 55:7368–88
Kamionkowski M, Spergel DN, Sugiyama N. 1994. *Ap. J. Lett.* 426:L57–60
Keating B, O'Dell C, de Oliveira-Costa A, Klawikowski S, Stebor N, et al. 2001. *Ap. J. Lett.* 560:L1–4
Knox L. 1995. *Phys. Rev. D* 52:4307–18
Knox L. 1999. *MNRAS* 307:977–83
Knox L, Christensen N, Skordis C. 2001. *Ap. J. Lett.* 563:L95–98
Knox L, Scoccimarro R, Dodelson S. 1998. *Phys. Rev. Lett.* 81:2004–7
Kofman LA, Starobinskii AA. 1985. *Sov. Astron. Lett.* 11:271–74
Kogut A, Banday AJ, Bennett CL, Gorski KM, Hinshaw G, et al. 1996. *Ap. J. Lett.* 464:L29–33
Komatsu E, Kitayama T. 1999. *Ap. J. Lett.* 526: L1–4
Komatsu E, Seljak U. 2001. *MNRAS* 327:1353–66
Krauss LM, Turner MS. 1995. *Gen. Relativ. Gravity* 27:1137–44
Lee A, Ade P, Balbi A, Bock J, Borill J, et al. 2001. *Ap. J. Lett.* 561:L1–5
Lewis A, Challinor A, Turok N. 2001. *Phys. Rev. D* 65:023505
Lewis GF, Babul A, Katz N, Quinn T, Hernquist L, et al. 2000. *Ap. J.* 536:623–44
Liddle AR, Lyth DH. 1993. *Phys. Rep.* 231:1–105
Limber DN. 1954. *Ap. J.* 119:655–81
Loeb A, Barkana R. 2001. *Annu. Rev. Astron. Astrophys.* 39:19–66
Luo X. 1994. *Ap. J. Lett.* 427:L71–74
Ma C, Bertschinger E. 1995. *Ap. J.* 455:7–25
Miller AD, Caldwell R, Devlin MJ, Dorwart WB, Herbig T, et al. 1999. *Ap. J. Lett.* 524: L1–4
Narkilar J, Padmanabhan T. 2001. *Annu. Rev. Astron. Astrophys.* 39:211–48
Netterfield C, Ade P, Bock J, Bond J, Borrill J, et al. 2001. *Ap. J.* In press (astro-ph/0104460)
Oh SP, Spergel DN, Hinshaw G. 1999. *Ap. J.* 510:551–63
Ostriker JP, Steinhardt PJ. 1995. *Nature* 377: 600–2
Ostriker JP, Vishniac ET. 1986. *Ap. J. Lett.* 306: L51–54
Padin S, Cartwright JK, Mason BS, Pearson TJ, Readhead ACS, et al. 2001. *Ap. J. Lett.* 549:L1–5
