The cosmic microwave background: past, present and future

2007 J. Phys. A: Math. Theor. 40 6801

(http://iopscience.iop.org/1751-8121/40/25/S24)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **40** (2007) 6801–6809 doi:10.1088/1751-8113/40/25/S24

The cosmic microwave background: past, present and future

Joseph Silk

Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road,
Oxford OX1 3RH, UK

E-mail: silk@astro.ox.ac.uk

Received 19 December 2006, in final form 30 January 2007
Published 6 June 2007
Online at stacks.iop.org/JPhysA/40/6801

Abstract
The cosmic microwave background has provided an unprecedented cosmological window on the very early universe for probing the initial conditions from which structure evolved. Infinitesimal variations in temperature on the sky, first predicted in 1967 but only discovered in the 1990s, provide the fossil fluctuations that seeded the formation of the galaxies. The cosmic microwave background radiation has now been mapped with ground-based, balloon-borne and satellite telescopes. I describe its current status and future challenges.

PACS numbers: 98.80.Bp, 98.80.Es

1. Introduction

Alpher, Herman and Gamow made one of the most remarkable predictions in modern physics, that the universe once was hot. At the time (in the 1940s), this seemed to be the most important channel for synthesizing the chemical elements. Of course, we now know that their logic was slightly flawed, as they did not realize that only the light elements could be, and indeed needed to be, synthesized in the first few minutes of the big bang. Moreover, their implementation lacked an important feature, since they failed to make any prediction about the observability today of the cosmic black body radiation as a relic cosmic microwave background. Historical credit for this connection goes to Doroshkevich and Novikov [1], who however misinterpreted data on the sky brightness taken at Bell Laboratories to rule out a hot big bang.

Within a year, however, and independently of any theoretical predictions, Penzias and Wilson utilized the same Bell Laboratories radio telescope to serendipitously discover the cosmic microwave background. It was to take another 16 years before its black body spectrum was measured by the FIRAS spectrometer on the COBE satellite, a measurement for which John Mather received the Nobel Prize in 2006.

1751-8113/07/256801+09$30.00

Once the background was detected, the next step was to measure the predicted temperature fluctuations. The COBE satellite again made the breakthrough, with the DMR differential radiometer experiment headed by George Smoot, similarly given Nobel recognition in 2006. Smoot pioneered the discovery of large angular scale fluctuations, originally predicted by Sachs and Wolfe in 1967 [2] at a resolution in excess of 5°. Since the causal horizon of the last scattering surfaces subtends an angle of about 1°, this means that in the context of inflationary cosmology, whereby quantum fluctuations are boosted to macroscopic scales, we are essentially viewing quantum fluctuations in the sky. The original prediction was based on the observed large-scale clumpiness of the universe, but was overly optimistic by some four orders of magnitude. The inflationary theory of cosmology (in 1980) did not make any prediction about the fluctuation strength or, more precisely, in its original version it predicted far larger fluctuations than even those imagined by Sachs and Wolfe.

The final step came with the discovery and mapping of subdegree scale fluctuations in the CMB. These too were predicted in 1967 [3, 4] as a prerequisite for galaxy and cluster formation by gravitational instability in the expanding universe. In this case, however, the prediction was overly optimistic by only one order of magnitude. The major improvement was the inclusion of weakly interacting dark matter in 1984. Inflation theory did anticipate, as did an earlier argument due to Harrison and Zel'dovich, that the fluctuations should be scale invariant. This meant that the theorists were able to predict, to within a factor of 2, the fluctuations measured by the COBE DMR [5, 6]. Direct confirmation of the seed fluctuations required for structure formation took another decade or so, culminating in the WMAP all-sky images of the CMB sky at 15 arc min resolution (in 2003).

2. Some theory

Primordial sound waves in the photon–baryon plasma evolve into density fluctuations in the matter-dominated era. We observe that non-baryonic weakly interacting dark matter dominates over baryonic matter by a factor of approximately 6. This dominance allows fluctuation growth to be boosted by roughly a similar factor. The approximate prediction for fluctuations that crest and compress to maximum density on the horizon scale at last scattering is $\delta T/T \sim 3 \times 10^{-5}$ on a scale of a degree. This is the first and dominant peak in the radiation fluctuation power spectrum.