Parodi BR, Saha A, Sandage A, Tamman GA. 2000. *Ap. J.* 540:634–51
Peacock JA. 1991. *MNRAS* 253:1P–5P
Pearce FR, Jenkins A, Frenk CS, White SDM, Thomas PA, et al. 2001. *MNRAS* 326:649–66
Peebles PJE. 1968. *Ap. J.* 153:1–11

Peebles PJE. 1982. *Ap. J. Lett.* 263:L1–5
Peebles PJE, Yu JT. 1970. *Ap. J.* 162:815–36
Penzias AA, Wilson RW. 1965. *Ap. J.* 142:419–21
Percival W, Baigj CM, Bland-Hawthorn J, Bridges T, Cannon R, et al. 2001. *MNRAS* 327:1297–1306
Perlmutter S, Aldering G, Goldnater G, Knop RA, Nugent P, et al. 1999. *Ap. J.* 517:565–86
Persi FM, Spergel DN, Cen R, Ostriker JP. 1995. *Ap. J.* 442:1–9
Phillips NG, Kogut A. 2001. *Ap. J.* 548:540–49
Polnarev AG. 1985. *Sov. Astron.* 29:607
Prunet S, Sethi SK, Bouchet FR. 2000. *MNRAS* 314:348–53
Pryke C, Halverson N, Leitch E, Kovac J, Carlstrom J, et al. 2001. *Ap. J.* In press (astro-ph/0104490)
Readhead A, Lawrence C, Myers S, Sargent W, Hardebeck H, et al. 1989. *Ap. J.* 346:566–87
Rees M, Sciama D. 1968. *Nature* 217:511
Refregier A, Komatsu E, Spergel DN, Pen UL. 2000. *Phys. Rev. D* 61:123001
Riess A, Fillipenko AV, Challis P, Clocchiatti A, Diercks A, et al. 1998. *Ap. J.* 116:1009
Rubakov VA, Sazhin MV, Veryaskin AV. 1982. *Phys. Lett. B* 115:189–92
Sachs RK, Wolfe AM. 1967. *Ap. J.* 147:73–90
Sandage A, Tamman GA, Saha A. 1998. *Phys. Rep.* 307:1–14
Schramm DN, Turner MS. 1998. *Rev. Mod. Phys.* 70:303–18
Scott D, Srednicki M, White M. 1994. *Ap. J. Lett.* 421:L5–7
Seager S, Sasselov DD, Scott D. 2000. *Ap. J. Suppl.* 128:407–30
Seljak U. 1994. *Ap. J. Lett.* 435:L87–90
Seljak U. 1996a. *Ap. J.* 463:1–7
Seljak U. 1996b. *Ap. J.* 460:549–55
Seljak U. 1997. *Ap. J.* 482:6–16
Seljak U, Burwell J, Pen UL. 2001. *Phys. Rev. D* 63:063001
Seljak U, Pen U, Turok N. 1997. *Phys. Rev. Lett.* 79:1615–18
Seljak U, Zaldarriaga M. 1996. *Ap. J.* 469:437–44
Seljak U, Zaldarriaga M. 1999. *Phys. Rev. D* 60:043504
Seljak U, Zaldarriaga M. 2000. *Ap. J.* 538:57–64
Silk J. 1968. *Ap. J.* 151:459–71
Smoot GF, Bennett CL, Kogut A, Wright EL, Aymon J, et al. 1992. *Ap. J. Lett.* 396:L1–5
Smoot GF, Gorenstein MV, Muller RA. 1977. *Phys. Rev. Lett.* 39:898–901
Souradeep T, Ratra BV. 2001. *Ap. J.* 560:28–40
Spergel D, Goldberg D. 1999. *Phys. Rev. D* 59: 103001
Springel V, White M, Hernquist L. 2001. *Ap. J.* 549:681–87
Starobinskii A. 1985. *Sov. Astron. Lett.* 11:133–36
Stompor R, Balbi A, Borrill J, Ferreira P, Hanany S, et al. 2001. *Ap. J.* In press (astro-ph/0106451)
Strauss MA, Willick JA. 1995. *Phys. Rep.* 261: 271–431
Sunyaev R, Zel'dovich Y. 1972. *Comments Astrophys. Space Phys.* 4:173
Szapudi I, Prunet S, Pogosyan D, Szalay AS, Bond JR. 2001. *Ap. J. Lett.* 548:L115–18
Tegmark M. 1997a. *Ap. J. Lett.* 480:L87–90
Tegmark M. 1997b. *Phys. Rev. D* 55:5895–907
Tegmark M. 1998. *Ap. J.* 502:1–6
Tegmark M, Bunn EF. 1995. *Ap. J.* 455:1–6
Tegmark M, de Oliveira-Costa A. 2001. *Phys. Rev. D* 64:063001
Tegmark M, Efstathiou G. 1996. *MNRAS* 281: 1297–314
Tegmark M, Eisenstein DJ, Hu W, de Oliveira-Costa A. 2000. *Ap. J.* 530:133–65
Tegmark M, Hamilton AJS, Strauss MA, Vogeley MS, Szalay AS. 1998. *Ap. J.* 499:555–76
Tegmark M, Taylor AN, Heavens AF. 1997. *Ap. J.* 480:22–35
Tegmark M, Zaldarriaga M. 2000. *Ap. J.* 544: 30–42
Tegmark M, Zaldarriaga M, Hamilton AJS. 2001. *Phys. Rev.* D63:043007
Theureau G, Hanski M, Ekholm T, Bottinelli L, Goriguenheim U, et al. 1997. *Astron. Astrophys.* 322:730
Tuluie R, Laguna P. 1995. *Ap. J. Let.* 445:L73–76
Turok N. 1996. *Phys. Rev. Lett.* 77:4138–41