Parameter degeneracies compound any simple interpretation in terms of cosmological parameters. For example, lowering the Hubble constant pushes the last scattering surface further away and mimics the effect on the angular scale of the first peak of making the curvature of the universe more negative. Because of this and other parameter degeneracies, we need to measure more than the CMB alone. Results from the large-scale distribution of galaxies have greatly refined the accuracy of the inferred cosmological parameters.

3. Applications

Cosmic microwave background and large-scale structure constraints overlap. For the CMB, one inverts the Boltzmann equation. The universe is completely linear at last scattering, $z \sim 1000$, and so the calculations are straightforward. There are currently two large-scale galaxy redshift surveys, 2DFGRS and SDSS. Both explore the universe over scales from ∼10 to 200 Mpc and to a depth of $z \sim 0.2$. Corrections must be made for nonlinearity, which is a large effect on scales below ∼10 Mpc, and conversion must be implemented from redshift space to real space. Remarkable agreement is found between fluctuations in the large-scale galaxy distribution and the density fluctuation spectrum derived from the CMB in the region of overlap for a reasonable value of the bias. It is this last factor, the ratio of the mean square fluctuations in the galaxy counts to those in the dark matter on a fiducial scale of $8h^{-1}$ Mpc, that sets the relative normalization of the two independently derived fluctuation spectra. Bias depends on our theory of galaxy formation, and hence is an empirically derived quantity. We define σ_8 to be the inverse of the bias factor, and its current value is determined to be 0.74 ± 0.06 from the WMAP3 data in combination with other large-scale structure data.

3.1. Dark matter

Detection and identification of dark matter is one of the greatest challenges in astrophysics. It has played a significant role in setting the level of the predicted fluctuation level in the CMB due to the period of fluctuation growth in the matter-dominated era prior to last scattering. The amplitude of the acoustic peaks is sensitive to the ratio of dark matter to radiation density, and hence provides a measure of the dark matter density. The primordial acoustic waves also imprint baryon oscillations that are observed in the large-scale galaxy distribution. These yield an independent measure of the dark matter density. Yet a third determination come from weak gravitational lensing. The currently preferred value is $\Omega_m h^2 = 0.127 \pm 0.008$ with $h = 0.73 \pm 0.04$ [7].

3.2. Dark energy

The CMB yields a flat universe from the acoustic peak locations. The only way to reconcile zero curvature or $\Omega = 1$ with the measured matter density is to introduce a cosmological constant to take up the slack. Reinterpretation of Einstein's cosmological constant yields dark energy, generalized by theorists to be a time-dependent vacuum density. It is conveniently parametrized by the equation of state parameter $w(z) \equiv p/\rho$, with three possible prescriptions in which w is either a constant or time dependent: $w = -1$, $w < -1/3$ or $w(z)$.

Dark energy induces an acceleration of the expansion rate at recent epochs, when the universe becomes dark-energy dominated. This effect has been directly measured by use of distant supernovae as standard candles. The combination of WMAP3 and supernova data at $z > 1$ favours the simplest model, $w = -0.97 \pm 0.08$, corresponding to Einstein's cosmological constant. The inferred value of dark energy is $\Omega_\Lambda = 0.72 \pm 0.04$ [8]. There is no evidence for any deviation from a constant value of Λ to a redshift of unity [9].

4. Surprises

We make progress in cosmology from experimental results that do not fall into complete agreement with our prior prejudices. There are some curious anomalies or hints in the data that suggest we may not yet have reached the final model of cosmology. Theoretical explanations are guided by observations.

4.1. Axis of evil

All-sky maps of the cosmic microwave background reveal persistent anomalies. The quadrupole is low, the lowest multipoles are aligned, there is a north–south asymmetry, and there is at least one significant cold spot [10]. All of these anomalies are present with something like a confidence level of 2σ.