Tytler D, O'Meara JM, Suzuki N, Lubin D. 2000. *Physica Scripta Vol. T* 85:12–31
Vishniac ET. 1987. *Ap. J.* 322:597–604
Vittorio N, Silk J. 1984. *Ap. J. Lett.* 285:L39–43
Wandelt B, Hansen F. 2001. *Phys. Rev. D* In press (astro-ph/0106515)
Wandelt BD, Gorski KM. 2001. *Phys. Rev. D* 63:123002
Wang X, Tegmark M, Zaldarriaga M. 2001. *Phys. Rev. D* In press (astro-ph/0105091)
Weinberg S. 1971. *Ap. J.* 168:175–94
White M. 1998. *Phys. Rev. D* 57:5273–75
White M, Hu W. 1997. *Astron. Astrophys.* 321: 8–9
White M, Scott D, Silk J. 1994. *Annu. Rev. Astron. Astrophys.* 32:319–70
White SDM, Efstathiou G, Frenk CS. 1993. *MNRAS* 262:1023–28
Wilson ML, Silk J. 1981. *Ap. J.* 243:14–25
Winitzki S, Kosowsky A. 1997. *New Astron.* 3:75–100
Wright EL, Hinshaw G, Bennett CL. 1996. *Ap. J. Lett.* 458:L53–55
Wu JHP, Balbi A, Borrill J, Ferreira PG, Hanany S, et al. 2001. *Ap. J. Suppl.* 132:1–17
Zaldarriaga M. 2000. *Phys. Rev. D* 62:063510
Zaldarriaga M, Seljak U. 1997. *Phys. Rev. D* 55:1830–40
Zaldarriaga M, Seljak U. 1998. *Phys. Rev. D* 58:023003
Zaldarriaga M, Seljak U. 1999. *Phys. Rev. D* 59:123507
Zaldarriaga M, Spergel DN, Seljak U. 1997. *Ap. J.* 488:1–13
Zel'dovich Y, Kurt V, Sunyaev R. 1969. *Sov. Phys.–JETP* 28:146

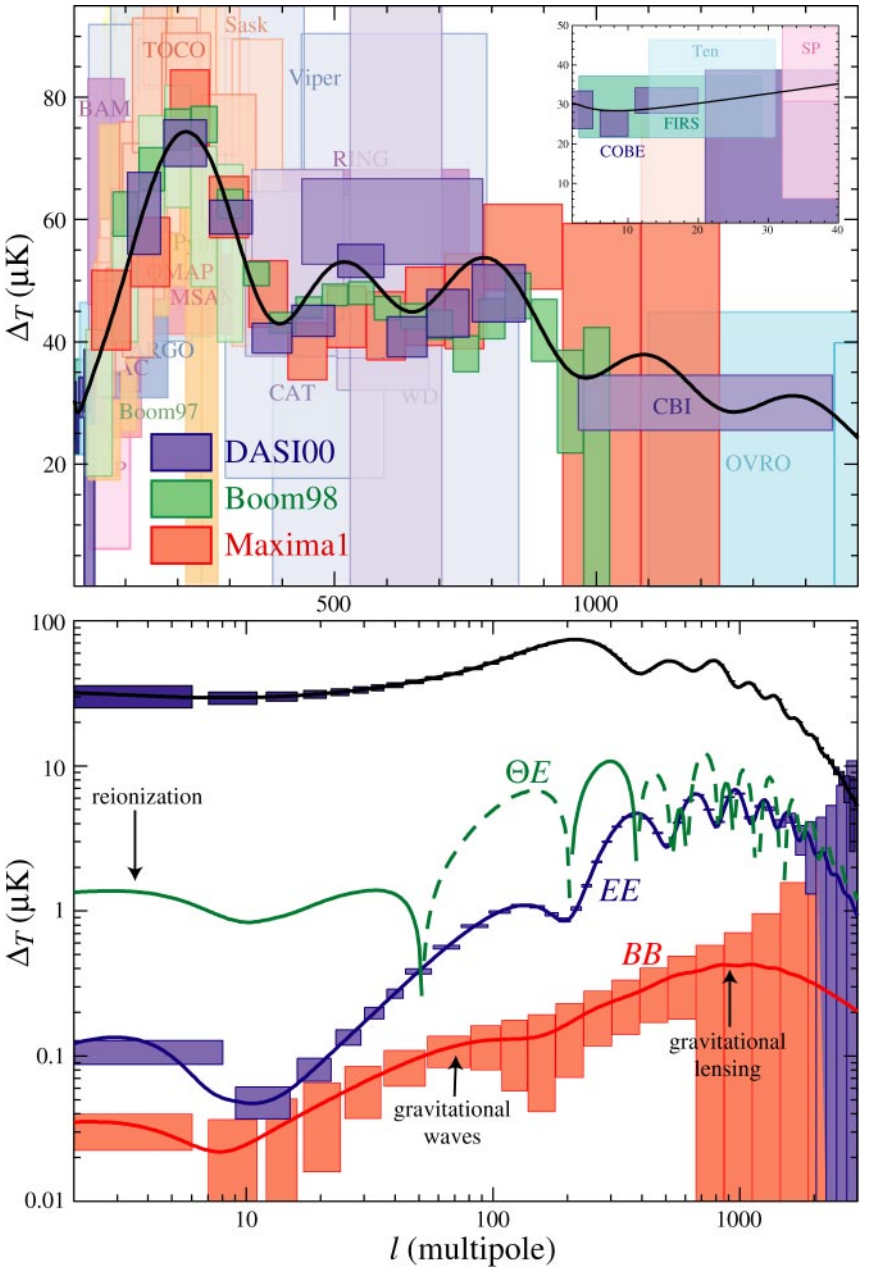

Figure 1 (*Top*) Temperature anisotropy data with boxes representing 1-σ errors and approximate l-bandwidth. (*Bottom*) Temperature and polarization spectra for $\Omega_{\rm tot} = 1$, $\Omega_\Lambda = 2/3$, $\Omega_b h^2 = 0.02$, $\Omega_m h^2 = 0.16$, $n = 1$, $z_{\rm ri} = 7$, $E_i = 2.2 \times 10^{16}$ GeV. *Dashed lines* represent negative cross correlation and *boxes* represent the statistical errors of the Planck satellite.

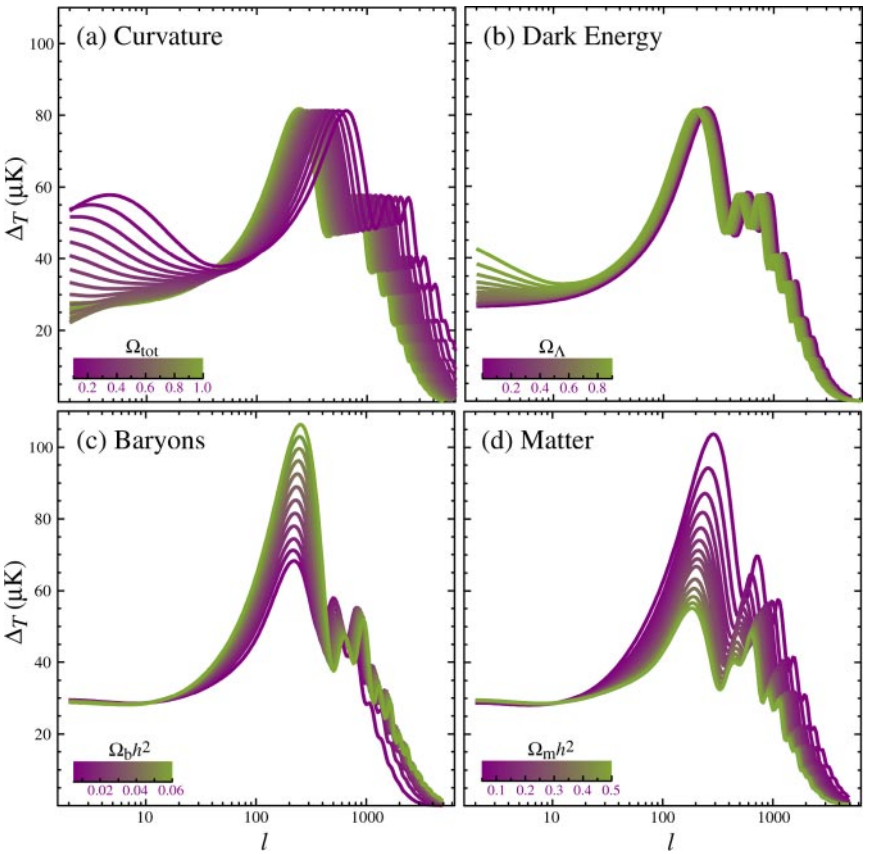

Figure 4 Sensitivity of the acoustic temperature spectrum to four fundamental cosmological parameters. (*a*) The curvature as quantified by $\Omega_{\rm tot}$. (*b*) The dark energy as quantified by the cosmological constant Ω_Λ ($w_\Lambda = -1$). (*c*) The physical baryon density $\Omega_b h^2$. (*d*) The physical matter density $\Omega_m h^2$. All are varied around a fiducial model of $\Omega_{\rm tot} = 1$, $\Omega_\Lambda = 0.65$, $\Omega_b h^2 = 0.02$, $\Omega_m h^2 = 0.147$, $n = 1$, $z_{\rm ri} = 0$, $E_i = 0$.

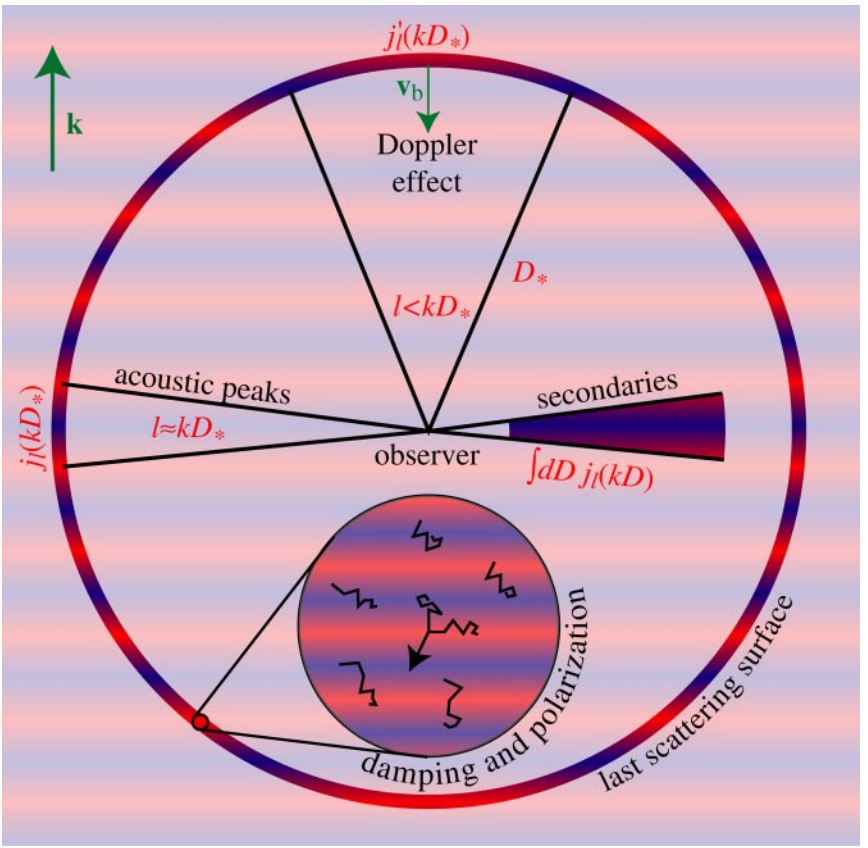

Figure 6 Integral approach. Cosmic Microwave Background anisotropies can be thought of as the line-of-sight projection of various sources of plane wave temperature and polarization fluctuations: the acoustic effective temperature and velocity or Doppler effect (see §3.8), the quadrupole sources of polarization (see §3.7) and secondary sources (see §4.2, §4.3). Secondary contributions differ in that the region over which they contribute is thick compared with the last scattering surface at recombination and the typical wavelength of a perturbation.