4.2. *Topology*

There is little guidance from theory about the topology of the universe. A non-trivial topology could help account for some of the large angular scale anomalies. More generally, non-trivial topology generates global anisotropies that could provide a unique signature. One difficulty is that in an open universe there are an infinite number of possible topologies. However if space is flat or compact, as is consistent with, if not required by, the CMB data, the choice is much more restrictive.

In fact, there are 18 locally homogeneous and isotropic Euclidean spaces in three spatial dimensions. One of these is the analogue of the infinite plane. Of the non-trivial topologies, 17 are multiply connected and 10 of these are compact. Only 6 of the compact topologies are orientable. If the universe is indeed compact, the choice is even simpler, reducing to one of the spherical multiply connected manifolds with finite fundamental domains.

There are (weak) arguments that the universe should be compact, from considerations of the wavefunction of the universe, and physics may reasonably require orientability along geodesics. Power is plausibly suppressed on scales larger than the fundamental domain, and one has a natural explanation for the low quadrupole. Orientability also leads to a breakdown of global isotropy, thereby allowing the possibility of alignments of the lowest moments [12].

4.3. *A compact universe?*

Attempts to fit the CMB, large-scale structure and supernova data simultaneously lead to asymmetric confidence level contours in the two-dimensional parameter space defined by Ω_m and Ω_Λ. While a flat universe is fully consistent with the data, most of the allowed area lies in the closed universe domain, with the deviation from the critical density being small: $\Omega_\kappa = -0.01 \pm 0.01$. A Bayesian analysis would even favour a slightly open universe. However, if, for example, the global Hubble constant were unexpectedly low, due to effects of large-scale inhomogeneity on the local measurements, Ω_m increases and the data would then favour a compact universe.

There is of course a geometric measure of the curvature of the universe. The angular scale of the CMB acoustic peaks favours a flat universe only if the Hubble constant is taken to be close to the locally measured value of 72 or 73 km (s Mpc)$^{-1}$. However if the global Hubble constant were significantly lower, and this is not inconsistent with the geometric measurements, the last scattering surface would be further away and its angular size decrease in a flat universe would simulate the effect of an open universe. Correcting for this again means that the universe would preferentially be closed.

It is essential to consider alternatives to the standard model in confronting the data, in order to test the robustness of our conclusions. Allowance for generic initial conditions of the primordial density fluctuations means that isocurvature modes are as equally admissible as adiabatic modes. Fitting to the acoustic peaks sets a limit of about 30% on the maximum contribution of such modes.

4.4. *Non-Gaussianity*

Another manifestation of generic initial conditions is possible primordial non-Gaussianity. This may be probed via at least two approaches. One involves measuring higher order moments of the microwave background. Another involves the large-scale galaxy distribution. One can look at higher moments in either 2D or 3D, but there is also a three-dimensional approach that is potentially more powerful. This involves counting massive and rare objects such as galaxies and galaxy clusters and studying their number density as a function of cosmological epoch [11]. For example, the number of massive galaxies should decrease rapidly with redshift, beyond a certain lookback time. Indeed, the number above a certain mass becomes exponentially small for a Gaussian distribution.

One can apply similar reasoning to quasilinear features in the galaxy distribution on very large scales. Comparison of simulations with Gaussian theory has been remarkably successful in demonstrating the existence of a cosmic web that is characteristic of the dominance of cold dark matter. However in surveys such as SDSS and 2DFGRS that are complete to 200 Mpc or slightly beyond, observed features such as the so-called Great Wall may be compared with Gaussian simulations. The distribution of overdensities seems to be more prominent than is found in the simulations [13]. The possible mismatch is aggravated by the revised normalization found for the matter power spectrum by the WMAP3 results, but compounded by issues of bias, that is, how galaxies are selected from otherwise reliable dark matter simulations of large-scale structure.

4.5. The dark ages

We are now beginning to probe the dark ages of the universe. These represent the epoch between matter-radiation decoupling and the formation of the first stars and galaxies. The CMB has hitherto proved to be the most powerful tool for these studies. The epoch of reionization has been measured via the polarization of the CMB, which allows a determination of the optical depth of the universe between now and decoupling. The most recent value of the optical depth to electron scattering, about 0.09 from WMAP3, requires that the universe was reionized by about $z \sim 9$. Because of the reduction in normalization found in the WMAP3 data ($\sigma_8 \approx 0.74$ instead of 0.94) [8], there is little change in the requirements on the nature of the ionizing sources, despite their lower redshift, which originally were required to be beyond $z \sim 17$ by the first year WMAP data. The reduction in optical depth was primarily due to improved polarization mapping of the CMB sky.

The reionization of the universe requires a population of early, ionizing sources that exceeds most estimates of the contribution from normal galaxies. One needs a population of low mass dwarf galaxies or of accreting intermediate black holes. Either could provide the necessary early boost in the number of ionizing photons per baryon expected from the known galaxy population that is needed to overcome the rapid recombination rate of hydrogen at such an early epoch.

5. Challenges for the future

Inflationary cosmology represents the principal major addition to cosmology since the discovery of the eponymous models of Friedmann, Einstein, de Sitter, Lemaitre and Eddington in the 1920s and 1930s. Inflation makes important predictions, in particular about the flatness of space and the power spectrum of the primordial fluctuations, that have been largely verified by the CMB temperature fluctuations, complemented by large-scale structure studies. Nevertheless, there is so much freedom in the theory, indeed there is no unique theory but hundreds of alternative theories of inflation, that cosmologists would dearly like to find a generic prediction to verify that inflation actually occurred.

5.1. The gravity wave background

The closest we have to a generic property of inflation is the presence of a gravity wave background. Gravity waves induce divergence-free temperature fluctuations and can be represented as tensor metric perturbations. Compressible (adiabatic or isocurvature) temperature fluctuations are scalar metric perturbations. The search for the gravity wave component has been called the quest for the holy grail of inflation and is performed by detecting the vector potential (curl-like or B mode) polarization. Hitherto, only the scalar potential (divergence-like or E mode) polarization signal has been detected. The B mode signal is at least an order of magnitude weaker.

One difficulty is that in typical models the amplitude of the gravity wave temperature component is approximately proportional to the dimensionless ratio $(m_{\rm GUT}/m_{\rm Planck})^2$, and $T_{\rm GUT}$ could be very small at low reheating temperatures. In single field models of inflation, however, there is the expectation that the tensor-to-scalar ratio is given by $T/S = 7(1-n)$, where n is the spectral index of a power-law fit to the power spectrum of the primordial fluctuations. The WMAP3 results prefer a value $n \approx 0.97$, and this would therefore give some reason for optimism were inflation simple. However, the systematic uncertainties are sufficiently large at present that one cannot really distinguish between inflationary models.

Several experiments are under development or being planned to search for the elusive polarization signal as well as to refine the power spectrum spectral index measurements. Foregrounds provide a potentially severe problem, and the ultimate experiment will almost certainly be done in space.

5.2. *Dark energy confronted*

An important goal for fundamental physics is to improve our knowledge of dark energy. The CMB provides an important constraint on the sum of dark matter and dark energy. It provides one of only two techniques for direct measurement of acceleration via the integrated Sachs–Wolfe (ISW) effect. The alternative, using supernovae, has hitherto been more successful, but is likely to run into systematics that will only be overcome with space-based surveys.

Over the next decade, ground-based projects should nevertheless make substantial inroads into clarifying the nature of dark energy. The Dark Energy Survey (2009+), led by FermiLab, will use 4–5 bands in the visible spectrum via a new camera installed on the refurbished 4m Blanco Telescope at Cerro Tololo to map 10 000 square degrees and obtain photometric redshifts to $z > 1$. A similar survey, PANSTARS, led by the University of Hawaii, will use two (and eventually four) dedicated 1.8 m telescopes (2006–2009+) at Mauna Loa, in a project partly funded by the US Air Force to map 3000 square degrees in five optical bands.