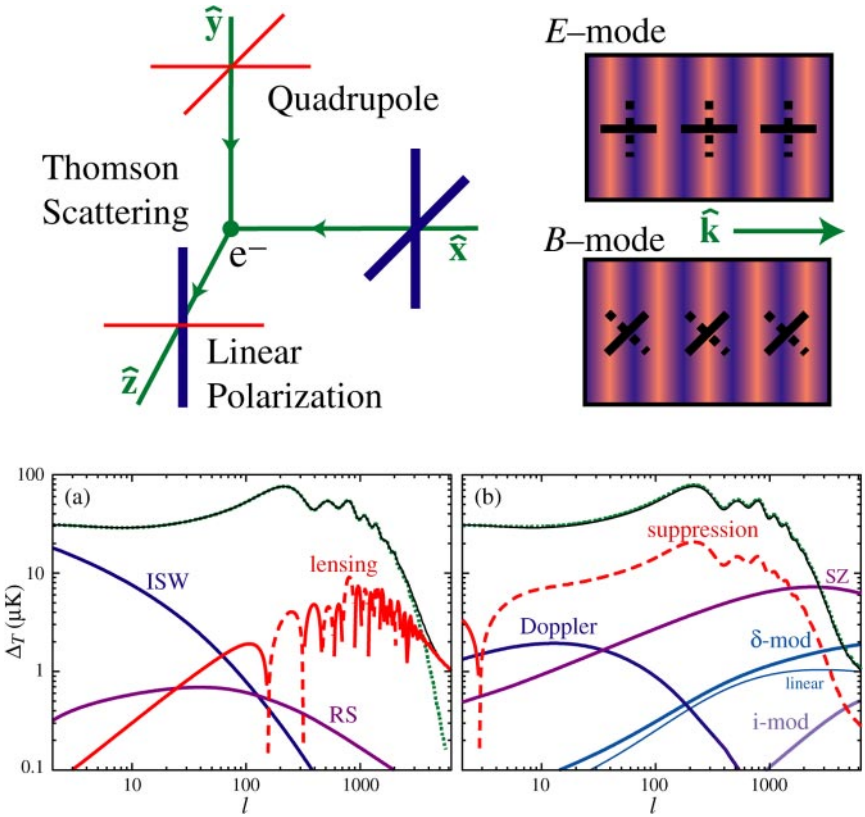

Figure 7 (*Top*) Polarization generation and classification. Thomson scattering of quadrupole temperature anisotropies generates linear polarization. The component of the polarization that is parallel or perpendicular to the wavevector **k** is called the *E*-mode, and the one at 45° angles is called the *B*-mode. (*Bottom*) Secondary anisotropies. (*a*) Gravitational secondaries: Integrated Sachs-Wolfe, lensing, and Rees-Sciama (moving halo) effects. (*b*) Scattering secondaries: Doppler, density (δ), and ionization (i) modulated Doppler, and the Sunyaev-Zel'dovich effects. Curves and model are described in the text.