Ultimate precision requires spectroscopy. This is being obtained in complementary surveys over smaller areas. Initially (2006/9), the AAOmega spectrograph will map 1000 square degrees with the AAT 4m telescope. This will be followed with the GWFMOS/HyperSuprime spectrograph on the Subaru 8m telescope at Mauna Kea, which will explore $z \sim 1-3$ (beginning in 2012). The ultimate ground-based surveys will use the 8m LSST, now in the design phase, to map 20 000 square degrees to $z \sim 3$ (2014), obtaining photometric redshifts to per cent accuracy with the aid of supplementary NIR data from surveys such as those planned for the ESO/UK VISTA telescope. Eventually, the Square Kilometre Array will obtain spectroscopic redshifts at a rest wavelength of 21 cm, appropriately redshifted, for every galaxy containing HI in the observable universe. In practice, this will mean a million redshifts over 20 000 square degrees (to be completed by 2020). The ultimate goal of all of these experiments is to determine w, the ratio of pressure to density in the dark energy component, to a precision of 2 or 3%, including its possible dependence on epoch. The key is measurement of the baryon acoustic oscillations, the analogue of the CMB acoustic temperature fluctuations. The power of baryon acoustic oscillations is that one can split the universe up into slices at different depths. Galaxy survey tomography thereby allows one to study the evolution of the density fluctuations and this is sensitive to dark energy.

Another probe of new physics may be buried in the low ℓ structure of the cosmic microwave background. As mentioned above, there appear to be several anomalies, all more or less at the 2 sigma level. These include the north–south asymmetry of the power in the temperature fluctuations and a large cold spot in the cosmic microwave background. The distribution of galaxies on very large scales also reveals several very large structures, such as one dubbed the Great Wall. Similar features appear in simulations, but are generally less prominent especially if the lower normalization of WMAP3 is adopted. It is possible that the two phenomena are connected.

Bias is an issue: if it is larger than anticipated from measurements on smaller scales, this could account for some or all of the anomalous large-scale structure signal. One could also be seeing such exotica as trans-Planckian relics, signatures of a non-trivial topology in a compact universe, or even rare massive structures that are remnants of primordial non-Gaussian signatures. Any or all of these possibilities await confirmation with future data sets that would also include, for example, all-sky polarization maps.

5.3. Probing the dark ages

There is another major CMB frontier to be breached at small angular scales. By $\ell \sim 2000$, primary anisotropies are falling steeply. Any signal detected at higher ℓ will necessarily of secondary origin. The most important of these is the cumulative Sunyaev–Zel'dovich effect due to intervening galaxy clusters along the line of sight to the last scattering surface. Multifrequency observations can distinguish this signal, at least in principle. Current observations, most notably by the CBI and ACBAR experiments, reveal a possible excess at small angular scales with the expected spectral signature. However, the interpretation as an integrated cluster signal is very sensitive to the adopted normalization of the fluctuation power spectrum, varying as σ_8^7. So far, one can argue that the data require either a value $\sigma_8 \sim 1$, a normalization that is in conflict with the most recent observations, or recourse to models in which the numbers of high redshift massive clusters are boosted. Such a boost may occur for non-Gaussian initial conditions or alternatively in late dark energy models wherein the effective threshold for nonlinear collapse is lowered below the value of 1.686 traditionally used as a linear theory filter on the density fluctuation amplitude [14].

The dark ages, before the first galaxies formed, may also contribute to the high ℓ signal. The polarization map from the WMAP experiment determines the optical depth of the universe and demonstrates that reionization most likely occurred at $z \sim 8$, before luminous quasars and galaxies were in place. The source of the ionizing photons is probably due to a large number of early-forming dwarf galaxies that contained massive stars. Determination of the galaxy luminosity function in the Hubble ultradeep field demonstrates that massive galaxies are sparse but dwarfs are common: the galaxy luminosity function continues to rise to a redshift of at least 5 and possibly 8 [15]. Such dwarfs could provide the necessary source of ionizing photons.

Another more exotic possibility is a population of intermediate mass black holes, whose presence is inferred indirectly from theoretical reasoning. One argument is that the first clouds in which baryonic dissipation occurred included a population of clouds of atomic hydrogen that cooled via Lyman alpha emission. Such clouds are prime candidates for forming IMBHs. Another argument is that the supermassive black holes that power quasars and are ubiquitously found in the cores of spheroids require a precursor population of seed IMBHs in order to have formed efficiently in the early universe. The resulting miniquasars provide an alternative source of ionizing photons that could have reionized the universe. Experiments such as LOFAR will be able to search for highly redshifted 21 cm emission whose statistical properties could help distinguish a non-thermal from a thermal ionizing photon source.

6. The ultimate goals of cosmology

We are now at a stage in cosmology where we have converged on a standard model, that of ΛCDM. The model parameters are measured to within a few per cent. There are several challenges that remain for the theorists. The origin of dark energy and its recent emergence as a dominant energy density are not understood. The fact that the universe appears to have a Euclidean geometry is another coincidence that requires the dark energy and matter densities to closely match each other.

Key questions to be resolved in the future include whether the balance between matter and dark energy is exact, thereby requiring the curvature of the universe to be a fundamental parameter of cosmology. Measurement of a possible deviation in $\Omega - 1$ from zero would have immense significance. This would distinguish an infinite universe from a finite universe, at least for the simplest topology. In fact, there is a fundamental lower bound on how small a deviation of the curvature from unity could ever be measured, of about 10^{-5}, due to the presence of fluctuations on large scales.

Hitherto, measurements of dark energy have found that the cosmological constant provides a satisfactory interpretation. We observe $w \approx -1$ to a redshift of unity with a precision of at least 10%. This of course is not good enough, as deviations from $w = -1$ as well as an epoch dependence could still be present. Future experiments are designed to measure w with a precision of 2 or 3%. This is about as well as one could ever hope to do, given the systematic uncertainties inherent in our measures of the galaxy population.

Seeking deviations of the primordial power spectrum index n from unity is another important goal. In single field models of inflation, such a deviation is simply connected to T/S, the gravity wave signal from inflation. Confirmation of a deviation of n from unity would strengthen the case for a space experiment designed to measure the tensor mode at low ℓ.

Perhaps our greatest challenge is to confirm that inflation occurred. It is an elusive goal, as the hundreds of alternative inflationary potentials can explain just about any observational result. One has to decide on the level of precision required for future experiments. The combination of measuring curvature and n, inferred at present to be approximately 0 and 0.97 respectively, to high precision would provide a strong case. However measurement of T/S is the only direct signature that would provide an iron-clad argument. Unfortunately, for the inflation advocates predictions of T/S can fall below any foreseeable experimental sensitivity.

The CMB is destined to continue to play a central role in cosmology. An important message to the consumer is that this role is not stand-alone and is greatly strengthened via coordination with large-scale structure mapping carried out with galaxy redshift and lensing surveys.

References

[1] Doroshkevich A G and Novikov I D 1964 *Dokl. Akad. Nauk USSR* **154** 809–11
[2] Sachs R K and Wolfe A M 1967 *Astrophys. J.* **147** 73–90
[3] Silk J 1967 *Nature* **215** 115
[4] Silk J 1968 *Astrophys. J.* **151** 459
[5] Vittorio N and Silk J 1984 *Astrophys. J. Lett.* **285** 39
[6] Bond J R and Efstathiou G 1984 *Astrophys. J. Lett.* **285** 45

[7] Lahav O and Liddle A 2006 Particle Data Book *Preprint* astro-ph/0601168 (*J. Phys.* G33)
[8] Spergel D *et al* 2006 *Preprint* astro-ph/0603449 (*Astrophys. J.* at press)
[9] Riess A *et al* 2007 *Preprint* astro-ph/0611572 (*Astrophys. J.* at press)
[10] Land K and Magueijo J 2006 *Preprint* astro-ph/0602214
[11] Mathis H, Diego J and Silk J 2004 *Mon. Not. R. Astron. Soc.* **353** 681
[12] Niarchou A and Jaffe A 2006 *Preprint* astro-ph/0602214
[13] Baugh C *et al* 2004 *Mon. Not. R. Astron. Soc.* **351** L44
[14] Bartelmann M, Doran M and Wetterich C 2006 *Astron. Astrophys.* **454** 27
[15] Bouwens R and Illingworth G 2006 *Nature* **443** 189